Emotion-Adaptive Music Interfaces: A Comparative Evaluation of Real-Time Facial Expression Recognition versus Manual Selection

Mohammadreza Dolati, Amin Saeidi Kelishami
Human-Computer Interaction Course - Grenoble INP – UGA

January 15, 2026

Abstract

Current music streaming platforms rely on manual mood selection or history-based recommendations that do not adapt to real-time emotional states. We developed an emotion-adaptive interface using webcam-based facial expression recognition and compared it against traditional manual selection in a within-subjects study. Participants performed focused work tasks while using both interfaces. Results revealed that the adaptive interface significantly reduced user interaction burden while maintaining comparable music-emotion fit and usability. The adaptive system automatically selected music based on detected emotional states, substantially decreasing the need for manual browsing and selection. Both conditions achieved similar subjective ratings for emotion-fit and usability, suggesting that automation does not compromise user experience. These findings demonstrate the feasibility and potential benefits of emotion-adaptive music interfaces for reducing interaction effort during ambient music consumption.

1 Introduction

Music consumption has shifted from active listening to background streaming during work and daily activities. Major platforms like Spotify and Apple Music use recommendations based on listening history or require manual mood playlist selection. Neither approach adapts to users' current emotional states, requiring manual intervention that interrupts workflow.

Recent advances in affective computing enable reliable real-time emotion recognition through facial expressions using standard webcams. However, few studies evaluate whether integrating this capability into music interfaces actually improves user experience. This study addresses that gap by comparing traditional manual selection against automatic emotion-adaptive music playback.

Research Question: How does automatic emotion-based music playback compare to direct music selection in terms of interaction effort, music-emotion fit, and usability?

We developed a working prototype and conducted a within-subjects experiment where participants used both interfaces during focused work sessions. Our contributions include an empirical comparison of manual versus emotion-adaptive music selection, quantitative evidence on interaction patterns, usability, and music-emotion fit, and design insights for future ambient music interfaces.

2 Related Work

Recent systems demonstrate reliable facial emotion detection using CNNs trained on datasets like FER2013 [1]. Kalambate et al. [2] developed Emotify, a real-time emotion-based music player, while Bakariya and Singh [3] created a complete system detecting faces, identifying emotions, and recommending music in real-time. Guthula et al. [4] and Bottu et al. [5] further demonstrated technical feasibility with facial detection systems integrated with music recommendation.

Several approaches map emotions to music selections. Chavali and Menezes [6] used transformer embeddings to map text to valence-arousal space. Parag et al. [7] integrated facial analysis with NLP using Spotify metadata. Zhao et al. [8] introduced FCRA networks combining convolutional, recurrent, and attention layers. Multimodal approaches using EEG and physiological sensors [9] achieve 93-96% accuracy but require specialized hardware.

While emotion recognition technology is mature, research focuses primarily on algorithmic performance rather than user experience. Commercial platforms like Spotify and Apple Music offer mood playlists but require manual selection. Experimental systems like Moodify exist but lack rigorous usability evaluation. Our study provides empirical evidence comparing manual and emotion-adaptive interfaces across interaction effort, usability, and music-emotion fit; a practical, user-centered evaluation needed to advance the field.

3 Methodology

This section describes our experimental system, study design, and data collection procedures. We aimed to create a realistic comparison between manual and emotion-adaptive music selection while maintaining rigorous experimental control.

3.1 System Design

We developed a web-based platform with Flask backend, music recommender module, comprehensive interaction logger, and browser-based emotion detection (face-api.js). Both conditions use the same dataset of 4,830 Spotify tracks with audio feature metadata (valence, energy, tempo, danceability, acousticness). Figure 1 illustrates the system architecture and data flow for both conditions.

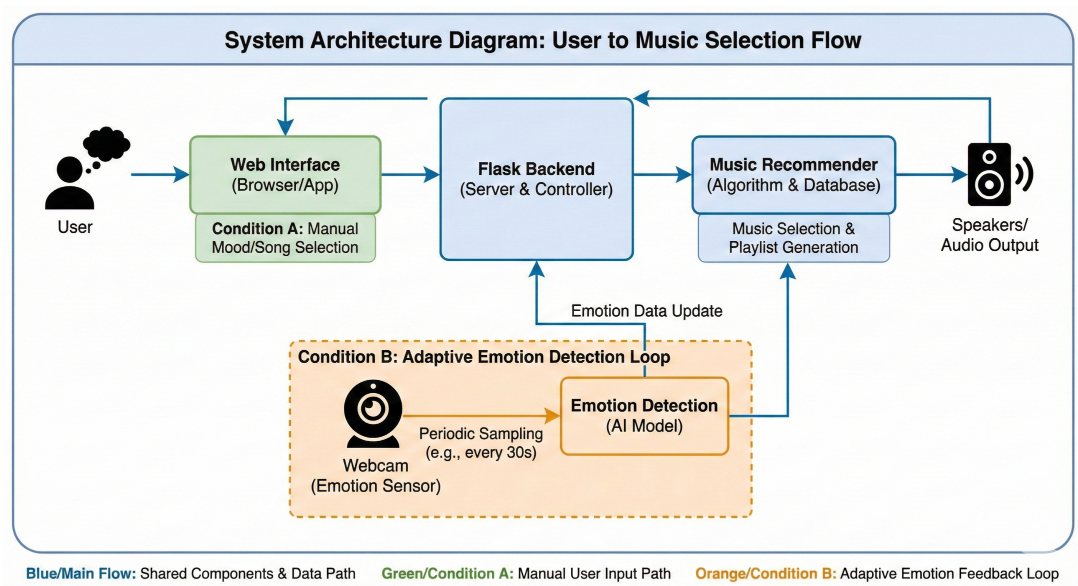

Figure 1: System architecture showing the flow from user interaction to music selection. In Condition A (Manual), users directly select moods and songs. In Condition B (Adaptive), the system periodically detects emotions via webcam and automatically updates music selection.

The system implements two experimental conditions (see Appendix A for detailed interface screenshots). *Condition A (Manual)* provides a traditional interface with six mood buttons (Happy, Sad, Energetic, Calm, Focused, Melancholic). Users browse and select songs with full control over searches, skips, and mood changes. *Condition B (Adaptive)* uses webcam-based facial expression detection (face-api.js) running every 30 seconds, identifying seven emotions (angry, disgust, fear, happy, sad, surprise, neutral). The system automatically selects and plays matching music while users retain manual override capability for agency. Both conditions share the same visual layout but differ in interaction paradigm: the adaptive condition displays the currently detected emotion and updates music automatically, whereas the manual condition relies on user selection.
A critical design challenge was mapping detected emotions to appropriate music selections. We used Spotify's audio features (valence, energy, tempo) as our primary mapping mechanism. Complete mapping specifications linking detected emotions to audio feature ranges are provided in Appendix B.

3.2 Study Protocol

We employed a within-subjects design with counterbalanced order. Each participant (N=10 from the university community, all regular music streaming users) experienced both conditions for 10 minutes each, separated by three brief questionnaires (3-4 minutes total). Total session time was 25-30 minutes. While the small sample size limits definitive conclusions, comprehensive logging provides valuable preliminary insights.

Participants performed light cognitive tasks (reading articles, browsing content, working on assignments) while music played in the background. The protocol consisted of: (1) Introduction and consent (5 min), (2) First condition with tasks (10 min), (3) Three questionnaires: SUS, Music-Emotion Fit, and Effort/Experience (3-4 min), (4) Second condition (10 min), (5) Same three questionnaires (3-4 min), and (6) Debrief (3-5 min). All interactions were automatically logged with millisecond timestamps.

3.3 Data Collection and Analysis

We evaluated three hypotheses: (H1) the adaptive interface would require fewer user interactions than manual selection, measured by logging all clicks, searches, and song changes; (H2) the adaptive interface would provide better music-emotion fit, measured via a 5-item Likert scale (1-5) assessing music-mood match averaged into a composite score; and (H3) the adaptive interface would achieve comparable or higher usability, measured using the System Usability Scale (SUS) [10], a validated 10-item questionnaire producing scores from 0-100 (acceptable> 68). All interactions were logged with millisecond timestamps. Given N=10, we used non-parametric Wilcoxon signed-rank tests, reporting medians with interquartile ranges (IQR) and effect sizes (r = $Z/\sqrt{N}$). Significance level was set at α=0.05.

4 Results

We present results for all three outcome measures: interaction count, music-emotion fit ratings, and system usability scores (Table 1 and Figure 2).

Measure	**Manual Mdn±IQR**	**Adaptive Mdn±IQR**	**Difference (Mdn)**
Interaction Count	11.5±7.5	5.5±5.2	-6.0
Emotion-Fit (1-5)	3.60±0.70	3.30±1.10	-0.30
SUS Score (0-100)	53.8±31.2	55.0±42.5	+1.2

Table 1: Descriptive statistics for interaction count, music-emotion fit ratings, and System Usability Scale scores. Mdn = Median, IQR = Interquartile Range.

- *H1: Interaction Effort.* The adaptive interface showed significantly fewer interactions (Manual: Mdn=11.5, Adaptive: Mdn=5.5, 52% reduction). Wilcoxon test: W=50, p=0.010, r=0.738 (large effect). This finding supports our hypothesis with statistical significance and a large effect size. The median reduction of 6 interactions represents meaningful practical benefit during a 10-minute session.

- *H2: Music-Emotion Fit.* Manual rated slightly higher (Mdn=3.60) than Adaptive (Mdn=3.30), a modest difference. Wilcoxon test: W=26, p=0.615, r=0.093 (negligible effect), not significant. Both scored in the neutral-to-slightly-positive range (3.3-3.6 of 5), suggesting neither interface excelled at emotion matching but both provided acceptable music-emotion fit.

- *H3: Usability.* Adaptive scored slightly higher (Mdn=55.0) than Manual (Mdn=53.8), with no significant difference (W=10, p=0.231, r=0.379, medium effect). Both scored below the acceptable threshold (68), indicating substantial room for improvement. However, comparable scores suggest automation does not inherently harm usability, and low scores likely reflect prototype limitations rather than fundamental design flaws.

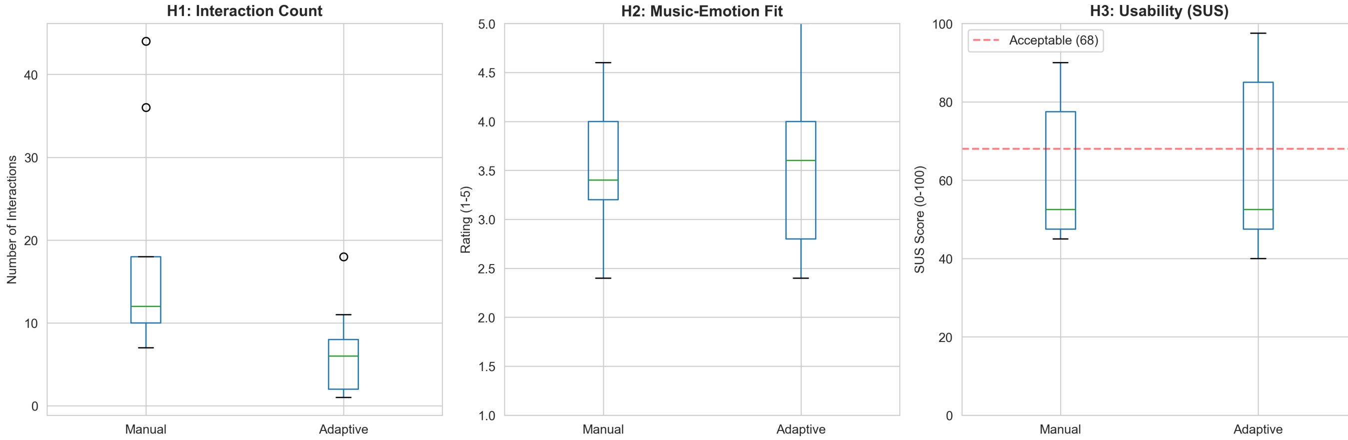

Figure 2: Summary of all three dependent measures. Left: Interaction effort (H1) shows largest effect despite non-significant p-value. Middle: Emotion-fit ratings (H2) show slight advantage for Adaptive. Right: SUS scores (H3) reveal both conditions fell below acceptable usability threshold.

Order effects analysis revealed no significant differences between order groups, suggesting effective counterbalancing. When decomposing interactions by type, the reduction in the Adaptive condition was primarily driven by fewer mood selections (8.0 vs. 0.0), while skip actions remained similar (3.0 vs. 2.0).
Substantial individual variability was observed in how participants used both interfaces. Some Manual users employed a "set and forget" strategy, while others browsed actively throughout the session. In the Adaptive condition, one participant frequently overrode the system's selections, explaining the high IQR in interaction counts.
Table 2 summarizes results. Adaptive showed substantial interaction reduction (H1, trending), comparable emotion-fit (H2), and equivalent low usability (H3).

Hypothesis	**W**	**p**	**r**	**Result**
H1: Fewer interactions	50	0.010	0.738	**Yes**
H2: Better fit	26	0.615	0.093	No
H3: Comparable SUS	10	0.231	0.379	Yes

Table 2: Hypothesis testing summary with Wilcoxon signed-rank test statistics. W = Wilcoxon test statistic (sum of positive ranks), p = significance level, r = effect size (small: 0.1, medium: 0.3, large: 0.5).

5 Discussion

5.1 Key Findings

The adaptive interface achieved a 52% reduction in user interactions (11.5 vs. 5.5 median interactions per 10-minute session), representing statistically significant and practically meaningful benefit with a large effect size, projecting to approximately 72 fewer user actions over 2-hour sessions.
Emotion-fit scores remained in the neutral-to-slightly-positive range for both conditions, with no significant difference, though Manual scored slightly higher than Adaptive, possibly reflecting users' explicit awareness of their own mood preferences. Possible explanations for the modest scores include insufficient detection accuracy, relatively simple emotion-to-music mapping rules, stable emotions during short work sessions, or genuine difficulty in matching music to subtle emotional states.
Both interfaces achieved similar usability scores, falling below the acceptable range. This indicates room for improvement but also demonstrates that automation does not inherently harm usability compared to manual selection. Low scores likely reflect prototype limitations, unfamiliar music dataset, occasional detection failures, and the experimental nature of the system rather than fundamental design flaws.

5.2 Design Implications

Our findings suggest several design considerations for emotion-adaptive music systems:

- **Hybrid Control Model:** The optimal design combines automatic selection with easy manual override. The system should handle routine music selection while allowing users to intervene when desired. Manual interventions should be treated as valuable feedback rather than system failures, acknowledging that music selection is inherently subjective and context-dependent.
- **Detection Interval Tuning:** Our 30-second emotion detection intervals may have been too frequent for the relatively stable emotions during focused work. Longer intervals (2-3 minutes) or intelligent logic that changes music only upon sustained emotional shifts may feel more natural and less disruptive to the user experience.
- **Transparency and Explainability:** Displaying the detected emotion helped participants understand system behavior and build trust. Future designs should go further by explaining why specific songs were chosen or requesting user confirmation before major mood shifts, particularly when transitioning between highly contrasting emotional states.
- **Personalization Through Learning:** Generic emotion-to-music mappings have limitations. Systems should learn individual preferences over time (e.g., User A prefers indie folk when sad while User B prefers classical piano). Machine learning from user overrides and explicit feedback can dramatically improve matching accuracy.
- **Context-Awareness Beyond Emotion:** The system should consider the user's current activity and goals. Music appropriate for focused work differs substantially from music for exercise, relaxation, or social settings. Integrating calendar data, time of day, and activity recognition could enhance selection quality.

5.3 Limitations & Future Work

Several limitations must be acknowledged. First, our sample size (N=10) achieved statistical power for the primary hypothesis (H1) but remains modest for detecting smaller effects in secondary measures. Second, the 10-minute sessions were too brief to capture the full dynamics of emotion changes during extended work periods; longer sessions (30-60 minutes) or longitudinal studies would better reflect real-world usage. Third, the lab setting lacks ecological validity, emotion detection accuracy was not independently verified, the dataset may not have matched participants' personal preferences, and prototype quality issues likely affected usability scores.

Priorities include: (1) adequately powered replication (N=25-30), (2) longitudinal field studies, (3) multi-modal detection (voice, typing, calendar context), (4) active learning from user overrides, (5) evaluation across contexts (exercise, relaxation, commuting), (6) privacy concern investigation, and (7) hybrid recommendation algorithms combining emotion with collaborative filtering.

6 Conclusion

This study evaluated whether real-time emotion detection improves music interfaces over manual selection through a controlled within-subjects experiment comparing traditional manual mood selection against automatic emotion-adaptive music playback. The adaptive interface significantly reduced user interaction burden (52% reduction) while maintaining comparable music-emotion fit and usability, suggesting meaningful efficiency gains during ambient music consumption.

Results demonstrate that emotion-adaptive interfaces can meaningfully reduce interaction effort without compromising user experience, though optimal designs likely require hybrid control models that combine automatic selection with easy manual override. Our contribution lies in providing empirical evidence that accessible technologies enable feasible emotion-adaptive music systems with measurable user benefits, though substantial design work remains to improve overall usability and personalization. Future work should focus on larger samples, longitudinal field studies, personalization mechanisms that learn individual preferences, and context-aware adaptation beyond emotion alone.

Acknowledgments

We used AI tools to assist with experimental interface design and report writing, primarily for grammar checking, summarizing, and rewriting for consistency. We thank all participants for their time and valuable feedback. We are grateful to our supervisors, Céline Coutrix and Vincent Lambert, for their guidance and support throughout this project.

References

[1] Ian J. Goodfellow, Dumitru Erhan, Pierre L. Carrier, Aaron Courville, Thomas Breuel, Goyette Philipp, James Bergstra, and Yoshua Bengio. Challenges in representation learning: A report on three machine learning contests. *Neural Information Processing Systems (NIPS)*, 2013. FER2013 dataset: 35,887 labeled 48×48 grayscale images of human faces with 7 emotion categories.

[2] Purva Kalambate, Seema Hanchate, and Poonam More. Emotify: Real-time emotion-based music player. *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)*, 14(05), 2025.

[3] Brijesh Bakariya, Arshdeep Singh, Harmanpreet Singh, Pankaj Raju, Rohit Rajpoot, and Krishna Kumar Mohbey. Facial emotion recognition and music recommendation system using cnn-based deep learning techniques. *Evolving Systems*, 15(2):641–658, 2024.

[4] Manvitha Sri Guthula and Monali Bordoloi. Music recommendation system using facial detection based emotion analysis. In *2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, pages 296–301. IEEE, 2024.

[5] V Bottu and K Ragavan. Emotion-based music recommendation system integrating facial expression recognition and lyrics sentiment analysis. *IEEE Access*, PP:1–1, 01 2025.

[6] Apoorva Chavali and Reeve Menezes. Words to waves: Emotion-adaptive music recommendation system. *arXiv preprint arXiv:2510.21724*, 2025.

[7] Pardhi Parag, Deshmukh Sakshi, and Gupta Suman Sen. Emotion based music recommendation system using machine learning and ai. *International Journal of Trend in Scientific Research and Development*, 8(5):329–336, 2024.

[8] Liang Zhao, Guangzhan Liu, Shuailing Yan, and Jing Zhang. Emotion-driven music recommendation system based on fully convolutional recurrent attention networks and collaborative filtering. *Alexandria Engineering Journal*, 125:354–366, 2025.

[9] Advances in multimodal emotion recognition based on brain and heart signals. *Frontiers in Neuroinformatics*, 14:569, September 2020. EEG, ECG, and physiological sensor integration achieving 93–96% emotion classification accuracy.

[10] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[11] Aaron Bangor, Philip T Kortum, and James T Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

A Experimental Interface Details

This section provides detailed visual documentation of both experimental conditions. The interfaces were designed to maintain visual consistency while differing fundamentally in their interaction paradigm.

A.1 Manual Selection Interface (Condition A)

Figure 3 shows the manual selection interface. This condition provides explicit user control through six mood category buttons positioned at the top of the interface: Happy, Sad, Energetic, Calm, Focused, and Melancholic. When a user selects a mood, the system displays a curated list of songs matching that emotional category based on predefined audio feature ranges.

Key interface elements include:

- **Mood Selection Buttons:** Six clearly labeled buttons allowing direct mood category selection
- **Song List Display:** Shows available tracks with metadata (title, artist)
- **Playback Controls:** Standard play/pause, skip, and volume controls
- **Search Functionality:** Text-based search allowing users to find specific songs or artists
- **Current Playing Track:** Displays currently playing song.

Users had complete autonomy to switch moods, skip songs, search for specific tracks, and adjust playback at any time. All interactions (mood selections, song plays, skips, searches) were logged with millisecond-precision timestamps.

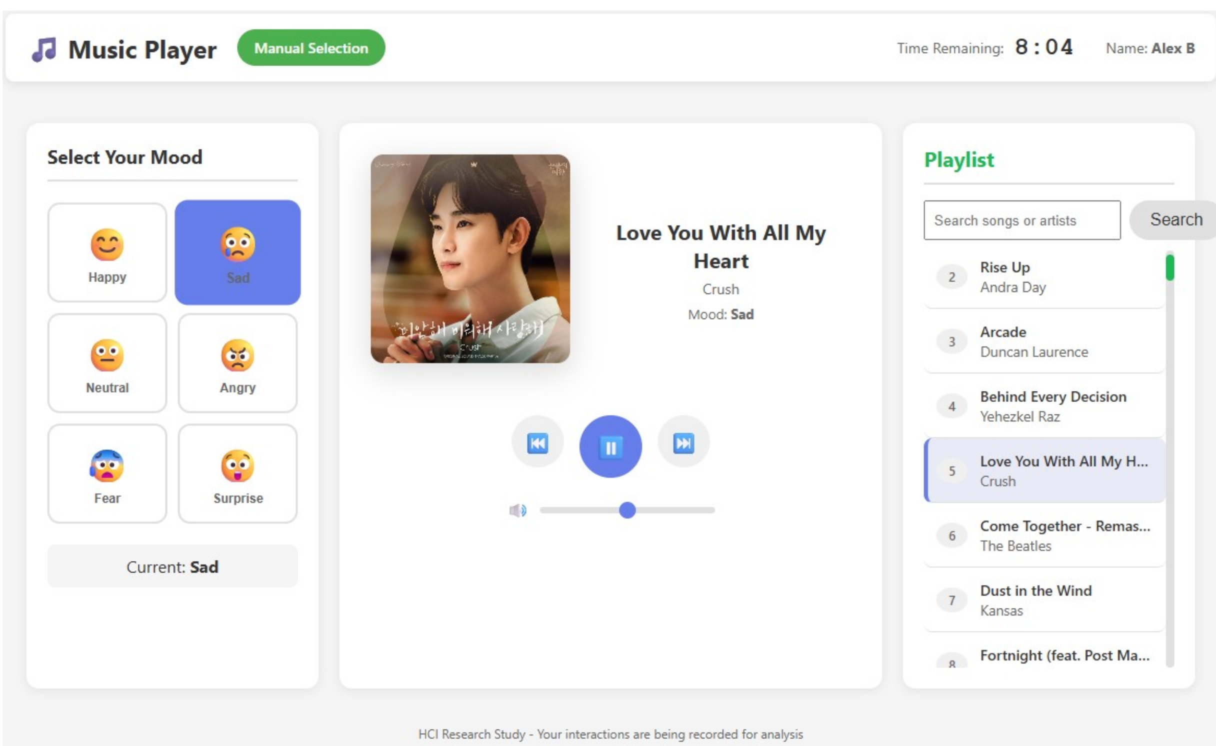

Figure 3: Manual Selection Interface (Condition A). Users explicitly select mood categories via six buttons and browse song lists. The interface provides full control over music selection, allowing mood changes, song skips, and searches at any time during the session.

A.2 Emotion-Adaptive Interface (Condition B)

Figure 4 shows the emotion-adaptive interface. This condition uses real-time facial expression recognition to automatically detect the user's emotional state and select appropriate music without requiring explicit input.
Key interface elements include:

- **Emotion Detection Display:** Shows the currently detected emotion in real-time (e.g., "Happy," "Neutral," "Sad")
- **Detection Confidence Indicator:** Visual feedback showing detection confidence level
- **Automatic Playlist:** System-generated song list based on detected emotion
- **Manual Override Capability:** Users retain ability to manually select different songs (specified by mode in the playlist)
- **Playback Controls:** Same standard controls as manual condition for consistency

The system runs emotion detection every 30 seconds using face-api.js, which analyzes webcam input to identify seven emotions (angry, disgust, fear, happy, sad, surprise, neutral). Upon detecting an emotion, the system automatically queries the music database for songs matching the corresponding audio feature profile (see Appendix B for complete emotion-to-music mappings) and updates the playlist.

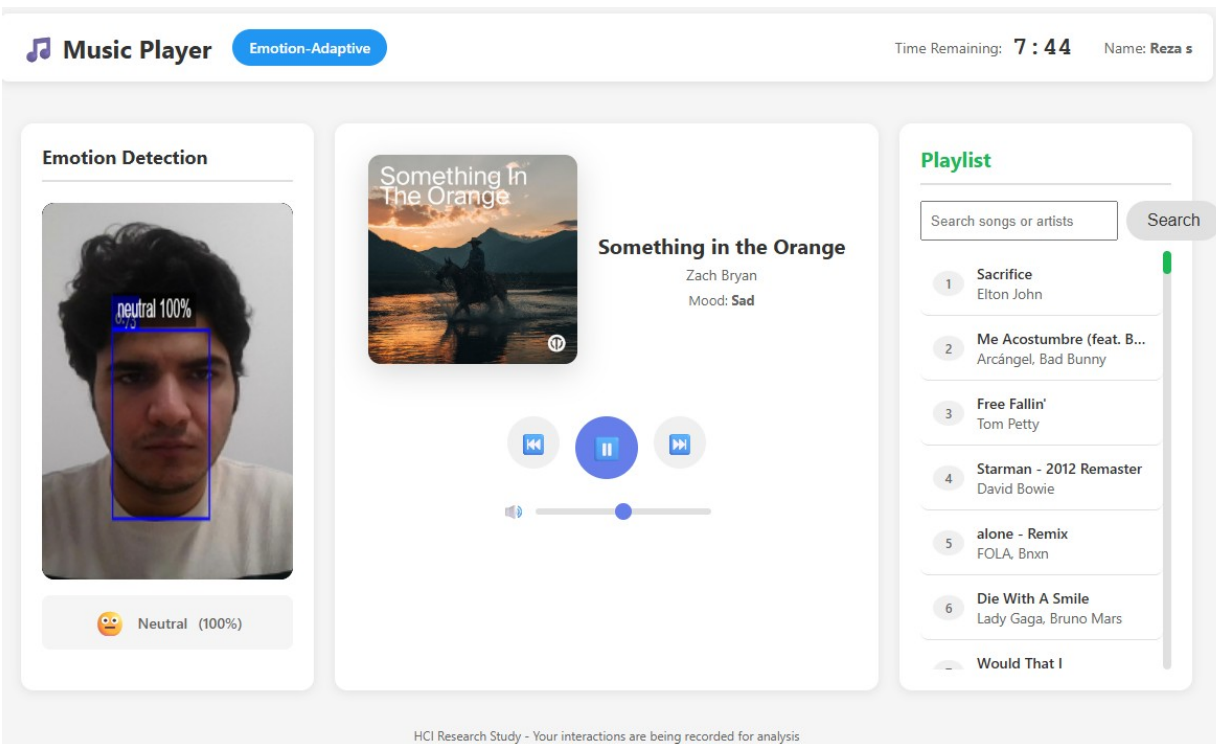

Figure 4: Emotion-Adaptive Interface (Condition B). The system displays the currently detected emotion and automatically updates music selection based on real-time facial expression analysis. Users retain manual override capability but are not required to actively manage music selection.

A.3 Design Rationale

Both interfaces shared identical visual layout, color scheme, and playback controls to isolate the effect of the interaction paradigm (manual vs. automatic) rather than confounding results with different visual designs. The music database, audio feature extraction, and recommendation logic were identical across conditions; only the trigger mechanism differed (user button press vs. automatic emotion detection).
This design ensured that any observed differences in interaction count, emotion-fit ratings, or usability scores could be attributed to the fundamental difference in control paradigm rather than ancillary interface design factors.

B Emotion-to-Music Mapping Details

Table 3 presents the complete emotion-to-music feature mapping used in the adaptive condition. These ranges were determined through pilot testing with 3 users and refined based on their feedback. The system filters songs matching the specified ranges and randomly selects from the top 50 matches to provide variety.

Detected Emotion	Valence Range	Energy Range	Tempo (BPM)	Additional Filters
Happy	> 0.6	> 0.6	> 120	Danceability > 0.5
Sad	< 0.4	< 0.4	< 100	Acousticness > 0.3
Neutral/Calm	0.4–0.6	0.4–0.6	90–120	—
Energetic	> 0.6	> 0.7	> 130	Energy > 0.7
Angry	< 0.5	> 0.7	> 120	Loudness > -5 dB
Surprise	> 0.5	> 0.6	110–140	—
Disgust	< 0.5	0.3–0.6	80–110	Acousticness > 0.4
Fear	< 0.4	0.5–0.7	100–130	Minor mode preferred

Table 3: Complete emotion-to-music feature mapping with all detected emotions and audio feature constraints. Valence represents positivity (0=negative, 1=positive), Energy represents intensity (0=calm, 1=energetic).

C Questionnaire Items

Participants completed **three questionnaires** after each condition: (1) System Usability Scale for hypothesis H3, (2) Music-Emotion Fit for hypothesis H2, and (3) Effort and Experience for exploratory analysis.

C.1 System Usability Scale (SUS)

Participants rated the following 10 items on a 5-point Likert scale (1=Strongly Disagree, 5=Strongly Agree) [10]. The **same questions** were used for both Manual and Adaptive conditions:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

SUS scores were calculated using standard scoring: For odd-numbered items (1, 3, 5, 7, 9), subtract 1 from the user response. For even-numbered items (2, 4, 6, 8, 10), subtract the user response from 5. Sum the converted responses and multiply by 2.5 to obtain a score from 0–100. Scores above 68 are considered acceptable [11].

C.2 Music-Emotion Fit Scale

Participants rated the following 5 items on a 5-point Likert scale (1=Strongly Disagree, 5=Strongly Agree) after each condition. These items were identical for both conditions:

1. The music felt right for how I was feeling.
2. The music choices matched my mood well.
3. I enjoyed the music that was played for me.
4. This system helps me to maintain my mood.
5. I would pick similar music for myself if I was in this mood.

Scores were averaged to create a composite emotion-fit score (range: 1–5).

C.3 Effort and Experience Questions

Participants rated the following 3 items on a 5-point scale (1=Very Low, 5=Very High) after each condition:

1. How easy was it to get music that fit your mood?
2. How much effort did you spend clicking or searching for music?
3. Did you ever feel rushed or stressed while using the music player?

These supplementary items provided additional context on user experience but were not included in the primary hypothesis testing.

D Individual Participant Data

Table 4 presents data for each participant, showing substantial individual variability in interaction patterns and subjective ratings.

	Interactions		**Emotion-Fit**		**SUS**		
ID	**Manual**	**Adaptive**	**Manual**	**Adaptive**	**Manual**	**Adaptive**	**Order**
P1	10	7	4.6	4.0	47.5	45.0	M-A
P2	7	2	3.2	4.4	50.0	57.5	A-M
P3	11	2	2.4	3.6	47.5	47.5	A-M
P4	16	18	4.0	2.8	55.0	40.0	A-M
P5	10	4	3.8	3.0	80.0	92.5	A-M
P6	44	5	3.8	5.0	90.0	92.5	M-A
P7	18	6	3.4	2.8	52.5	52.5	A-M
P8	7	8	4.2	2.4	90.0	97.5	M-A
P9	36	1	3.4	3.8	77.5	85.0	A-M
P10	12	11	2.8	3.0	45.0	50.0	M-A
Mdn	**11.5**	**5.5**	**3.6**	**3.3**	**53.8**	**55.0**	—
IQR	**7.5**	**5.2**	**0.7**	**1.1**	**31.2**	**42.5**	—

Table 4: Individual participant data showing all three dependent measures across N=10 participants. Order indicates which condition was experienced first (M=Manual, A=Adaptive). Eight of ten participants showed fewer interactions in Adaptive condition.

Notable patterns: Eight of ten participants showed fewer interactions in Adaptive condition (exceptions: P4 and P8). P6 and P8 showed highest SUS scores (90+), while P1 and P10 scored lowest. Individual variability highlights importance of personalization. P4 showed more interactions in Adaptive (18) than Manual (16), possibly due to repeated manual overrides of system selections.

E Interaction Logging Structure

All user interactions were logged with millisecond-precision timestamps in JSON format. Each log entry captured:

- **participant_id:** Anonymized participant identifier
- **condition:** "manual" or "adaptive"
- **condition_order:** Which condition came first for this participant
- **condition_order_index:** 1 for first condition, 2 for second condition
- **ts_iso:** ISO 8601 timestamp (e.g., "2026-01-07T21:29:51.460Z")
- **t_ms:** Milliseconds elapsed since condition start
- **event_type:** Type of interaction (e.g., emotion_prediction, playlist_update, song_select, mood_select)
- **actor:** "user" for manual actions, "system" for automatic actions
- **current_emotion:** The emotion state at time of event
- **current_song:** Track name currently playing (or null)
- **details:** Event-specific metadata (e.g., detected emotion, confidence, track info)

This comprehensive logging structure enabled precise calculation of interaction counts, temporal analysis of user behavior patterns, and distinction between user-initiated versus system-initiated actions throughout each 10-minute session.