Face the Music

Amelia Simonoff

DD/MM/2022

A thesis presented for the degree of

Bachelor of Arts in Music

At the

THE UNIVERSITY OF CHICAGO | COLLEGE

1 Abstract

Since time immemorial, music has both elicited and strengthened emotions. Normally, music elicits emotions, from the performer in the listener. However, little has been done going the other direction: Emotions do not usually create music. Emotions are often used during the composition and creation of music, but that is still through and intermediate: the composer. In this project, I wanted to go the other way: What would happen if the emotions of the listener affected the music?

To that end, I use facial recognition technology to read a user's emotions, as current facial recognition technology has progressed to the point that it can be done in nearly real-time. In this project, I use the Python library DeepFace to read peoples' emotions in real-time, and then the Spotify Web API and Web Playback SDK to play music related to those emotions. I set up this project on a Raspberry Pi, continually running a Python script that reads data from the camera and then pulls data from Spotify.

I then set up the Raspberry Pi in the Media Arts, Data, and Design (MADD) Center at the University of Chicago.

2 Background

"To take part in music is central to our very humanness..."

— Christopher Small, *Musicking*

In our lives, music exists in a wide variety of contexts, and there are as many ways to interact with it as there are people. Music is an undertaken action: we create, perceive, and respond to music, and there are as many ways to do so as there are people (Small 1998).

Music, with how incredibly prevalent it is around is in the modern day, may not always elicit emotions (Juslin 2019a, ch. 16). However, music is a language that is able to communicate emotional ideas succinctly, which speech is unable to do (Henson 1977). These are usually basic emotions, such as joy, love, happiness, fear, anger, sorrow, and neutral (Sundberg 1983). The ways in which it does so is affected by affected the current internal state of the listener (Harrer and Harrer 1977).

Patrik N. Juslin has created a model for how music elicits emotions in an individual, known as the BRECVEMA model: Brainstem reflex, Rhythmic entrainment, Evaluative conditioning, Contagion, Visual imagery, Episodic memory, Musical expectancy, and Aesthetic judgement (Juslin 2019a, ch. 17). Juslin describes a continuum between musical sensation (in the brain) and emotional perception (in the conscious mind), which occers from primitive reaction formation in various parts of the brain which affects conscious thought. Each of these mechanisms affects the effected emotion differently: for example, music that stimulates the brainstem (such as sudden loud sounds) is more likely to induce arousal or surprise, while music that triggers episodic memory can induce any possible emotion, but especially nostalgia and longing (Juslin 2019a, ch. 25).

David Huron has also created a model for describing how music elicits certain physical responses in listeners, such as frisson (goosebumps) through the elements of anticipation inherent in continuous music through time (Huron 2006). Known as the ITPRA Theory of Expectation, the Imagination response motivates an organism to do things that increase the likelihood of future beneficial outcomes; the Tension response prepares an organism for an event by changing arousal and attention; the Prediction response provides positive and negative feedback based on outcomes, the Reaction response addresses the worst-case situations by making a protective response; and the Appraisal response providespositive and negative reinforcements. The first two mechanisms occur before the outcome, while the last three mechanisms occur after the outcome is heard. This model also explains the fact about how music listeners report that experiencing negative emotions, such as sadness, anger, or fear, is often pleasurable (Juslin 2019a, ch. 32)

These two models try to both physical and emotional reactions to music. However, there are also other mechanisms with whic hmusic can elicit emotions, such as cognitive goal appraisal.

However, music can only induce emotions over time: one note, out of context, cannot create an emotional state all on its own (Schubert 2001). Also, the emotions that music induces are shaped the society and culture of the listener, which is often overlooked. Music also induces stronger emotions when it can be controlled, such as in private spaces or through busking (Sloboda and O'Neill 2001). Music in private spaces can be more closely controlled and chosen by the listener, and so it is more likely to cause a strong emotional response.

In public spaces, such as in gyms or Muzak in elevators, music is usually thrust upon the listeners, who must tolerate what they are subjected to. However, in large public gatherings, such as crowds at a concert, the musical emotions are at their strongest due to the volume of the music, passion of the performers, and the synergetic emotional responses of the people in the crowd.

There already exists an abundance of research describing how listeners perceive emotions in music, but it is rare that the listener can take an active role in shaping the emotive quality of the sound. In fact, under the current paradigms, the listener would have to themselves be a performer or a composer. I was then interested in creating music based on an emotion, instead of the other way around, where the listener and their emotions are active participants due in the generation of music and emotions.

For that, I needed a way to read the emotions of users in near real-time, and then a way to either synthesize music of a certain emotion or pull from an already-existing database. Modern facial recognition software is a great candidate for accomplishing the first step, and so I used an already-existing software suite to accomplish this (known as a 'pre-trained model').

Facial recognition software, first pioneered in the 1960's, has been available since 2001 for real-time video (Yamaguchi 2012). There has been a lot of progress since, and modern implementations are incredibly lightweight and are available to the public through open source licenses. However, facial recognition is not without its controversies. Worries include privacy violations about a new level of surveillance afforded by this technology, imperfect results for this technology for indiiduals who aren't light-skinned males, and lack of data protection concerns. This project does not record nor save any information, and is only available to the public in one location, bypassing these concerns.

Unfortunately, real-time synthesis of original musical sound is not yet available. While artificial-intelligence based musical synthesis has already passed the Turing Test (AILabs 2018), the models are not quite versatile enough yet to produce a continuous stream of music according to the changing emotions of the user. As such, I used a corpus of recorded music from streaming services, such as Apple Music or Spotify.

The algorithms providing recommendations from a previous song is a million-dollar problem in the music streaming industry: algorithms which provide a satisfying next song to the consumer increase customer retention through increasing satisfaction. This is paramount to increase revenue for these streaming servise. As such, platforms such as Apple Music, Pandora, or Spotify, all have well-researched algorithms for their musical recommendations, which are especially fine-tuned for commercial music (Drott 2018). As such, I chose to use one of those algorithms for obtaining recommendations for my own project by leveraging the metadata of songs in its responses. By feeding the algorithm the genre preferences of the user, their top songs, and their emotion (see Methods), I receive musical suggestions that the user is likely to enjoy which also correspond to their current emotion.

Spotify, the most-used music streaming service in the world (Mulligan 2022), has many playlists with emotional labels, such as "Happy Beats" or "Sad Piano", etc.. However, Spotify does not label their music based on emotion outright, and instead classifies music by a large variety of parameters (see Methods).

3 Methods

The general overview of the project is outlined in Figure 1. The application consists of a few parts. A backend, written in Python, conducts the video inputs and outputs. It also processes the video frames and queries Spotify based on the emotions in the frames. In order to output the audio, a React application hosts the Spotify Web Playback applet, which demonstrates the current song playing and allows for user login. Finally, the front-end React application is supported by a Node.js server backend, which provides the information for the React application to display.

All Python processes in this application (known as "workers" in this context) run asynchronously and communicate using a Queue datastructure. This is because the time for each processing step in the pipeline is variable, but results must be displayed at the same time.

3.1 Video Input and Output

I use a Python environment running OpenCV to pull video data from the webcam. Each frame is passed to DeepFace, which is a wrapper for multiple state-of-the-art emotion recognition models.

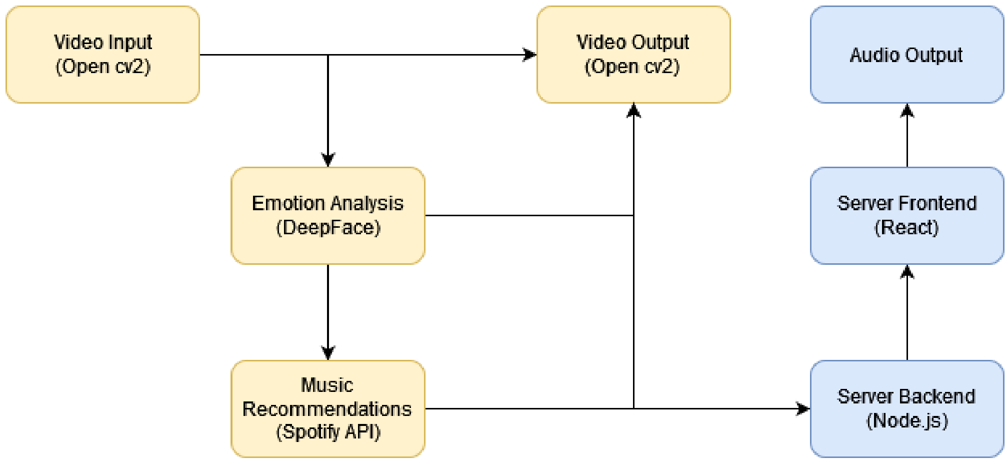

Figure 1: High-level overview of the project. Each block is a submodule, each arrow is a communication pipeline. Yellow: Python, blue: JavaScript.

Each video frame is rendered, and a user's emotion, if available, is displayed on the frame. Pulling and rendering video frames takes significantly less time than analyzing an emotion in a frame, and so often multiple video frames are displayed with the same analyzed emotion.

3.2 Server Backend and Frontend

Concurrently, a Node.js process supporting a React application is running to display the Spotify Web Playback SDK in a browser. The backend once again authorizes the user, and the browser functions as a device for Spotify playback, which is used as the default playback destination for the Spotify worker.

3.3 Emotion Analysis

I rewrote the available open-source code for real-time analysis of incoming video frames.

Each frame is classified by the probability that it is one of the following emotions: `angry`, `disgusted`, `scared`, `happy`, `sad`, `surprised`, and `neutral`. Then, the most likely emotion is chosen as the `dominant emotion`, which is then sent for downstream processing.

3.4 Music Recommendations

Parsed emotions are also passed to a third worker, which communicates with Spotify through the Spotify Web API. Upon initialization, the worker authenticates the user using the Spotify OAuth2 service, and then queries the API for the user's top songs and artists. The worker then makes a list of the genres of the user's top songs, and compares it with the available genres for recommendation.

After initialization, the worker loads the parameters for the recommendation request for each emotion. Every ten seconds, the worker queries Spotify for a list of recommendations given a random number of user top songs, artists, and genres as seeds, as well as the parameters given by the current emotion. The worker then plays a randomly chosen song from the returned list of suggestions, starting between one-sixth and one-third of the way through the song. I chose these values because they are close to the start of the song but are past any lead-ins or introns, which might not accurately represent the emotion codified by the song.

It is not a simple matter to assign emotions to recorded songs. Spotify does not directly contain emotional data, such as 'happy' or 'sad.' However, since its algorithms have created playlists such as "Happy Beats" or "Sad Piano," we can ask how these playlists are created and assigned.

Spotify's recommendations query requires a list of metadata than can be tuned. Each parameter has a maximum, minimum, and target value (e.g. `min_acousticness`, `max_danceability`, `target_valence`): `acousticness`, `danceability`, `energy`, `instrumentalness`, `liveness`, `loudness`, `popularity`, `speechiness`, `tempo`, and `valence`. See params.py for the tuned parameters for each emotion and see Parameter Tuning for a description of how I tuned each one. The descriptions of each musically revelant parameter are below (based on the Spotify reference):

- `Acousticness`: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
- `Danceability`: How suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.

- `Energy`: A measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

- `Instrumentalness`: Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

- `Liveness`: Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.

- `Loudness`: In decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

- `Popularity`: How popular a song is on Spotify. 0 represents a completely unknown song, and 100 represents one of the most popular tracks on the platform.

- `Speechiness`: Detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent classical music and other non-speech-like tracks.

- `Tempo`: The overall estimated tempo of a track in beats per minute (BPM). It derives directly from the average beat duration.

- `Valence`: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

The emotions of the currently-playing music are also passed to the video output, where they are rendered on each frame.

3.5 Parameter Tuning

Following the BRECVEMA model, discussed in Juslin 2019a, ch. 17, different aspects encoded in music can stimulate different emotions through the stimulation of different physical processes (Juslin 2019a, ch. 25, Table 25.1). Furthermore, David Huron's ITPRA Theory of Expectation describes how predictable and novel concepts can lead to frisson and surprise responses (Huron 2006). The following list describes how I tuned the parameters for each emotion. The complete list of tuned parameters can be found in params.py.

I did not vary the `instrumentalness`, `liveness`, `loudness`, or `popularity` parameters. I did not find the `instrumentalness` parameter to have a strong effect on the emotional valence of music. While music that is heard live is likely to cause a stronger emotional response, recorded live music often does not have the sound quality needed for a strong response, and so specified the `liveness` parameter to be low. Since the user can personally control the volume of the music I did not vary the `loudness` parameter. I also found that the `popularity` parameter wass not relevant to the emotional quality of the music.

- **Angry**: Anger, a basic emotion, is usually observed due to the stimulation of the amygdala in particular and limbic system in general. Furthermore, listeners can become physiologically entrained to the music, with their hearbeats and breathing changing due to the rhythms. It follows that angry music might be loud and, ideally, have a beat at around 2Hz for physical entrainment (120 bpm).
  - `Acousticness`: Music containing instrumental sounds, such as heavy guitar, is more likely to contain the needed tempi. I tuned this parameter to be on the lower end of the range, with a high maximum value.

- `Danceability`: Interestingly, music with low danceability parameters is more likely to be too slow for the physical entrainment response. Too-high values are more likely to stray into the 'happy' territory. I tuned these to contain median values.
- `Energy`: Energy must also be high enough for a high tempo, but not high enough for 'happy' emotions. I tuned these parameters to start at an energy value of 0.4 with no maximum.
- `Speechiness`: High-speechiness music can often evoke angry emotions through the composer's crafting of lyrics and language. However, I found that angry music is more likely to be evoked by music with a strong instrumental component, and so I tuned this parameter to be on the lower end of the range.
- `Tempo`: Physical entrainment responses respond most strongly to a 120 bpm tempo, and so that is the target that I chose.
- `Valence`: The valence of each song must be low enough to call on a fast song with a low emotional valence. As such, I specified this parameter to be near the bottom of the range.

- **Disgusted**: Disgust, another basic emotion originating from the limbic system, is a primal avoidance reflex, which is usually triggered when a positive expectation is not being met. Music that is disgusting must, as such, not meet the expectations of the listeners, in whatever sense. Disgust could also be classified as unpleasurable surprise response. Musicians often use predictablitiy to evoke listener pleasure through the predicion effect, with melodies consisting mostly of small pitch intervals and becoming more pleasurable the more a listener hears them. As such, I tuned the parameters to focus on discordant music that is surprising but in a way that is not pleasurable, where the listeners could not be able to predict the music.

  - `Acousticness`: This parameter did not have a noticeable effect on the disgust response to the music.
  - `Danceability`: I set this parameter to be in the middle of the range as that would allow for the greatest range in possible musical type.

- `Energy`: For the listener to perceive a surprising event and then a payoff in the short time dedicated to each song, there must be enough energy in the song. As such, I raised this parameter slightly above the middle of the range.
- `Speechiness`: This parameter did not have a noticeable effect on the disgust response to the music.
- `Tempo`: This parameter did not have a noticeable effect on the disgust response to the music, as surprising changes can happen at any tempo.
- `Valence`: I set this parameter to be low for the algorithm to recommend songs a surprising event with a negatively-valued payoff.

- **Scared**: Fear is one of the most primal emotions. It can be elicited by the slow buildup of the music towards something unknown, with an increase in one's pulse and other physiological states. The parameters for this kind of music were tuned to low-`valence` and mid-`tempo`.
  - `Acousticness`: This parameter did not have a noticeable effect on the fear response to the music.
  - `Danceability`: Songs with a high `danceability` parameter are less likely to have a buildip to something unknown, instead building towards a concrete resolution. As such, I tuned this parameter to have a maximum of only 0.4.
  - `Energy`: I tuned this parameter to be on the low end of the range for there to be more uncertainty in the music.
  - `Speechiness`: I found that songs with low `speechiness` are more likely to evoke a physical response. While words can, of course, tell a scary story, I tuned this parameter to be on the low end of the range for a more evokative instrumental sound.
  - `Tempo`: This parameter did not have a noticeable effect on the disgust response to the music, as fearful responses can happen at any tempo.
  - `Valence`: The emotional valence of the music must be low enough to cause a fearful response due to the emotional content of the music, and so that is where I tuned the parameter.

- **Happy**: An increase in excitement is observed in music with high rhythmic entrainment and tempo, usually thought of in a major key. For this, the `danceability` parameter was set to a maximum, as musical rhythms inducing a high level of motor responses (especially in a group setting) are perceived with a high emotional valence.
  - `Acousticness`: This parameter did not have a noticeable effect on the joy response to the music.
  - `Danceability`: Music with a high rhythmic entrainment is more likely to be perceived as happy, and so I set this parameter to maximum.
  - `Energy`: Music with a high energy is more likely to cause motor responses, and so I set this parameter to maximum.
  - `Speechiness`: This parameter did not have a noticeable effect on the joyful response to the music, as both music with words and without words can be perceived as happy with the correct combination of tempo and beat.
  - `Tempo`: This parameter did not have a noticeable effect on the joy response to the music.
  - `Valence`: The emotional content of happy music must have a high valence by definition. As such, I set this parameter to maximum.
- **Sad**: Sad music is thought of as music without fast changes, surprising events, or rhythmic entrainment. Commonly, this is stereotyped as slow and somber, although faster kinsd of music (such as folk songs) can also evoke a sad response. As such, I tuned the parameters for this kind of music to be low-`valence` and low-`tempo`, with minimal amounts of `danceability`.
  - `Acousticness`: I increased the value of this parameter for the music to be more likely to contain instruments.
  - `Danceability`: Sad music must not evoke a physical entrainment or cause motor responses, and so I set this parameter to minimum.

- Energy: Music with a high energy is more likely to cause motor responses, and so I set this parameter to minimum.
- Speechiness: This parameter did not have a noticeable effect on the sad response to the music.
- Tempo: I set this parameter to be lower for slower music.
- Valence: The emotional content of sad music must have a low valence by definition. As such, I set this parameter to minimum.

- **Surprised**: Surprised is a basic reflex originatng from the brainstem. Surprise can be evoked through schematic mechanisms, violating a listener's schema for a certain musical piece; dynamic, so that a work-specific element of surprise is violated, or veridical, violating expectations about the work itself. Surprise can also be positively valuenced due to pleasurable outcomes after the moment of surprise. As such, pleasurable surprising music must have high-`valence`, as music with a low-`valence` would more likely be understood as fearful or disgusting.
  - `Acousticness`: For a suprising emotional response in the short amount of time dedicated to each song, there must be enough instrumental context. As such, I increased the value of this parameter for the music to be more acoustic.
  - `Danceability`: This is a fragile parameter, as music with it set too low or too high is not likely to cause a surprising change. As such, I set it to a narrow range of values between 0.4 and 0.6.
  - `Energy`: The music for a surprise must have enough energy to cause a surprising change. As such, I set it to be elevated from the minumum.
  - `Speechiness`: This parameter did not have a noticeable effect on the surprise response to the music.
  - `Tempo`: I set this parameter to be lower, as I found that is where surprising changes are more likely to occur.
  - `Valence`: Surprising changes can occur at any valence. I slightly raised the minimum on this parameter to exclude music without any changes.

- **Neutral**: Neutral music is the hardest to define — what one person may consider neutral may be the most exciting to another. However, the parameters must be tuned in such a way that the music does not evoke other emotions.
  - `Acousticness`: I slightly lowered the maximum on this parameter to exclude purely acoustic music.
  - `Danceability`: Neutral music must not evoke a strong motor response, and so I lowered the maximum on this parameter to 0.5 to exclude it.
  - `Energy`: Neutral music must not be slow enough to cause a sad response or fast enough to cause an anger or joy response. As such, I set this parameter to a narrow range around the middle.
  - `Speechiness`: Music with only words is unlikely to be perceived as neutral, and so I lowered the maximum on this parameter.
  - `Tempo`: This parameter did not have a noticeable effect on the neutral responses to the music.
  - `Valence`: Neutral music must, by definition, be in the middle of the valence range. As such, I set this parameter to a narrow range around the middle.

4 Code Availability

The code is available on Github under the GNU General Public License, version 3.

5 Observations

I am planning to set up this project in the MADD Center at the University of Chicago. This will be a public art installation, and anyone will be able to interact with the project.

Over my course of testing and observations, I have found that people are more entertained with the emotional recognition portion of the project than with the music generation. Knowing that there are seven emotions, people would try to contort their faces to display all of them, which is not as easy as it sounds (disgust is notoriously hard to produce consistently).

6 Discussion and Conclusion

Music and emotions are, of course, indelibly linked. In this project, investigating how emotions affect music, I found that people are more interested in the analysis of their emotions than the music elicited by them.

References

AILabs. 2018. "AI music composition passed Turing test" (February). https://ailabs.tw/human-interaction/ai-music-composition/.

Clynes Manfred and Nettheim Nigel. 1983. Music, Mind, and Brain: The Neuropsychology of Music. Edited by Manfred Clynes. Chap. IV: The Living Quality of Music: Neurobiologic Patterns of Communicating Feeling. New South Wales State Conservatorium of Music.

Emotions can be communicated through physical expression, such as a change in finger pressure over time. These can then be converted to essentic forms with recognizeable emotional characteristics. Essentic forms (i.e., musical pieces with specific dynamic expressive communications) can then be created and synthesized by computers.

Cohen Annabelle J. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 11: Music as a Source of Emotion in Film. Oxford University Press.

Music adds to the diegetic functions of the film while being non-diegetic itself. Music, among other things, induces the mood of the scene in film. Emotions of characters can be communicated through the musical scores. Music also controls emotional responses, "directs attention to an object and ascribes its meaning to that object", and commands interest. Music represents emotion in the abstract, while the screen represents the object of emotional direction.

Drott Eric. 2018. "Why the Next Song Matters: Streaming, Recommendation, Scarcity." Twentieth Century Music 3 (15): 325–357.

Music must be curated. Music as a service, especially Spotify etc., is a SaaS. Recommendations provide the illusion of customer sovereignity. Commercials must flow from the music.

Gabrielsson Alf. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 19: Emotions in Strong Experiences with Music. Oxford University Press.

Lists reports of intense, positive, negative, and mixed reactions to music. Reactions are all mostly for basic emotions, with some complex emotions. It is difficult to distinguish between emotions expressed in and aroused by music.

Harrer G. and Harrer H. 1977. Music and the Brain: Studies in the Neurology of Music. Edited by MacDonald Critchley and R. A. Henson. Chap. 12: Music, Emotion and Autonomic Function. William Heinemann Medical Books Limited.

Emotional reactions to music depend on the attitude of the listener towards the music, their current mood, etc.. Different music can produce cardiovascular or motor responses. Responses greater in performers than in listeners. Pulse can synchronize to the beat of the music (or some subbeat). Muscular movement can be recorded throughout a musical piece (EMG). Low dosages of tranquilizers suppress muscle movements during musical listening, but not emotional experiences. Large doses suppress both.

Henson R. A. 1977. Music and the Brain: Studies in the Neurology of Music. Edited by MacDonald Critchley and R. A. Henson. Chap. 14: The Language of Music. William Heinemann Medical Books Limited.

Music is a language that communicates musical ideas well and succinctly, which speech is unable to. Music cannot be described in words. There may sometimes be ”meanings beyond musical ideas”, but that is up to the listener to decide. Some of these ideas evoke emotions: ”...music is then the image of these emotions...”.

Huron David. 2006. Sweet Anticipation: Music and the Psychology of Expectation. The MIT Press.

1: Anticipation is a very low-level concept in biological systems, and the ability to anticipate future events is important for survival. He posits the ITPRA Theory of Expectation: the Imagination response motivates an organism to do things that increase the likelihood of future beneficial outcomes; the Tension response prepares an organism for an event by changing arousal and attention; the Prediction

response provides positive and negative feedback based on outcomes, the Reaction response addresses the worst-case situations by making a protective response; and the Appraisal response providespositive and negative reinforcements. First two pre-outcome, last three post-outcome.

2. Surprise can be positively valenced due to pleasurable outcomes. Three kinds of surprise: laughter, generated through innocuous risks; awe, generated through the freeze response; and frisson, which is demonstrated through loudness and sudhen menical changes. Frisson: gaining command, awe: taking command, laughter: transformation.

5. Listeners expect melodies to consist mostly of small pitch intervals, with a reversal in direction after larger jumps. 8. People like music the more they listen to it.

13. Predictability can be created through schematic, dynamic, and veridical devices. Musicians often use predictability to evoke listener pleasure through the prediction effect.

14. Surprise can be evoked through schematic mechanisms, violating a listener's schema; dynamic, so that a work-specific element of surprise is violated; or veridical, violating expectations about the work itself.

15. Tension can be evoked by anticipation, suspension, and delay through eg. ritardandos. Premonition is also incredibly important.

Huron David and Margulis Elizabeth Hellmuth. 2011. Handbook of Msuic and Emotion: Theory, Research, Applications. Edited by Patrk N. Juslin and John A. Sloboda. Chap. 21: Musical Expectancy and Thrills. Oxford University Press.

Music is often sought out to change one's mood, second only to a conversation with a close friend. Prediction events in music elicit dopaminergic reward responses similarly to any other expected event. Musical frisson is elicited through mechanisms similar to fear.

Juslin Patrik N. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 14: Communicating Emotion in Music Performance: A Review and Theoretical Framework. Oxford University Press.

Performers can communicate emotions to listeners with high accuracy through cues. Complex emotions are a combination of basic emotions, which can vary in time and depend on note ordering. Timing patterns alone communicate some emotions, but less effective than tempo or dynamics. Emotions are communicated in music similar to using prosodic cues in the voice. Muic might stimulate emotions through an "emotional contagion" or through violation of musical expectations.

———. 2019a. Musical Emotions Explained: Unlocking the secrets of musical affect. Oxford University Press.

———. 2019b. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 15: Does Music Arouse Emotions? How Do We Know? Oxford University Press.

This book is the easiest to read so far. Music arousing emotions can be measured through the following mechanisms: self-reporting, psychophysiological measurements, specific patterns of neural activation (most likely a network of the amygdala, hippocampus, para-hippocampus, temporal poles, and maybe the pre-genual cingulate cortex), or action tendencies.

———. 2019c. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 16: The Prevalence of Emotional Reactions. Oxford University Press.

Music is incredibly prevalent but doesn't always elicit emotions. Music usually arouses positive emotions: joy, interest, contentment, and love; but may also elicit complex emotions. Mixed emotional responses can also occur to music. Music doesn't induce unique emotions. Interestingly, music can also arouse emotions at the individual but also interpersonal and intergroup levels.

Juslin Patrik N. 2019d. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 17: How Does Music Arouse Emotions? Oxford University Press.

Music is, indeed, a stimulus which then, through a complex mechanism, arouses emotions, just as any other stimulus. Emotions can be triggered by a "musical event," containing music, listener, and context. Mechanisms for emotion effectations are the BRECVEMA model: Brainstem reflex, Rhythmic entrainment, Evaluative conditioning, Contagion, Visual imagery, Episodic memory, Musical expectancy, and Aesthetic judgement. Could also be cognitive goal appraisal. He goes deeper in depth in each of these in the subsequent chapters.

———. 2019e. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 25: Predictions, Implications, Complications. Oxford University Press.

He has a table of each of the BRECVEMA aspects with, among other things, induced affect, induction speed, degree of volitional influence, availability to consciousness, and dependendence on musical structure. B: arousal, surprise; R: arousal increase/decrease; E: basic emotions; C: basic emotions; V: all possible emotions; E: all possible emotions (esp nostalgia and longing); M: anticipation, anxiety, surprise, releif, tension. All are medium dependent on musical structure except E, E (low) and M (high). Musical emotions involve A1, cingulate, PFC, PAG, emotional processing. Real-world musical events often multiply mediated, not just one-to-one emotion:music. BRECVEMA response framework explains individual differences between listeners, and also explains no response (if no relevant information present). Personal musical induction mechanisms still influenced by context. Musical emotion evocation is moderately universal across cultures: low-level mechanisms conserved, responses not.

———. 2019f. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 29: Aesthetic Criteria: Meet the Usual Suspects. Oxford University Press.

This chapter and this part all deal with the Aesthetic judgement part of the BRECVEMA model. To be aesthetically valuable, music should be heard as expressive. Not all good music is beautiful but all good music is expressive. Novel music is interesting. People prefer music that arouses emotions, but emotional arousal does not require aesthetic judgements. Preferences for specific styles can either help or hinder appreciation of the music at hand.

Juslin Patrik N. 2019g. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 30: A Novel Approach Towards Aesthetic Judgement. Oxford University Press.

Music judged as aesthetically pleasing, greater than some aesthetic threshold, induces positive emotions, while displeasing emotions induce negative emotions. Not all art strives to evoke positive emotions.

———. 2019h. Musical Emotions Explained: Unlocking the secrets of musical affect. Chap. 32: The Last Chorus: Putting It All Together. Oxford University Press.

Music can effect pleasurable sad (or otherwise negative) emotions. Listeners can prepare to feel the strongest emotions from a certain musical listening.

Juslin Patrik N. and Sloboda John N. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 4: Psychological Perspective on Music and Emotion. Oxford University Press.

Emotions are hard to define, but can be measured using self-reporting, expressive behaviors, or physiological measures. Emotions are eliciting through proximally eliciting stimuli in an unconscious manner. Sources of emotions in music can be either intrinsic to the music or extrinsic, with complex interactions.

Mulligan Mark. 2022. “Music subscriber market shares Q2 2021.” MIDiA (January). Accessed March 23, 2022. https://www.midiaresearch.com/blog/music-subscriber-market-shares-q2-2021.

Scherer Klaus R. and Zentner Marcel R. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 16: Emotional Effects of Music: Production Rules. Oxford University Press.

Music can evoke emotions through: A. Central route production: 1. appraisal theory, which evaluates the con/de-structiveness of a musical stimulus and its compatibility with internal and external standards;

2. by bringing up memories with an emotional valence; or 3. empathetic responses to the performers. B. Peripheral route production: 1. proprioceptive feedback; or 2. by facilitating pre-existing emotions.

Schubert Emery. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 17: Continuous Measurement of Self-Report Emotional Response to Music. Oxford University Press.

Emotions can only be induced over time.

Sloboda John A. and O'Neill Susan A. 2001. Music and Emotion: Theory and Research. Edited by Patrik N. Juslin and John N. Sloboda. Chap. 18: Emotions in Everyday Listening to Music. Oxford University Press.

Music is ubiquitous and intrinsically socially and culturally linked, which people often forget. People listen to music to fill functional niches, and can influence mood. People like music moe when they can control it (i.e., in private spaces, or busking). Music is directly tied to personality and identity, and can be used daily as self-therapy.

Small Christopher. 1998. Musicking: The Meanings of Performing and Listening. Chap. Prelude: Music and Musicking. Hanover: Wesleyan University Press.

Music exists in a very wide variety of contexts. Claims that music is an activity, not a thing or an object, and so "music" DNE. Music currently is valued as the object itself instead of the methods of creating, perceiving, or responding to it. Interactions during musical events are assumed to be composer to performer to listeners, and not across or backwards. Musical works exist in order to give performers something to perform. "To take part in music is central to our very humanness...".

Stravinsky Igor. 1936. An Autobiography. Simon and Schuster.

Quote: "Most people like music because it gives them certain emotions, such as joy, grief, sadness, an image of nature, a subject for daydreams, or – still better – oblivion from "everyday life." They want a drug – "dope." It matters little whether this way of thinking of music is expressed directly or is wrapped up in a veil of artificial circumlocutions. Music would not be worth much if it were reduced to such an end. When people have learned to love music for itself, when they listen with other ears, their enjoyment will be of a far higher and more potent order, and they will be able to judge it on a higher plane and realize its intrinsic value.".

Sundberg Johan. 1983. Music, Mind, and Brain: The Neuropsychology of Music. Edited by Manfred Clynes. Chap. VII: Speech, Song, and Emotions. New South Wales State Conservatorium of Music.

Sundberg discusses vocally generated emotions, focusing on phonation frequencies. He also discusses singing with an emotional expression, with the following observed criteria, increasing: 1. Tempo: sorrow, neutral, joy, anger, fear. 2: Pauses between syllables: neutral, joy, anger, sorrow, fear. 3: Voice amplitude: fear, neutral, sorrow, joy, fear. 4: Tone onsets/offsets: anger/fear, joy, neutral, sorrow. Joy requires more than just signal characteristics (eg pitch, loudness, outset, etc.).

Yamaguchi Osamu. 2012. "Perception and Machine Intelligence: First Indo-Japan Conference, PerMIn 2012, Kolkata, India, January 12–13, 2011, Proceedings." Chap. Face Recognition Technology and Its Real-World Application, edited by Malay K. Kundu; Sushmita Mitra; Debasis Mazumdar; Sankar K. Pal, 28–34. Springer Science / Business Media.