Dean&Francis

ISSN 2959-6122

# The Impact of Musical Stimuli on the Interpretation of Facial Expressions: An Eye-Tracking Study

**Minghui Luo**

The Hill School (860 Beech Street, Pottstown, PA, US 19464)

**Abstract:**

Facial recognition are a cornerstone of human social interaction and emotional bonding. People rely on facial recognition to analyze others' emotions, intentions, and social cues. Facial emotions are particularly effective in conveying information in social contexts. It has the benefits of speed, directness, and universality across cultures. Music, as an emerging area of emotional study, has the power to evoke and alter individual emotions which has the potential to affect emotion recognition. This study examines the impact of different types of music on the judgement of facial emotions. The present research adopts a mixed design, with two types of music as the between-subject variable and three types of facial emotions as the within-subject variable. Participants were asked to evaluate facial expressions while listening to different musical pieces, and an eye-tracker recorded their viewing behaviors. The eye-tracking data, serving as the dependent variable, offers quantifiable details of the observation process. Analysis indicates that under sad music, participants are more likely to rate neutral faces positively. While under happy music, they tend to rate neutral faces more negatively. Eye-tracking results reveal a tendency for participants to focus more on the eye region when viewing faces based on increased total fixation duration, fixation counts, and quicker time to first fixation. Sad music prompts faster gaze towards the eyes, with a shorter time to first fixation compared to happy music. Moreover, when viewing negative emotional faces, participants' gaze towards the nose is quicker than when viewing positive faces, regardless of the music types. This study contributes a novel perspective to cross-modal emotion recognition research which highlights the interaction between music and facial expression interpretation process. The results suggest new insights for emotion recognition training in music therapy, clinical psychology, and education.

**Keywords:** Happy Music, Sad Music, Facial Expression, Face, Emotion, Face Recognition, Eye Tracking

ISSN 2959-6122

# 1. Introduction

Facial recognition is a key skill in human social interactions, enabling individuals to identify others and gather information about their identities, emotions, and intentions. This ability has been critical to human evolution, enhancing social interactions and emotional bonds (Bruce & Young, 2015). Facial expressions are a direct way of communicating emotional information, surpassing language and other non-verbal behaviors in speed and intuitiveness (Schupp et al., 2015). In complex social environments, facial expressions are important sources of emotional signals (Elfenbein & Ambady, 2012). Recognizing emotional faces could facilitate assessing the situation and making appropriate responses rapidly (Ekman, 2016). For instance, detecting an angry face can alert individuals to potential threats, while recognizing a happy face can foster social interaction and cooperation (Adolphs, 2010). Emotional face interpretation is also cross-cultural. Since this form of emotional perception could overcome cultural differences (Barrett et al., 2019). Above all, emotional face recognition is a vital skill and is closely linked to individual emotional regulation and social interaction.

Recently, there has been growing interest in how music influences emotional perception and recognition. As a medium for emotional expression and regulation, music can significantly impact an individual's emotional responses to stimuli such as facial expressions. Thoma et al. (2013) suggest that music can significantly affect emotional states, especially in emotional regulation and expression. Hendricks et al. (2015) found that music can influence emotional perception and mental health through emotional regulation, i.e. reducing anxiety and depression symptoms. However, the impact of music on facial expression interpretation is a subject of debate. Studies by Hickok & Heffner (2017) and Müller & Jonason (2015) indicate that positive music can enhance the speed of recognition, especially for positive facial expressions. Research by Schubert & Doležal (2012) and Kallinen & Ravaja (2015) suggest that positive music may hinder the recognition of facial emotions when the emotions conveyed by the music and the facial expression are inconsistent. Kawakami & Sato (2014) note that while pleasant music can facilitate the recognition of happy faces, it may also impede the recognition of negative emotional faces. Therefore, it is crucial to explore the effects of music on emotional face recognition.

Previous research on music and emotional face recognition has typically focused on traditional measurements such as emotional assessment and reaction time, which rely on participants' self-reports or emotional reactions to music. These approaches are susceptible to subjective biases from participants. Emotional expression is often complex and multidimensional, making it difficult for a single rating to accurately capture the diversity and depth of emotions (Brattico et al., 2013). Additionally, traditional research methods have overlooked physiological changes or neural responses that may be triggered by emotional experiences (Koelsch, 2010). These physiological indicators can provide deeper insights into emotional reactions. This study employs eye-tracking data as a dependent variable. Since eye gaze data can reflect an individual's attention distribution when viewing emotional faces. It offers more detailed insights of the cognitive processes. The results help understand how attention and emotional regulation mechanisms change during emotional face recognition in different musical contexts. This study designs a mixed experiment to explore the impact of music on individuals' emotional face recognition patterns. Music type and facial emotion type are independent variables. Participants' eye-tracking data and ratings of emotional expression are dependent variables This research is expected to offer a new perspective for cross-modal studies on emotion recognition, enriching the theoretical framework of emotional perception. The findings could also provide new approaches for emotional recognition training of music therapy.

# 2. Method

## 2.1 Participants

This study recruited 20 individuals (mean age = 28.7, standard deviation = 5.89) randomly in Shanghai, China. The sample consisted of 11 males and 9 females. Participants were randomly assigned to one of the two groups: Group A (happy music) and Group B (sad music). They were informed about the general purpose and procedures of the experiment. And all subjects provided their consent by signing the consent form.

## 2.2 Stimuli

During the trial, one piece of sad music and one piece of cheerful music, both without lyrics, were utilized as auditory stimuli. Participants assigned to the happy group were exposed to the cheerful music, while those in the sad group listened to the sad music. Visual stimuli consisted of twelve human facial expressions sourced from the database of the Institute of Psychology at the Chinese Academy of Sciences. The selection included four positive expressions, four negative expressions, and four neutral emotional faces. The emotional face images were evenly divided between males and females. Each expression category featuring two females and two males.

### 2.3 Design and Procedure

This study used a mixed-design approach. Music type serves as the between-group variable and emotional facial expressions as the within-group variable. Participants were required to read and sign an informed consent form before taking part in the study. They were then seated in front of an eye tracker (Tobii Pro 4C) connected to a computer screen. Group A participants wore headphones and listened to happy music with unrelated images displayed on the screen for 40 seconds to immerse themselves in the auditory environment. Following this, an introductory page was shown. Group B participants experienced a similar procedure, with the music changed to a sad piece.

All participants underwent a calibration process, tracking a moving dot on the screen to align the eye tracker with their gaze. After successful calibration, participants were notified via on-screen instructions. They were instructed to keep their eyes on the screen throughout the trial. A series of 12 faces, depicting positive, negative, and neutral expressions, were randomly presented on the screen. The duration for each image was 6 seconds. After viewing each image, participants rated the facial expressions on a 9-point scale. 1 is the most negative. 5 is neutral. And 9 is the most positive.

The experiment lasted approximately five minutes, after which participants received a gift. The area of interest (AOI) on the images was divided into eyes (E), nose (N), mouth (M), and full face (F). Eye-tracking parameters such as total fixation duration (TFD), fixation counts (FC), and time to first fixation (TFF) were collected to analyze participants' eye movements in response to the stimuli. Additionally, participants' ratings of the facial expressions were recorded to assess their subjective perception of the emotions displayed.

### 2.4 Data Analysis

To assess the impact of music type on participants' emotion recognition, between-group t-tests were conducted to compare the ratings of emotional faces and eye gaze behaviors within each area of interest (AOI) between Group A and Group B. Additionally, the influence of emotional face type on expression recognition was examined through within-group t-tests on the eye-tracking data within the AOIs for both groups. These analyses allowed for a detailed evaluation of how different music types may modulate the recognition of emotional facial expressions and the associated eye movement patterns.

## 3. Result

### 3.1 Between-group T-test analysis of Ratings for Group A and Group B

The results are listed in Table 1. The between-group comparison of ratings for negative and positive facial expressions did not reveal significant difference between Group A and Group B. However, when rating neutral expressions, Group A ($M = 4.63, SD = 0.53$) rated these expressions significantly more negative than Group B ($M = 5.13, SD = 0.53; p < 0.05$). This indicates that the type of music may have influenced the perception of neutral expressions.

### 3.2 Between-group ANOVA and T-test analysis of TFD for Group A and Group B

The results of the ANOVA and T-test for Total Fixation Duration (TFD) and Fixation Counts (FC) are presented in Table 2 and Table 3 respectively. According to the one-way ANOVA tests, when observing facial images, all participants allocated more fixation duration and counts to the eye areas ($TFD: M = 1.88, SD = 0.94; FC: M = 8.14, SD = 3.63$) than to other facial components ($p < 0.05$). Additionally, T-test results indicated that Group A participants ($M = 1.49, SD = 0.89$) spent a significantly lower amount of time looking at the eye areas ($p < 0.05$) compared to Group B participants ($M = 2.23, SD = 0.89$). A comparison of the facial images revealed that Group A subjects ($M = 1.61, SD = 0.81$) allocated less time ($p < 0.05$) to the eyes than Group B subjects ($M = 2.31, SD = 0.91$) when observing the negative facial images.

### 3.3 Between-group ANOVA and T-test analysis of TFF for Group A and Group B

As presented in Table 4, participants' Time to First Fixation (TFF) for the eye region ($M = 1.03, SD = 1.22$) was significantly shorter ($p < 0.05$) compared to other facial regions when viewing facial images. This finding is consistent with the ANOVA tests for Group A ($M = 1.08, SD = 1.45$). A between-group t-test revealed that the TFF for the nose region in positive images (M = 2.48, SD = 1.10) was significantly longer ($p < 0.05$) than that in negative images ($M = 1.49, SD = 1.01$). Further analysis indicated that for both Group A and Group B, the TFF for the nose region was longer in positive images than in negative images. This suggests that emotions of the images influenced the initial allocation of attention, particularly to the nose region, during the early stages of facial expression processing.

ISSN 2959-6122

**Table 1: Rating Results of Group A and Group B**

| | Total | Positive Expression | Negative Expression | Neutral Expression |
|---|---|---|---|---|
| Total | 5.05 | 7.00 | 3.26 | 4.90 |
| Group A | 4.95 | 6.69 | 3.53 | 4.63 |
| Group B | 5.15 | 7.28 | 3.03 | 5.14 |

**Table 2: TFD(s) for Group A and Group B**

| | Total | | | | Group A | | | | Group B | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | F | E | M | N | F | E | M | N | F | E | M |
| Total | 0.58 | 3.52 | 1.88 | 0.77 | 0.61 | 3.13 | 1.49 | 0.77 | 0.56 | 3.86 | 2.23 | 0.78 |
| Positive Expression | 0.47 | 3.51 | 1.82 | 0.88 | 0.50 | 3.25 | 1.52 | 0.91 | 0.45 | 3.75 | 2.09 | 0.85 |
| Negative Expression | 0.58 | 3.52 | 1.98 | 0.73 | 0.62 | 3.07 | 1.61 | 0.64 | 0.55 | 3.92 | 2.31 | 0.81 |

**Table 3: FC(freq) for Group A and Group B**

| | Total | | | | Group A | | | | Group B | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | F | E | M | N | F | E | M | N | F | E | M |
| Total | 3.08 | 15.71 | 8.14 | 3.18 | 3.52 | 15.94 | 7.53 | 3.67 | 2.69 | 15.49 | 8.69 | 2.74 |
| Positive Expression | 2.56 | 15.84 | 8.01 | 3.72 | 2.97 | 16.63 | 7.72 | 4.47 | 2.19 | 15.14 | 8.28 | 3.06 |
| Negative Expression | 3.13 | 15.57 | 8.37 | 3.01 | 3.53 | 15.56 | 7.84 | 3.09 | 2.78 | 15.58 | 8.83 | 2.94 |

**Table 4: TFF(s) for Group A and Group B**

| | Total | | | | Group A | | | | Group B | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | F | E | M | N | F | E | M | N | F | E | M |
| Total | 1.92 | 0.29 | 1.03 | 1.62 | 1.69 | 0.37 | 1.22 | 1.25 | 2.11 | 0.22 | 0.87 | 1.95 |
| Positive Images | 2.48 | 0.28 | 1.08 | 1.57 | 2.23 | 0.23 | 1.09 | 1.22 | 2.70 | 0.32 | 1.07 | 1.87 |
| Negative Images | 1.49 | 0.32 | 0.98 | 1.47 | 1.19 | 0.40 | 1.09 | 1.32 | 1.76 | 0.24 | 0.89 | 1.60 |

## 4. Discussion

This study explores the impact of music on the interpretation of facial expressions. The experiment utilized music type as the independent variable, dividing participants into two groups: a happy music group (Group A) and a sad music group (Group B). Participants viewed facial expression images while listening to the corresponding music. They were also asked to rate the facial expressions on a scale from 1 (most negative) to 9 (most positive). The viewing process were tracked via an eye-tracker throughout the trial. Data analysis indicated that Group A participants rated neutral faces more negatively, whereas Group B participants rated them more positively. Eye tracking analysis revealed that when viewing emotional faces, participants allocated more Total Fixation Duration (TFD) and Fixation Counts (FC) to the eye region than to other facial regions. Time to First Fixation (TFF) was also shorter for the eye region. Moreover, Group A showed a significantly smaller TFD for the eye region when observing negative faces compared to Group B. Additionally, the TFF for the nose region was significantly shorter for negative emotional faces than for positive ones.

During the facial expression rating task, Group A, listening to happy music, provided more negative ratings. Group B, listening to sad music, provided more positive ratings. This phenomenon can be explained by the mood-congruent judgment effect (Forgas, 2007), where an individual's emotional state influences their interpretation of surrounding stimuli. People tend to set their own emotional status as the reference for neutral status. In this trial, affected by the music, the mood of participants in Group A was more positive. Hence their standard for neutral expression was closer to the positive side. Vice versa, the standard for neutral expression of Group B was closer to the negative side. On the other hand, Bradley (2001) suggests that under strong emotional states, individuals are prone to an emotion amplification effect when processing neutral faces. When exposed to pleasant music, participants' elevated emotional states may intensify the negative interpretation of neutral faces. In contrast, when exposed to sad music, participants' negative emotions may lead them to rate neutral faces more positively, possibly as a means to seek more positive stimuli to mitigate their sadness. For faces with clear positive or negative expressions, the emotional cues are explicit enough, requiring no additional emotional inference. Consequently, there is no significant difference in ratings between the two groups for positive and negative faces.

According to the highest FC and TFD on the eye region, all participants focused more on the eyes when viewing faces. This preference is attributed to the eyes' critical role in recognizing emotional faces (Schultz, 2005). Emotional information is largely communicated via the eyes, especially concerning the intensity and subtlety of emotions. Research has also shown that the eye region has a preferential processing effect for emotional recognition (Leppänen, 2004), which is significant in both emotional judgments and interpersonal communication. The eyes, being the most informative part of the face, provide crucial cues for social cognition, including emotion recognition, emotional bonding, and social interaction. Though other facial areas also contain emotional cues, the eyes are typically the primary focus of attention due to the immediate and clear emotional information they could convey.

Group B participants exhibited a higher Total Fixation Duration (TFD) on the eyes compared to Group A participants. The potential cause is sad music induces stronger negative emotions, leading participants to engage in more detailed processing of emotional information, especially facial expression nuances (Batty, 2006). The increased TFD in the eye region suggests that Group B participants processed emotional information with greater depth. Emotional states influence not only attention allocation but also the distribution of cognitive resources (Vuilleumier, 2004). In a sad emotional state, individuals are more sensitive to negative information and may process the details of facial expressions more carefully, particularly in the eye region, which is rich in emotional cues.

The within-subject t-test shows that participants directed their attention to the nose region more quickly when viewing negative images. This phenomenon could be linked to the increased alertness and threat detection mechanisms that are activated in response to negative emotions (Schupp, 2012). Negative facial expressions, such as anger, fear, and sadness, often trigger the threat detection system, prompting rapid attention to potential dangers. In the interpretation of facial expressions, the nose region is typically a key area for processing threat-related information (Huang, 2011). It can be a significant cue of negative emotional expressions. Consequently, during the initial stages of emotional recognition, the nose region can elicit quicker responses due to subtle shifts in emotional states (Gupta, 2015). This early focus on the nose area in response to negative emotional faces mirrors an individual's rapid alertness and defense mechanism when confronted with potential threats.

The current study has limitations that should be considered in future research. With 20 participants in a mixed-design experiment, the sample size is within an acceptable range. However, to further increase the external validity and generalizability of the results, future studies should aim to expand the sample size and recruit participants. Moreover, this study was limited to 12 trials, which may constrain the outcomes. Future research should aim to increase the number of experimental materials and trials to achieve more precise findings. Last but not least, this study used facial images as visual stimuli. Future studies could utilize real people to replace images to more closely simulate social interactions.

## 5. Conclusion

This study investigates the influence of music on individuals' emotional expression recognition. It uses a mixed experimental design. Music type and facial expressions serve as independent variables, and participants' eye tracking data and ratings for expressions serve as dependent variables. Data analysis indicates that under sad music conditions, participants provided more positive ratings for neutral faces, while happy music resulted in more negative ratings Analysis of eye tracking data reveals that participants, under the influence of sad music, directed their attention to the eye region of faces earlier than when listening to happy music. Moreover, the gaze on the nose region was quicker for negative emotional faces than for positive ones. The study also confirmed the important role of the eyes in emotional face recognition. Regardless of music type or facial expression, the eye region was the primary and most frequently attended area by participants. This research contributes a novel perspective to cross-modal emotion recognition studies, highlighting the interaction between music and emotional faces interpretation. It expands the theoretical framework of emotional

ISSN 2959-6122

perception and emotional face recognition. Furthermore, the findings regarding the impact of music on emotional face recognition patterns suggest that targeted musical interventions could enhance individuals' capacity to recognize and comprehend others' emotions. These insights may provide innovative strategies for emotional recognition training in clinical psychology and education, potentially assisting individuals with emotional or social challenges in accurately identifying others' emotions.

## References

Adolphs, R. (2010). The social brain: Neural bases of social knowledge. Annual Review of Psychology, 60, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Barrett, L. F., Mesquita, B., & Gendron, M. (2019). The experience of emotion in the face: Implications for emotion recognition. Psychological Science, 30(1), 22-30.

Batty, M., & Taylor, M. J. (2006). Early processing of the six basic facial emotional expressions. Cognition, 100(1), 71-92.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions to aversive and rewarding stimuli. Psychological Science, 12(5), 417–423. https://doi.org/10.1111/1467-9280.00276

Brattico, E., Bogert, B., Tervaniemi, M., & Näätänen, R. (2013). Music and emotion: From the auditory perception to emotional response. Psychology of Music, 41(4), 528-545. https://doi.org/10.1177/0305735611411888

Bruce, V., & Young, A. W. (2015). Face perception. Psychology Press.

Ekman, P. (2016). Emotions revealed: Recognizing faces and feelings to improve communication and emotional life (2nd ed.). St. Martin's Press.

Elfenbein, H. A., & Ambady, N. (2012). On the universality and cultural specificity of emotion recognition: A meta-analysis. Psychological Bulletin, 138(2), 200-235.

Forgas, J. P. (2007).The role of emotion in social judgments: An introductory review. European Journal of Social Psychology, 37(1), 1-23. https://doi.org/10.1002/ejsp.336

Gupta, R., & Voyer, D. (2015). Early attention to emotional facial expressions: Evidence for enhanced processing of negative emotions. Journal of Experimental Psychology: Human Perception and Performance, 41(3), 580–594.

Hendricks, M., Williams, S., & Lanza, S. (2015). Music and mood regulation in clinical populations: An investigation of its effects on depression and anxiety. Journal of Music Therapy, 52(2), 212-230. https://doi.org/10.1093/jmt/thu057

Hickok, S. A., & Heffner, C. L. (2017). Music and emotion recognition: The influence of musical training and context on facial expression recognition. Journal of Experimental Psychology: General, 146(2), 164–176.

Huang, Y., et al. (2011). Visual attention to emotional facial expressions: Evidence from the eye movement study. Psychological Science, 22(9), 1089-1097.

Kallinen, K., & Ravaja, N. (2015). The effect of background music on emotional facial expression recognition in complex auditory environments. Psychology of Music, 43(5), 611-627.

Kawakami, S., & Sato, M. (2014). Music-induced positive emotions impair recognition of facial expressions of anger. Cognition & Emotion, 28(2), 206-214.

Koelsch, S. (2010). Towards a neural basis of music perception. Trends in Cognitive Sciences, 14(11), 556-564. https://doi.org/10.1016/j.tics.2010.09.001

Leppänen, J. M., & Hietanen, J. K. (2004).Emotion and face perception: Fast and slow spatial frequency cues. Emotion, 4(2), 156-174.https://doi.org/10.1037/1528-3542.4.2.156

Müller, R., & Jonason, P. K. (2015). The effects of happy music on facial emotion recognition and the role of arousal. Psychology of Music, 43(4), 498-510.

Schubert, E., & Doležal, A. (2012). Can music enhance recognition of emotional facial expressions? The role of familiarity and congruence. Music Perception, 30(1), 39-50.

Schultz, W. L., et al. (2005).Affective modulation of the early visual evoked potentials to faces and objects. Emotion, 5(3), 347-358.https://doi.org/10.1037/1528-3542.5.3.347

Schupp, H. T., et al. (2012). Neural circuits of emotion: A review of the literature on emotion and attention. Neuroscience & Biobehavioral Reviews, 36(8), 2093-2106.

Schupp, H. T., et al. (2015). Emotion and attention: A special issue on the interplay of emotion and attention. Psychophysiology, 52(4), 419-426. https://doi.org/10.1111/psyp.12312

Thoma, M. V., La Marca, R., Brönnimann, R., Finkel, L., & Ehlert, U. (2013). The effect of music on the human stress response. Psychoneuroendocrinology, 38(11), 2843-2851. https://doi.org/10.1016/j.psyneuen.2013.06.009

Vuilleumier, P., et al. (2004). Neural mechanisms of emotional attention. Current Opinion in Neurobiology, 14(2), 215-222.