2019

Feel the Moosic: Emotion-based Music Selection and Recommendation

Patrick Helmholz

Michael Meyer

Susanne Robra-Bissantz

32ND BLED ECONFERENCE
HUMANIZING TECHNOLOGY FOR A SUSTAINABLE SOCIETY
JUNE 16 – 19, 2019, BLED, SLOVENIA, CONFERENCE PROCEEDINGS
A. Pucihar (ed. et al.)

University of Maribor Press

Feel the Moosic: Emotion-based Music Selection and Recommendation

PATRICK HELMHOLZ, MICHAEL MEYER & SUSANNE ROBRA-BISSANTZ

Abstract Digital transformation has changed all aspects of life, including the music market and listening habits. The spread of mobile devices and music streaming services has enabled the possibility to access a huge selection of music regardless of time or place. However, this access leads to the customer's problem of choosing the right music for a certain situation or mood. The user is often overwhelmed while choosing music. Context information, especially the emotional state of the user, can help within this process. The possibilities of an emotional music selection are currently limited. The providers rely on predefined playlists for different situations or moods. However, the problem with these lists is, that they do not adapt to new user conditions. A simple, intuitive and automatic emotion-based music selection has so far been poorly investigated in IS practice and research. This paper describes the IS music research project "Moosic", which investigates and iteratively implements an intuitive emotion-based music recommendation application. In addition, an initial evaluation of the prototype will be discussed and an outlook on further development will be given.

Keywords: • Music • Emotion • Mood • Recommendation • Context • Digital transformation •

CORRESPONDENCE ADDRESS: Patrick Helmholz, Chair of Information Management, Technische Universität Braunschweig, Braunschweig, Germany, e-mail: p.helmholz@tu-braunschweig.de. Michael Meyer, Chair of Information Management, Technische Universität Braunschweig, Braunschweig, Germany, e-mail: m.meyer@tu-braunschweig.de. Susanne Robra-Bissantz, Chair of Information Management, Technische Universität Braunschweig, Braunschweig Germany, e-mail: s.robra-bissantz@tu-braunschweig.de.

DOI https://doi.org/10.18690/978-961-286-280-0.11 ISBN 978-961-286-280-0
Dostopno na: http://press.um.si

1 Introduction

Smartphones, mobile broadband connectivity and music streaming services have massively changed the way people listen to music in everyday life. Music listeners now have the ability to access an almost unlimited number of songs and consume them anytime as well as anywhere. Listening to music accompanies us in all everyday situations (DeNora, 2011). Whether at parties or weddings, on the drive to work, in the gym or alone at home, music has become part of our social and physical environment (Pettijohn, Williams, & Carter, 2010). The time we spend listening to music every day increases from year to year. On average, people now listen to over four hours of music per day (Nielsen, 2017). Every second person uses music streaming services and around 80 percent of smartphone owners frequently use their devices to listen to music (Clement, 2018). Consequently, listening to music is the most important accompanying activity in our society (DeNora, 2011).

Music streaming has become the way music is consumed. From 2013 to 2018 the number of music streams listened to per year increased nine times and just from 2017 to 2018 it increased by 50 percent (IFPI, 2018; Nielsen, 2019). Streaming now accounts for 75 percent of music industry revenue (RIAA, 2018).

The mass of choices created by the digitalization and streaming of music, as well as the availability of music independent of time and place, can lead to user problems. Since music is emotional, it can strengthen and change the emotions of the listener. As a result, people enjoy listening to music that fits their mood and situation. It is not always trivial to find the right music for the listening situation and the user's mood. (Sloboda, 2011).

Users are often overwhelmed by the large number of digital music titles and the possibilities for selecting music based on mood are currently very limited. Music streaming services, such as Spotify or Apple Music, offer pre-defined playlists to structure music for their user base and to facilitate the selection. This can be done in different ways. The most common way to classify music has been the genre for a long time, although this distinction is usually not clear. Nowadays, playlists should be compiled according to the situation and mood. Accordingly, music platforms increasingly offer such playlists.

The problem with these playlists is, that they do not adapt to new user situations. If the situation changes, the user has to select a new playlist that fits to his needs. A change in the situation or mood leads to a new search and an associated high as well as time-consuming interaction with the service. However, if the service knows the situation or mood of the user, it can react accordingly and generate a new, adapted playlist. Thus, such a service would be called a smart, context-based music player.

Situations and activities often lead to certain user emotions or are associated with concrete moods. As a result, the categorization and selection of music according to mood or emotion is a modern possibility of context-based and user-centered music classification (Jamdar, Abraham, Khanna, & Dubey, 2015). The research field of emotional music recommendation in the IS discipline is relatively new, but on demand. The increasing sale of modern mobile devices such as smartphones and smartwatches is creating new opportunities for user interaction and the collection of user data. The research project Moosic deals with the development and investigation of an emotional music player which is easy and fast to use and enables an immediate emotional change in the music playback with one click.

In the context of this contribution the context-based – here especially emotion-based – music playback will be dealt with in more detail. It also shows how music can be classified on the basis of emotions and what possibilities the current prototype of Moosic offers to recommend music. Furthermore, a comparative user study to different user interfaces of Moosic will be introduced and discussed. Finally, the paper is summarized and an outlook on the future significance of context-based music recommendations in the digital world, as well as on the automatic capture of emotions as an extension of the prototype is given.

2 Theoretical Foundation

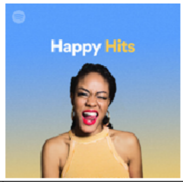
2.1 Music Recommendation

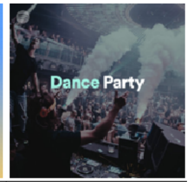
Music has an influence on our mood and can support or change it (Gaston, 1951). In order to bring the musical influence on our mood in line with our situation, music must first be categorized (Brinker, Dinther, & Skowronek, 2012).

As already mentioned, this can happen in different ways. The most commonly used way to classify music nowadays is by genre. The entire categorization of artists and albums, as well as the creation of chart rankings, is based on the classification by genres. Playlists such as "Hip-Hop", "Rock" or "Country", which are genre-based, are referred to as content-based playlists. However, these genres are often very broadly based and the boundaries between them are still blurred, making the problem of automatic classification of music a non-trivial one (Scaringella, Zoia, & Mlynek, 2006). Moreover, Daniel Ek – Founder and CEO of Spotify – said as early as 2015 that music search and classification is moving away from the genres. People are no longer looking for hip-hop or country, but for activities or a particular experience. Therefore, the playlists should be generated according to the situation and mood.

Playlists such as "Happy Hits" or "Sports Power" are composed accordingly and are referred to as context-based playlists. In addition to these two types of playlists, there are also those that cannot be clearly assigned. These playlists have both content-based and context-based characteristics. This is referred to as hybrid playlists. Examples would be playlists like "Dance Party" or "Chill Hits". The following Table 1 shows exemplary playlists from Spotify for the three different types with their key statistics. All playlists are from the Top 50 Spotify playlists.

Table 1: Examples for different playlist types





Playlist type	Content-based	Context-based	Hybrid
Playlist title	Rock Classics	Happy Hits	Dance Party
Follower	5,137,461	3,395,018	2,940,899
Avg. monthly listeners	689,253	964,063	110,936
Monthly listeners to followers ratio	13 %	28 %	4 %
Spotify ranking	12	31	45

(Data retrieved from Chartmetric.io on March 01, 2019)

Initial analyses have shown that the proportion of such context-based playlists is growing, even though content-based playlists are currently still more common. However, these available context-based playlists are already more popular. The analysis was based on the 973 playlists from the "Genres and Moods" section of the music streaming service Spotify. About 57 percent of the playlists examined are content-based and only about 37 percent context-based. Hybrid playlists account for about 7 percent. Nevertheless, if one considers the average number of followers, context-based and hybrid playlists are much more popular than content-based ones. The growth in followers from 2017 to 2018 is also more than 20 percent higher for these two playlists than for content-based playlists.
Accordingly, it can be deduced that there is already a rethinking about the composition of playlists in music streaming services, which is mainly influenced by the user base. The increasingly widespread digital personal assistants are also getting functions to get music for the current situation and mood. For example, Alexa (the virtual assistant of Amazon) now supports the selection of a playlist on request and takes current conditions and user preferences by interviewing the user into account (Welch, 2018).

2.2 Emotions and Music

In the course of the digital transformation, new possibilities arise to support the users on the hardware and software side. Modern mobile devices make it possible to precisely capture the context of a user. This is made possible by a multitude of sensors, which are installed in smartphones as well as smart watches and accompany the user inconspicuously in his everyday life. Thus, it is possible to move mobile applications even closer to the user needs in order to provide him with a concrete benefit in the situation. Due to technological progress, the cost and size reduction of devices and the further development of sensors in the context of digital transformation, situation-oriented applications are constantly being driven forward (Yurur et al., 2016). Current smartphones and smartwatches are able to determine the current emotion of the user. With the help of various sensors, like pulse, skin temperature or skin conductance, these mobile devices are able to draw conclusions about the emotional situation of the user. Due to the strong influence of emotions on the user, these will be examined in more detail below in relation to the music played. In the concrete case of the prototype Moosic, a service was created, which gives the user the opportunity to adjust the music playback to his current emotional state. This is done within the first prototype via user input (see also Chapter 3).

Emotions are the reaction of the human body to an occurring stimulus, such as an event of certain importance. The reasons for such an emotional reaction can be of different nature. For example, an emotional response can be triggered when a user is prevented from satisfying his needs or achieving his goals. Furthermore, the occurrence of an emotional response can be the result of an existing emotional situation or a previous emotional situation. In the perception of an event in the human environment, it also becomes apparent that emotions involve a degree of pleasure or displeasure (Brave & Nass, 2009; Cabanac, 2002). The emotional reaction to an event or a concrete situation can therefore be positive or negative. Emotions can also lower the threshold of the occurrence of other emotions. Emotions are a typically human trait which can influence many aspects of our lives. Thus, the perception, rational thinking and even the decision-making of a user are not free of emotions (Brave & Nass, 2009; Hussain & Bieber, 2009; Reeves & Nass, 1996). While emotions are triggered by a certain event and only exist in a relatively short time period, moods are caused indirectly and trend to exist over a longer term.

Thus, emotions last seconds and moods can remain for days. Moods are also able to influence judgement and decision-making, while at the same time they can lower the threshold for emotions (Brave & Nass, 2009). Although there is a difference between moods and emotions, in this paper, both terms are treated synonymously here. For instance, the Circumplex-Model-of-Affect by Russell (1980) offers a possibility to classify emotions and to represent them in a model (see Chapter 2.3).

Due to the strong influence of emotions on many different aspects of our lives, this prototype will use emotions as a basis for music selection and playback. The modern sensory possibilities of mobile devices offer the possibility to understand individual situations of a user and to provide suitable music for them. Since emotions are the result of a concrete situation, they can be used as a kind of situational variable to adapt the music playback to a special customer situation. Emotions and music are strongly connected. Music is capable of triggering emotions, amplifying, weakening or even changing them (Sloboda, 2011). Conversely, a concrete situation can influence the user's needs to listen to a specific type of music. Thus, situations cause emotions and emotions can act as a guide for the music selection. As already mentioned, common classifications of music often concentrate on genres and try to classify music into artificial and inseparable categories. A system-based and automatic (emotional) classification of music is already possible and appears to be more natural and humanly (Jamdar et al., 2015).

2.3 Emotion-based music recommendation

„The idea [of Context-Aware Music Recommender Systems (CAMRS)] is to recommend music depending on the user's actual situation, emotional state, or any other contextual condition that might influence the user's emotional response and therefore the evaluation of the recommended items." (Ricci, 2012, p. 865). CAMRS have been researched for some time, but the increasing mobile use of services is creating new challenges. Due to the great importance of emotions in and for music, emotional music recommendation has become particularly important as a sub-area of CAMRS. The research field of emotional music recommendation in the IS discipline is relatively new, but on demand. Only a few IS prototypes investigate emotional music recommendations (see e.g. (Ayata, Yaslan, & Kamasak, 2018; Janssen, Broek, & Westerink, 2012; Nathan, Arun, & Kannan, 2017)), but they do not sample a multidimensional emotion model that has already been verified. According to Russell's Circumplex-Model-of-Affect, emotions can be classified in a two-dimensional order. This model arranges the emotions in circular order according to arousal and valence. The dimension of arousal ranges from calming or soothing to exciting or agitating, whereas the dimension of valence ranges from highly negative to highly positive (see Figure 1, left side).

Therefore, the model is able to represent each emotional state in the form of a certain degree of these two dimensions (Kensinger, 2011; Russell, 1980). Basically, the emotions can be divided into the four quadrants Q1 - Angry, Q2 - Happy, Q3 - Sad and Q4 - Relaxed.

Based on Russell's circular emotion model, Thayer developed an alternative emotion model that arranges 11 emotional states using tiles (Thayer, 1991). Thayer's model is also based on the two dimensions arousal and valence (see Figure 1, right side). Both models were simplified as well as colour coded in the research project and for the prototype to simplify the user interface and accordingly the selection process by the user (see also Figure 3).

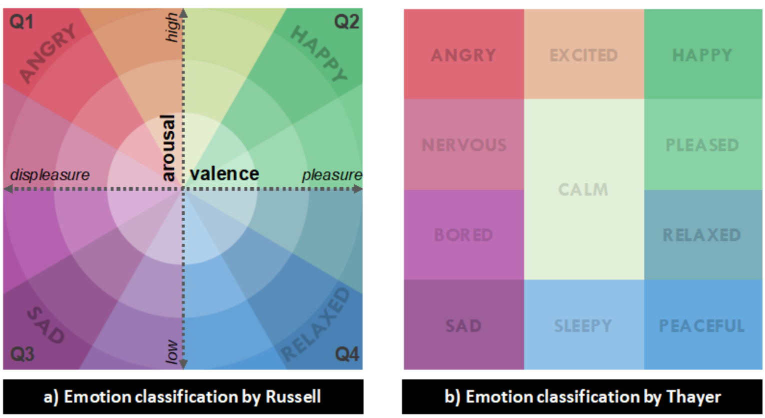


Figure 1: Human emotional classification model a) by Russell (1980) and b) Thayer (1991)

In the case of music, this two-dimensional scale can also be used to classify songs emotionally. Here energy corresponds to the human activation or arousal and valence to the human mood (Krause & North, 2014; Russell, 1980). Popular streaming services such as Spotify use these two music parameters to classify music (Jamdar et al., 2015). Energy in music is a perceptible measure of intensity and activity. Typically, energetic tracks feel fast, loud and intense. For example, Death Metal very often has high energy, while a ballad has low values on the scale. Valence, on the other hand, describes the musical positivity conveyed by a piece of music. Songs with high valence sound more positive (e.g. happy, cheerful, euphoric), while pieces with low valence sound more negative (e.g. sad, depressed, angry) (Kim, Lee, Kim, & Yoo, 2011). In order to better understand the emotional classification of music, Figure 2 shows the classification of different songs from various genres. Valence and energy are strong indicators for the acoustic mood and the overall emotional quality of a song (Krause & North, 2014).

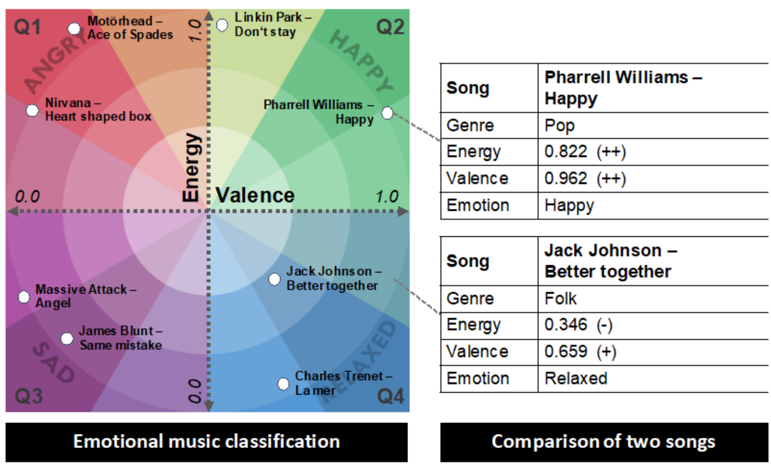


Figure 2: Emotional classification of music by energy and valence parameters with exemplary songs for each quadrant

Spotify determines and stores these emotional parameters automatically by algorithms for each individual song in addition to other parameters such as danceability or speechiness. These so-called high-level parameters are generated by using low-level and mid-level features of the music. Low-level features are timbre and temporal features whereas mid-level features are rhythm, pitch, and harmony. Both have been used predominantly in music classification, due to the simple procedures to obtain them and their good performance. However, they are not closely related to the intrinsic properties of music as perceived by human listeners (Fu et al., 2011). Accordingly, nowadays the derived high-level features are primarily used to recommend music and to create playlists. These parameters are set on a scale from 0.0 to 1.0. Consequently, songs in the first quadrant have a high energy and a low valence value and songs in the fourth quadrant, for example, have a low energy and a high valence value.

The right side of Figure 2 shows the two songs "Better together" by Jack Johnson and "Happy" by Pharrell Williams with their genre classification and their values for energy and valence. Furthermore, it is indicated in which emotional quadrant they are classified.

The first song is characterized by a low energy and a high valence value. Accordingly it can be classified as relaxed. The second song has both, a very high energy and a high valence value and can be classified as happy. The presented parameters energy and valence as well as the two-dimensional emotion models were used in this project to develop a prototype, which selects and plays music based on an emotional mood input (see Chapter 3).

3 Moosic Prototype

The prototype "Moosic" was developed as part of the research project. The application uses Spotify and the available API to create a playlist based on the user-selected genres and the emotional input of the user (see Figure 3).

The music selection works via a circular avatar, which shows the Spotify profile picture of the user. Moving the avatar allows the user to select his emotional state. During the implementation of the emotional input area, attention was paid to a user interface that is as simple as possible and therefore color-coded. The models were implemented in a somehow simplified way to not overwhelm the user while selecting the emotional state. Both, an input area based on Russell's model (see Figure 3, right side) and one based on Thayer's model were implemented (see Figure 3, left side).

The advanced settings of the application can be used to switch between the two surfaces. In addition, the user can select several genres as well as the popularity of the songs to further adapt the music selection by the system to his personal preferences. The position of the avatar (user input) in the frontend is interpreted by the system in the backend as x- and y-axis values between 0.0 and 1.0. The position of the avatar on the x-axis represents the value for valence and on the y-axis the value for arousal. These values are provided with a certain tolerance range, which is +/- 0.1 by default, but can be changed in the settings. The tolerance range results in minimum and maximum values for valence and energy, which are combined into a request to the Spotify API for the selected genres. To this request, Spotify returns a playlist with 20 matching songs, which is played randomly (see Figure 4).

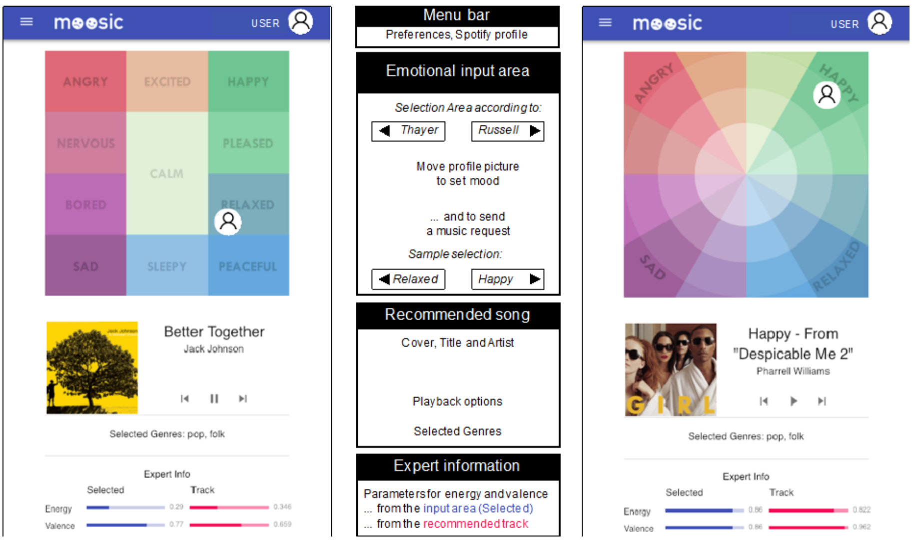


Figure 3: User interface of Moosic with input area according to Thayer (1991) left and Russell (1980) right using two example entries (Relaxed and Happy)

In addition to the emotional input area, the application provides information about the currently played song as well as the usual playback options. You can also activate an expert mode, which displays more information, like the search parameter input and the parameters for the actually played song. In addition, the expert mode allows the music selection to be further restricted.

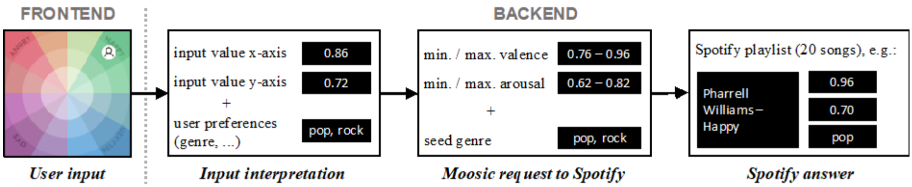


Figure 4: Input/output process and data processing of Moosic

4 Evaluation of Moosic

In a first experiment the prototype of Moosic was evaluated. The experiment was conducted in the form of a controlled laboratory experiment. Four participants carried out the experiment at the same time. In order to enable standardized test conditions that ensure comparability of the results, attention was paid to uniform instructions and equal the testing conditions. The probands were divided into two test groups, resulting in a Between-Subjects study design. Each participant received the same accompanying online questionnaire and was exposed to the same testing scenarios. Both, the general usability and the music playback of the prototype as well as the two different emotional input areas were tested. One test group got the prototype with the input area according to Russell's model (T1) and the other group according to Thayer's model (T2).

Totally 43 subjects (30 male / 13 female) took part in the experiment. The age of the subjects was between 20 and 34 years and the mean value is 24.4. About 86 percent of the participants use music streaming services (65% paid subscription, 21% free version). The composition of the test groups T1 (21 subjects) and T2 (22 subjects) was very similar based on sociodemographic data (see Table 2).

The majority of the total participants listened to music between 30 minutes and two hours a day (72 %). The most important channel is music streaming, as 65 percent have stated that they very often listen to music via this channel before YouTube (16 %) and classic radio (14 %). All other media and channels are negligible. With regard to music streaming, Spotify is clearly ahead with about 80 percent of the probands who use music streaming. About 92 percent of all test persons consider it as useful to receive music recommendations from the music provider. In their main activities, the respondents stated that they listen to music while on the move with mean value on a four-level frequency Likert scale of 3.35, followed by sports (2.95), housework and relaxation (2.56 each), work (2.28) and learning (2.12).

Table 2: Sociodemographics of the participants

Attribute	Overall	T1: Russell	T2: Thayer
Participants	43	21	22
Sex (male/female)	m: 30; f: 13	m: 14; f: 7	m: 16; f: 6
Age (by mean)	M: 24.4	M: 25.3	M: 23.7
Usage of music streaming services (yes/no)	y: 86%; n: 14%	y: 90,5%; n: 9,5 %	y: 81,8%; n: 18,2%

After the test had been carried out, the subjects were presented with standardized statements on the satisfaction with the prototype. These were statements about navigation, user interface, target fulfilment and mood input. These statements were evaluated on a four-level approval Likert scale. With the help of a mean value comparison significant similarities and differences can be determined (see Table 3).

Table 3: Results of the statements for the two different input surfaces T1 and T2 (on a four-level Likert scale by mean value comparison for each group)

Statement: The prototype …	T1: Russell	T2: Thayer
… is visually appealing.	2.86	2.91
… has an intuitive and understandable navigation.	3.00	3.18
… is user-friendly in its interface.	3.10	3.18
… is self-explanatory in its functionality.	3.05	3.09
… has responded to my individual needs.	2.57	2.82
… includes different expressions of the individual emotions.	2.76	3.23
… makes it easy to switch between different emotions.	3.62	3.45

… is applicable to all genres.	2.71	2.73
… can be used for different emotions.	3.14	3.27
… can be easily adapted to my emotion.	3.05	3.09
… represents an added value for music streaming services.	3.19	3.09
… has addressed my emotional receptivity to music.	2.62	2.55
… is valuable and recommendable.	2.90	2.91
Overall	**2.98**	**3.01**

The statement can be made that the test group of the Thayer surface (T2: 3.23) perceived the different expressions of the emotions better through the prototype than the participants who were provided with the Russell surface (T1: 2.76). This can possibly be explained by the fact that T2 verbalized significantly more emotions than T1. The surface has an effect on the perception of emotions by the participants. However, the fact that one prototype can better represent the currently perceived emotions than the other cannot be proven by the data collected. Nor is it possible to prove that one of the two prototype surfaces responds better to the individual wishes of the user.

Finally, it can be said that the participants stated a relatively high level of satisfaction with the prototype, which applies equally to both surfaces (T1: 2.98; T2: 3.01). Both test groups were willing to continue using the application for music selection and recommendation. However, with about 86 percent approval it was significantly higher for T1 than for T2 (64 %). A further qualitative study could possibly be used to gain further insights into the advantages of the two surfaces and to combine the advantages of both input surfaces into a new one.

5 Conclusions

Digital transformation is an all-encompassing phenomenon, which also influences the music market. Concrete manifestations of this transformation are the development and increasing use of streaming services such as Spotify, Apple Music, Amazon Music or YouTube (Music). These streaming service providers serve the user with a wide range of music titles, which makes the selection process of suitable titles more difficult. Music recommendations help to solve this problem. In addition, a change in user behavior can also be observed, as they increasingly consume their music in a mobile context thanks to increasingly powerful mobile devices (Clement, 2018). Typical mobile devices, such as smartphones, which contain a multitude of sensors, are able to observe and interpret the situation of a user (Yurur et al., 2016). For example, it is possible to measure the user's biofeedback and derive emotional states based on it. In this way, modern mobile devices in combination with smart and innovative services can deliver added value to the user. This makes it possible to adapt the selection of music to the current situation or to the current emotional state of the user. Music is able to influence our emotional state or the emotional state of the user influences the type of music he wants to listen to. Emotions are strongly linked to the listening behaviour of a user, which makes emotions a useful basis for music selection (Han, Rho, Jun, & Hwang, 2010).

Based on these findings and problems, a first prototype of an emotional music player was developed and evaluated. The presented prototype offers the user the possibility to enter his emotional state. Based on this input, a suitable playlist is generated. The user interface for entering the emotional state is based on two different two-dimensional models for classifying emotions according to Russell (1980) and Thayer (1991). Both models were implemented and integrated in a simplified and color-coded form.

Furthermore, an experiment with the prototypes was carried out, which already provided initial insights into its added value and its perception by the users. In addition to the test subjects' listening habits, connections between music and emotions, as well as the meaningfulness of emotions were queried as a data basis for a selection of music. An essential part of the experiment was the comparison of these two different surfaces. Here it was shown that the general satisfaction with the application is very high and the selection of music via the input of emotions is considered meaningful.

6 Future development and outlook

In the next steps of this research project the prototype will be extended by an automated emotion measurement, which can support or even replace the manual input of the user. The sensory possibilities described in chapter 2.2 will serve as a basis for the measurement. The measurement of various user characteristics with the help of smartphones and smartwatches is already possible today (Bachmann et al., 2015). The voice and facial expression can be used to draw conclusions about the emotional situation of the user (Essa & Pentland, 1995). However, these methods are less suitable in the mobile context of listening to music when the user does not interact directly with the system in the mean time. Furthermore, the measurement and evaluation of biofeedback offers a possibility to derive emotions. Promising sources of information can be heart rate, skin conductance or skin temperature (Picard & Klein, 2002).

In order to measure and use emotions for the music recommendation in a mobile context, unobtrusive and non-interfering possibilities of measurement are better suited. Smartphones and especially wearables, like smartwatches as well as activity trackers, which are equipped with various biometric sensors (Bachmann et al., 2015) should therefore be given priority. The smartphone based measurement of basic emotions, such as anger or joy, based on the change of the heart rate, was already proven (Lakens, 2013). A biofeedback-based and two-dimensional approach was described by Yamamoto et al. in 2009. This approach can be easily adapted for the applied emotion model(s) and the prototype. In this case heart rate corresponds to energy and skin temperature corresponds to valence (Yamamoto, Kawazoe, Nakazawa, Takashio, & Hideyuki, 2009). Therefore high heart-rate of the user can be interpreted as a high level of energy, whereas a high skin temperature would be classified into the positive area of the valence dimension. As smartphones and wearables become more and more powerful and even more sensors are implemented, biofeedback especially heart rate and skin temperature represent promising sources of user-based emotional information (Di Lascio, Gashi, & Santini, 2018).

Meanwhile, the self-input of emotion by the user is to be maintained and possibly extended by further elements. Because in addition to automatic measurement of the users' emotional state and music recommendation based on it, the user should still have the option to enter or specify his emotional state or preference manually. For this purpose, the use of emoticons and other adjectives is conceivable in order to offer the user a simpler way of expressing his emotional state (Meschtscherjakov, Weiss, & Scherndl, 2009). Further context of the user, such as the time of the day, weather or his location, can also help to better understand the user's situation and to ensure an even better music selection and recommendation depending on the situation. Furthermore, the consideration of different scenarios within the framework of a field experiment would be useful. Concrete use cases such as sports, shopping or learning are only a few examples in which an automatic emotion-based music recommendation could support the user.

In the experiment, the wish for an evaluation respectively feedback system was expressed several times. A feedback system would open up the possibility of a more customized music recommendation system which adapts even better to the preferences of a certain user by using his feedback data. Thus the system may learn that the user in a certain mood does not want to support this mood, but rather wants to initiate a change of his emotional state.

These functions and extensions will be implemented and tested iteratively in the next versions of the prototype as well as a slightly modified input surface.

References

Ayata, D., Yaslan, Y., & Kamasak, M. E. (2018). Emotion Based Music Recommendation System Using Wearable Physiological Sensors. IEEE Transactions on Consumer Electronics, 64(2), 196–203. https://doi.org/10.1109/TCE.2018.2844736

Bachmann, A., Klebsattel, C., Schankin, A., Riedel, T., Beigl, M., Reichert, M., … Ebner-Priemer, U. (2015). Leveraging Smartwatches for Unobtrusive Mobile Ambulatory Mood Assessment. Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, 1057–1062. https://doi.org/10.1145/2800835.2800960

Brave, S., & Nass, C. (2009). Emotion in Human–Computer Interaction. In A. Sears & J. Jacko (Eds.), Human-Computer Interaction Fundamentals (Vol. 20094635, pp. 53–68). https://doi.org/10.1201/b10368-6

Brinker, B. den, Dinther, R. van, & Skowronek, J. (2012). Expressed music mood classification compared with valence and arousal ratings. EURASIP Journal on Audio, Speech, and Music Processing, 2012(1), 1–14. https://doi.org/10.1186/1687-4722-2012-24

Cabanac, M. (2002). What is emotion? Behavioural Processes, 60(2), 69–83. https://doi.org/10.1016/S0376-6357(02)00078-5

Clement, M. (2018). Studie zur Zukunft der Musik. Retrieved from Lehrstuhl für Marketing & Media, Universität Hamburg website: https://www.wiso.uni-hamburg.de/projektbuero-angewandte-sozialforschung/aktuelles/2018-09-30-musikdialog/ergebnisse-der-studie-zum-musiknutzungsverhalten.pdf

DeNora, T. (2011). Music in Everyday Life. Cambridge University Press.

Di Lascio, E., Gashi, S., & Santini, S. (2018). Unobtrusive Assessment of Students' Emotional Engagement During Lectures Using Electrodermal Activity Sensors.

Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., 2(3), 103:1–103:21. https://doi.org/10.1145/3264913

Essa, I. A., & Pentland, A. P. (1995). Facial expression recognition using a dynamic model and motion energy. Proceedings of IEEE International Conference on Computer Vision, 360–367. https://doi.org/10.1109/ICCV.1995.466916

Fu, Z., Lu, G., Ting, K. M., & Zhang, D. (2011). A Survey of Audio-Based Music Classification and Annotation. IEEE Transactions on Multimedia, 13(2), 303–319. https://doi.org/10.1109/TMM.2010.2098858

Gaston, E. T. (1951). Dynamic Music Factors in Mood Change. Music Educators Journal, 37(4), 42–44. https://doi.org/10.2307/3387360

Han, B.-J., Rho, S., Jun, S., & Hwang, E. (2010). Music emotion classification and context-based music recommendation. Multimedia Tools Appl., 47(3), 433–460. https://doi.org/10.1007/s11042-009-0332-6

Hussain, S. S., & Bieber, G. (2009). Emotion Recognition on the Go : Providing Personalized Services Based on Emotional States. Proceedings of the 2009 Workshop: Measuring Mobile Emotions: Measuring the Impossible? Presented at the Bonn. Bonn.

IFPI. (2018). Global Music Report - Annual state of the industry. Retrieved from IFPI website: https://www.ifpi.org/downloads/GMR2018.pdf

Jamdar, A., Abraham, J., Khanna, K., & Dubey, R. (2015). Emotion Analysis of Songs Based on Lyrical and Audio Features. International Journal of Artificial Intelligence & Applications, 6(3), 35–50. https://doi.org/10.5121/ijaia.2015.6304

Janssen, J. H., Broek, E. L. van den, & Westerink, J. H. D. M. (2012). Tune in to your emotions: a robust personalized affective music player. User Modeling and User-Adapted Interaction, 22(3), 255–279. https://doi.org/10.1007/s11257-011-9107-7

Kensinger, E. A. (2011). Remembering Emotional Experiences: The Contribution of Valence and Arousal. Reviews in the Neurosciences, 15(4), 241–252. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kim, J., Lee, S., Kim, S., & Yoo, W. Y. (2011). Music mood classification model based on arousal-valence values. 2011 13th International Conference on Advanced Communication Technology (ICACT), 292–295.

Krause, A. E., & North, A. C. (2014). Contextualized music listening: playlists and the Mehrabian and Russell model. Psychology of Well-Being, 4(1), 1–16. https://doi.org/10.1186/s13612-014-0022-7

Lakens, D. (2013). Using a Smartphone to Measure Heart Rate Changes during Relived Happiness and Anger. IEEE Transactions on Affective Computing, 4(2), 238–241. https://doi.org/10.1109/T-AFFC.2013.3

Meschtscherjakov, A., Weiss, A., & Scherndl, T. (2009). Utilizing Emoticons on mobile devices within ESM studies to measure emotions in the field. Proc. MME in Conjunction with MobileHCI, 9.

Nathan, K. S., Arun, M., & Kannan, M. S. (2017). EMOSIC — An emotion based music player for Android. 2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 371–276. https://doi.org/10.1109/ISSPIT.2017.8388671

Nielsen. (2017). Time with Tunes: How Technology is Driving Music Consumption. Retrieved February 26, 2019, from http://www.nielsen.com/us/en/insights/news/2017/time-with-tunes-how-technology-is-driving-music-consumption

Nielsen. (2019). Total Album Equivalent Consumption in the U.S. Increased 23% in 2018. Retrieved February 20, 2019, from http://www.nielsen.com/us/en/insights/news/2019/total-album-equivalent-consumption-in-the-us-increased-23-percent-in-2018

Pettijohn, T. F., Williams, G. M., & Carter, T. C. (2010). Music for the Seasons: Seasonal Music Preferences in College Students. Current Psychology, 29(4), 328–345. https://doi.org/10.1007/s12144-010-9092-8

Picard, R. W., & Klein, J. (2002). Computers that recognise and respond to user emotion: theoretical and practical implications. Interacting with Computers, 14(2), 141–169. https://doi.org/10.1016/S0953-5438(01)00055-8

Reeves, B., & Nass, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places. Stanford, Calif. : New York: The Center for the Study of Language and Information Publications.

RIAA. (2018). Mid-year 2018 RIAA Music Revenues Report. Retrieved from https://www.riaa.com/wp-content/uploads/2018/09/RIAA-Mid-Year-2018-Revenue-Report.pdf

Ricci, F. (2012). Context-aware Music Recommender Systems: Workshop Keynote Abstract. Proceedings of the 21st International Conference Companion on World Wide Web, 865–866. https://doi.org/10.1145/2187980.2188215

Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161–1178. https://doi.org/10.1037/h0077714

Scaringella, N., Zoia, G., & Mlynek, D. (2006). Automatic genre classification of music content: a survey. IEEE Signal Processing Magazine, 23(2), 133–141. https://doi.org/10.1109/MSP.2006.1598089

Sloboda, J. A. (2011). Music in everyday life: The role of emotions. In P. N. Juslin & J. Sloboda (Eds.), Handbook of Music and Emotion: Theory, Research, Applications. Oxford University Press.

Thayer, R. E. (1991). The Biopsychology of Mood and Arousal. Oxford University Press.

Welch, C. (2018, December 6). Alexa can now find the right Amazon Music playlist by having a conversation with you. Retrieved March 2, 2019, from The Verge website: https://www.theverge.com/2018/12/6/18128757/alexa-amazon-music-playlist-conversations-feature

Yamamoto, J., Kawazoe, M., Nakazawa, J., Takashio, K., & Hideyuki, T. (2009). MOLMOD: Analysis of Feelings based on Vital Information for Mood Acquisition. IEICE Technical Report (Institute of Electronics, Information and Communication Engineers), 109(131), 63–68.

Yurur, O., Liu, C. H., Sheng, Z., Leung, V. C. M., Moreno, W., & Leung, K. K. (2016). Context-Awareness for Mobile Sensing: A Survey and Future Directions. IEEE Communications Surveys Tutorials, 18(1), 68–93. https://doi.org/10.1109/COMST.2014.2381246