Directing Physiology and Mood through Music: Validation of an Affective Music Player

ARTICLE *in* IEEE TRANSACTIONS ON AFFECTIVE COMPUTING · JANUARY 2013

Impact Factor: 3.47 · DOI: 10.1109/T-AFFC.2012.28

Marjolein Van Der Zwaag
Philips

Joyce Westerink
Philips

Directing Physiology and Mood through Music: Validation of an Affective Music Player

Marjolein D. van der Zwaag, Joris H. Janssen, and Joyce H.D.M. Westerink

Abstract—Music is important in everyday life, as it provides entertainment and influences our moods. As music is widely available, it is becoming increasingly difficult to select songs to suit our mood. An affective music player can remove this obstacle by taking a desired mood as input and then selecting songs that direct toward that desired mood. In the present study, we validate the concept of an affective music player directing the energy dimension of mood. User models were trained for 10 participants based on skin conductance changes to songs from their own music database. Based on the resulting user models, the songs that most increased or decreased the skin conductance level of the participants were selected to induce either a relatively energized or a calm mood. Experiments were conducted in a real-world office setting. The results showed that a reliable prediction can be made of the impact of a song on skin conductance, that skin conductance and mood can be directed toward an energized or calm state and that skin conductance remains in these states for at least 30 minutes. All in all, this study shows that the concept and models of the affective music player worked in an ecologically valid setting, suggesting the feasibility of using physiological responses in real-life affective computing applications.

Index Terms—Affective music player, music, physiology, skin conductance, skin temperature, mood

1 Introduction

Listening to music is popular as it provides entertainment, is readily available to anyone, can be listened to almost anywhere, and is one of the most popular ways to regulate mood [1], [2], [3]. The ability of music to direct mood is important, as positive moods enhance several cognitive functions [4], cause optimistic feelings to predominate [5], and, among other things, can enhance cognitive revalidation [6]. However, as our personal music databases increase in size, it becomes more difficult to select songs. On top of this, the listener often does not know who the composer or performer of the music is [1]. Therefore, it is not easy to select songs that direct you to a mood that you want at a certain point in time or that is consistent with the tasks that you are doing. To deal with this, music players have been proposed that help to direct affect toward a target state [7], [8]. In the present study, the affective music player (AMP) [7] will be extensively validated to show whether it can direct mood toward a desired state, and thereby facilitate personal music selection.

1.1 Music Listening in Everyday Life

In this section, we will briefly explain some important characteristics of the influence of music on mood. First of all, it is increasingly recognized that responses to music are personal and depend on preference [3], [9]. This personal response to music may depend on several aspects such as age, gender, and culture [1], [3]. To further illustrate this, Cassidy and MacDonald [9] found that self-selected compared to experimenter-selected high- and low-arousal music caused greater enjoyment, reduction of tension, and more efficient performance during a driving game. Therefore, it is important to personalize music selection.

Second, mood is often operationalized in energy and valence dimensions [10]. Energy is an important dimension for mood, and features prominently in different theories on musically evoked mood and task performance (e.g., arousal-based theory [11], [12] and the psychobiological approach [13]). As previous work has already focused specifically on the valence dimension of musical mood induction [7], we focus this work on energy induction.

Third, responses to music include changes in a variety of physiological responses [14]. Physiological changes have repeatedly been used to reflect changes in affect [15], [16], [17]. The physiological responses to musically induced affect are primarily sensitive to a single mood dimension [18]. Skin conductance is a relatively direct (i.e., with a lag of only 2 seconds) reflection of the sympathetic nervous system and therefore has been related to arousal [19]. Consistent with this, skin conductance is most clearly responsive to the energy of the music [14], [20], [21], [22], [23]. For example, higher skin conductance levels have been reported with high percussive, more arousing music compared to low percussive music, low arousing music [24]. The other mood dimension, valence [10], has been repeatedly related to skin temperature (ST) [7], [25], [26], [27], [28]. However, there is no clear agreement on the direction of this relation: Where several authors [20], [26], [29], [30] agree in reporting a lower skin temperature for negatively valenced experiences, others [7], [25] find the

- *M.D. van der Zwaag and J.H.D.M. Westerink are with the Brain, Body, and Behavior Department, Philips Research, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands. E-mail: mvanderzwaag@gmail.com.*
- *J.H. Janssen is with the Healthcare Information Management Department, Philips Research, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands.*

Manuscript received 4 Jan. 2012; revised 19 July 2012; accepted 25 July 2012; published online 28 Aug. 2012.
Recommended for acceptance by S. D'Mello.
For information on obtaining reprints of this article, please send e-mail to: taffc@computer.org, and reference IEEECS Log Number TAFFC-2012-01-0002.
Digital Object Identifier no. 10.1109/T-AFFC.2012.28.

opposite, namely lower skin temperatures for positively valenced music. Lundqvist et al. [25] explain this discrepancy by pointing toward the confounding influence of highly arousing elements in the music, leading to sympathetic stimulation. Indeed, changes in skin temperature do reflect changes in vasoconstriction of the blood vessels in the skin [31] as a consequence of sympathetic arousal.

1.2 Types of Music Selection

Several methods and systems have been described that make song selection easier for a user. These are based on different systems of music recommendation: 1) systems that find similar songs based on subjective input, 2) based on automatically extracted music features, or 3) systems that tag songs with an affect label based on song features [32], [33], [34], [35], [36], [37], [38]. Nevertheless, most of these models do not take into account that affective reactions to music are highly personal [1]. As a result, these music selection systems could be improved by taking the user's affective response to a song into account as well.

A couple of systems do already take into account users' affective responses to music (often based on physiological measurements to enable unobtrusive affect recognition). To start with, systems have been proposed that adjust music characteristics, such as tempo, to intensity or stress levels while exercising [37], [39]. The system of Wijnalda et al. [37] takes into account the heart rate and exercise goal of the user by adjusting the music tempo to the user's heart rate, by adjusting the music tempo to the pace of a runner, or by selecting music with a constant tempo so that the users can synchronize their pace with the music. The system keeps track of heart rate history to predict how the heart rate will progress in the next 30 seconds, which in turn influences the selection of the next song.

The affective remixer system of Chung and Vercoe [32] personalizes music selection by rearranging the order of songs. It does so by taking the users' physiological measurements into account to detect the affective state. It then remixes the music using a predictive music-arrangement system that was trained based on previous physiological responses of various users to a song. If the user wants to be more aroused the music is remixed so that the most arousing parts of the music will be presented first.

The physiology and purpose-aware automatic playlist generation (PAPA) system incorporates physiological responses to determine which song to play next [8]. The PAPA system uses the physiological measurements to find a song with the correct music features to change the physiological responses in the desired direction (which is based on a workout plan preselected by the user). If, for instance, heart rate is lower than the desired goal, a song with a higher tempo than the current song will be selected.

1.3 The Affective Music Player

The music players mentioned so far do not keep track of the personal affective or physiological responses to songs over time. The affectiveDJ [40] is an exception to this and saves the change in skin conductance difference between the start and end of a song. Afterward, the affectiveDJ uses this information to select songs that either relax or arouse the listener. The affectiveDJ failed to perform successfully in a user test, which was attributed to the high song repetition rate in a short period of time together with the fact that the songs were not individually chosen.

The AMP as developed by Janssen et al. [7], [41] is an extended and more sophisticated version of the affectiveDJ [40]. The AMP uses physiological responses as input to personalize music selection and direct the mood toward a target state selected by the user. The AMP thereby uses an affective loop that uses the physiological changes to adjust the music selection [7], [15], [16], [42], [43]. Thus, the behavior of the closed-loop system is described on the level of physiological change, and the AMP immediately selects songs based on the physiological input, without inferring the mood from the physiology. After a song has been listened to, physiological responses of the last minute of listening to the song are normalized using z-transformations to correct for the variance over the last 30 minutes of the session. Next, the song effect is calculated by subtracting the normalized physiological value during the last minute of the previous song from the last minute of the current song. The song effect is then corrected for the law of initial values (LIV, i.e., the idea that the physiological change depends on the prestimulus level [44]). This is done by regressing the physiological values of the song effect over the physiological values of the previous song, and taking residuals [7]. This LIV-corrected song effect is then tagged to the song, and together with the previous tags to the song, a prediction is made on how the song will influence the physiological state of the individual the next time it is played. The models of the AMP are based on accumulating probability density functions (PDFs). These models can naturally deal with noise in the environment on top of the influence of the music. The more a song is listened to, the more reliable the prediction will be of its effect on the physiology. In particular, the AMP can select a song that is known to increase or decrease skin conductance. Similarly, a precise target level might be specified so that a song may be selected that makes the exact required change from the present state [41].

To overcome the fact that a new song in the user's database does not yet have a prediction, a bootstrap needs to be implemented. For example, the user can provide their own bootstrap by tagging the song with a mood label, the prediction of a user or group of users can be adopted, or music features can be used to make a general prediction in advance of how a song might influence a user's physiology [35], [36]. There is no best bootstrap yet, but a combination of user input and music features to predict the mood that a song will induce would probably be most accepted by a user. Such an AMP is currently implemented in a working carrier application (see [45]).

The AMP has certain advantages over the other music players previously mentioned that aim to direct affect. Specifically, the AMP combines a wide range of design considerations based on previous research:

1. It recognizes that music is personal.
2. It incorporates physiological sensors to allow unobtrusive mood measurement and personalization of music selection.

3. It incorporates an affective loop so that the personal information on the user's physiological responses to a song can be updated after each song.
4. It uses probabilistic models to make a prediction of the influence of a song on the user's physiology.
5. By implementing the law of initial values (LIV), it can accommodate for the effect that the influence of a song on physiological state depends on the physiological level at the onset of a song.

1.4 The Present Study

This study aims to further validate the affective music player developed by Janssen et al. [7]. Janssen et al. [7] validated the affective music player for directing valence using skin temperature. For this, they used three participants and 27 songs for each participant. They found that by modeling participants' skin temperature responses to pieces of music, the effect of music on skin temperature could be predicted. These models were subsequently used to select music to direct skin temperature and mood in a predefined direction.

The present study expanded on this validation in the following ways. First, it tested whether the energy rather than the valence of a mood can be directed based on SCL. This is important as energy is an important feature in directing mood. Second, we used 10 participants and included 36 songs for each of them. The songs were selected to direct mood toward neutral, high-energy positive valence (happy), and low-energy negative valence (sad) moods, as happy and sad are the most opposite moods in terms of both valence and energy [10], [15]. Therefore, they result in the largest differences between the states. This is especially important in the present study because the high ecological validity of an office environment in the study also introduces noise in the measurements. The validation process is also very time consuming as models have to be trained before they can be validated. Therefore, the most opposite moods were chosen. Third, tests were done over a longer time (i.e., 30 minutes) than those used by Janssen et al. [7], to investigate whether moods can be induced and maintained over longer time periods. Finally, the SCL and mood were targeted to move in one of three directions: relatively higher SCL, relatively lower SCL, and dynamically changing the SCL to a neutral state (i.e., in between the other two directions) based on real-time analysis.

To our knowledge, there have been no attempts to direct mood in a *dynamic* way by means of song selection based on the evaluation of physiological responses in real time. In a slightly different domain, systems have been evaluated to adapt task difficulty to psychophysiological responses in an adaptive and real-time manner [46], [47]. Haarman et al. [46] determined the task difficulty of a simulated flight on set thresholds of psychophysiological values. If the skin conductance was found to be above this threshold, for instance, the level of turbulence and therefore the difficulty of the test was lowered. Novak et al. [47] used input from several physiological measurements, such as heart rate and skin conductance, and applied an adaptive threshold based on a discriminant analysis of these responses to determine whether the task is too easy or too difficult and then adapt it accordingly. This method of adaptive real-time estimation of the song to be played was also tested with the AMP.

First of all, as part of a training phase it was investigated for each participant individually to what extent happy, neutral, or sad songs can influence SCL in a certain direction each time that a particular song is presented. Based on the results of the training phase, it was predicted how each song would influence SCL the next time it was presented, following the methods of Janssen et al. [7]. Second, in an evaluation phase the songs with the highest prediction to increase, decrease, or to not change SCL were used to validate the affective music player. Accordingly, the songs for one of the following directions (up, down, or no change in SCL) were played successively in one session. Higher SCL was expected with happy songs compared to sad songs. In a dynamic neutral session, SCL was expected to be directed to a neutral state, in between the relatively happy and sad moods, with both happy and sad songs. The selection of the songs in this session was chosen automatically and dynamically. The training phase and the evaluation phase are described separately.

2 Part 1: Training Phase

2.1 Method

2.1.1 Participants and Design

Ten employees (five male/five female, age $\mathrm{M} = 26.5$ years, $\mathrm{SD} = 3.5$ years) of Philips Research received €100 in vouchers for taking part in both the training phase (part 1) and the evaluation phase (part 2) of the experiment. Participants were unaware of the research goals, and signed a written informed consent form before participating. For the training phase a within-subject design was employed so that all participants followed a set of nine sessions listening to 36 songs in each session. During the experiment the order of the songs within each session was counterbalanced based on the nine first rows, one for each session, of a 36×36 digram-balanced Latin square [48]. In this way, each song was played once every session, each song occurred only once in a certain place in the list of songs, and each sequence of two songs in a row did not occur more than once.

2.1.2 Music Stimuli

Music was selected per participant based on a music rating procedure. Participants rated 200 songs that were randomly selected from the participants' own music database. Participants rated their songs in terms of the perceived influence on energy (low energy to high energy) and valence (unpleasant to pleasant) on 7-point Likert scales. During the music rating, participants could listen to the song if they wished. Based on the music ratings, 12 songs were selected for each participant for each of three mood categories, resulting in 36 songs in total. The songs with the highest valence and energy ratings were selected for a happy (high-energy positive valence) category, intermediate valence and energy levels for a neutral category, and songs having the lowest valence and energy values were chosen for a sad (low-energy low valence) category. To prevent the experiment taking too long per session the

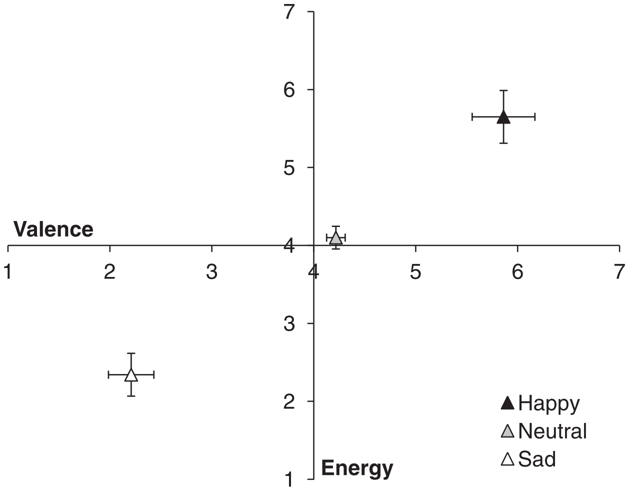


Fig. 1. The average valence and energy ratings for the individually selected songs of the training phase of the experiment.

selected songs were shortened, if necessary, to 4 minutes and faded out at the end of the song using Audacity software (version 1.2.6).

A repeated-measures MANOVA with Mood (Happy, Neutral, Sad), was applied on the valence and energy ratings of the selected songs to verify that songs in the three categories differed in subjective ratings. Multivariate results showed that the selected songs differed significantly in Mood ($F(2,4) = 27.16$, $p = 0.001$, $\eta_p^2 = 0.948$). Univariate tests showed that both valence and energy ratings significantly differed for the three Moods (valence ratings $F(2,18) = 48.70$, $p < 0.001$, $\eta_p^2 = 0.844$; energy ratings $F(2, 18) = 42.19$, $p < 0.001$, $\eta_p^2 = 0.844$). Pairwise comparisons revealed that all the moods (Happy, Neutral, Sad) significantly differed from each other in valence and energy ratings in the expected way (all $p's < 0.001$). Means and SEs are depicted in Fig. 1.

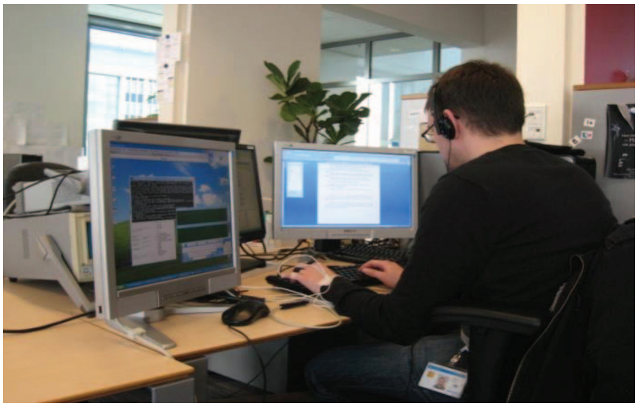
2.1.3 Physiological Measurements

The Nexus-10 (MindMedia B.V., Roermond, the Netherlands) was used to measure the participants' physiological signal of skin conductance. Skin conductance measurements were obtained using dry Ag-AgCl finger electrodes, attached to Velcro strips (sample frequency 128 Hz). The participants strapped the electrodes around the lower phalanxes of the index and middle fingers of the nondominant hand. To make sure the wires would not distract the participant from their work activities, the participant taped the wires of the sensor to the wrist with medical tape.

2.1.4 Procedure

Each of the nine experimental sessions took place while the participants conducted their normal office work activities at their desk. Each participant received one desktop PC to record the physiological measurements and run the experimental sessions. In this way, the participants could easily start a session themselves, and their work on their normal work PC would not be disrupted by the experimental programs. Participants could start a session whenever they liked, with a maximum of one per day. Participants were encouraged to complete a session without too many

(a)



(b)

Fig. 2. Two participants taking part in the experiment during their normal office work activities. The SCL sensors were connected to the fingers of the nondominant hand.

interruptions. Participants were asked not to start a session when they did not feel like doing it. In an introductory session, the participants were shown how to connect the physiological sensors to themselves and how to start the measurement and the music listening. None of the participants felt that the sensors disrupted their work activities, like typing. Fig. 2 shows the setup of the experiment with two of the participants, while they are running an experimental session during their normal work activities.

During each experimental session participants first attached the skin conductance sensor to their fingers. They then started the experiment by clicking on a "start session" icon on the desktop. The physiological recording then began and the first song started to play. The first song was a 2-minute silent period for data analysis which is explained in the data analysis section. Music was listened to via a circum-aural headphone. The participant could press a "break" button if anyone or anything interrupted the music listening. The music listening could be continued by pressing a "continue" button, and the session was resumed from the last minute of the song that was interrupted, and this song was also put at the end of the playlist. This ensured that all the songs were fully played during each session. When all the songs had been played, the session ended and the participants could disconnect the sensors. Lastly, the participants reported in a diary whether there were any circumstances that might have influenced their

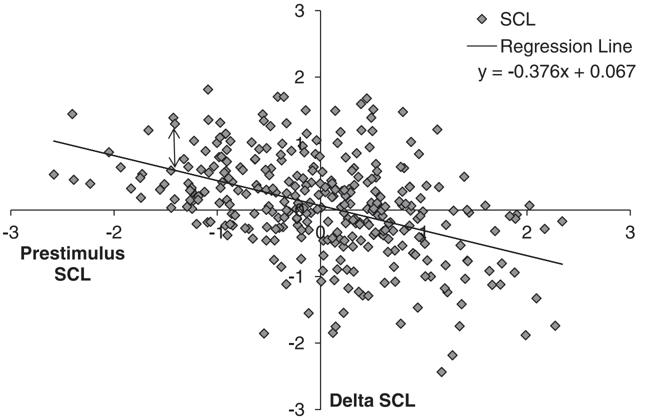


Fig. 3. The delta SCLs (i.e., the normalized SCL of the song minus the normalized SCL of the preceding song) of one participant for all songs played during all sessions of the training phase plotted against the normalized SCLs of the preceding song. The regression line is determined per participant to correct for the LIV. The vertical arrow shows one delta SCL residual indicating the effect of that song corrected for the LIV.

physiological state (e.g., a very short night of sleep, extreme stress). Each session had a maximum duration of 2 hours 22 minutes if no breaks were taken. As the participants were only allowed to conduct one session a day, the total training phase took at least nine days. In practice, participants preferred to spread the sessions over about three weeks, having about three sessions a week.

2.2 Data Preprocessing and Results

The data processing of the training phase followed the method described by Janseen et al. [7], [41]. We have summarized the most important steps below and refer to [7] for specific mathematical descriptions.

The SCL of the last minute of each song was normalized using a z-transformation. For this normalization, the mean and standard deviation of the SCL per session (i.e., listening to 36 songs) were taken. For each song, the average of the last minute of the SCL of the previous song was subtracted from the average of the last minute of the SCL of the song. For the first song in each session, the average of the second minute of the silent period was subtracted from the average of the last minute of the SCL of the first song. This difference in the song (delta SCL) best describes the effect of the song on SCL taking into account the influence of the previous song. The data were then corrected for the initial value of the SCL (according to the Law of Initial Values [44], [49]). Accordingly, first the delta SCLs of all songs were regressed over the SCLs of their preceding songs. The regression line was determined per participant based on the data points of all songs in all sessions, resulting in 324 data points (i.e., 36 songs $\times$ 9 sessions). Then, the delta SCL of the current song was corrected for the regression line by subtracting the value of the regression line, resulting in delta SCL residuals. Fig. 3 shows an example of the regression line of one participant and one residual indicating how the effect of a song is corrected for the LIV. Over the 10 participants, the average slope of the regression line was -0.39 ($SD = 0.12$), and the intercept was 0.04 ($SD = 0.03$).

Next, the delta SCL residuals per song were used to create a probability density function, which was implemented using kernel density estimation (KDE) to make a prediction of the song's effect on the SCL [50]. The KDE fitted a Gaussian distribution to every delta SCL residual of a song and averaged over these distributions (see Fig. 4 for an example of four songs with different impacts on the SCL of one participant). The probability that a song will increase the SCL is represented by the size of the area below the curve in the range >0. The probability that a song will decrease the SCL is represented by the area below the curve in the range <0.

To determine whether a song can increase or decrease the skin conductance level the PDF was divided into three areas. We calculated the probability that a song will increase SCL as the area under the PDF curve for residual sizes >0.25 ($[0.25, \infty>$). Similarly, we calculated the probability that a song will decrease SCL as the area under the PDF for residual sizes below -0.25 ($< \infty, -0.25]$), and the probability for a neutral effect as the area under the PDF curve in between these residual sizes ($< -0.25, 0.25>$). The results showed that per participant, on average, 9.2 songs could be selected within the decrease SCL range, 12.4 songs within the neutral SCL range, and 7.6 songs within the increase SCL range. This already suggests that songs can be used to direct SCL of the listener, but we will explicitly test this in the next phase.

3 Part 2: Evaluation Phase

The aim of this phase of the experiment was to consistently direct skin conductance in a target direction (relatively up/neutral/down) using music that in the training phase of the experiment was demonstrated to consistently influence SCL. We further expect a time-on-task effect. The time-on-task effect is a well-known effect showing that physiological signals tend to change over the course of time irrespective of the experimental conditions [51], [52]. We, therefore, expect that it will be difficult to lower the skin conductance, because following the time-on-task effect we would expect a tonic increase in the SCL. In the "up" condition, we predict that this effect will accelerate the upward trend of SCL; in the "down" condition, this trend will be inhibited; and in the "neutral" condition, SCL will follow the trend.

3.1 Method

3.1.1 Participants and Design

Eight of the 10 participants who took part in the training phase continued with the evaluation phase of the experiment. Two participants dropped out because they were not located at the Philips Research premises during the evaluation phase. In the evaluation phase, we employed a within-subject design in which SCL was directed in each of three directions: relatively up, neutral, or down. During a session, eight songs targeting one mood state (energetic, neutral, calm) were presented. Each mood state was presented six times, resulting in a total of 18 sessions. The order of mood state was counterbalanced over the 18 sessions with the constraint that one target mood state would not immediately follow the same state.

3.1.2 Song Stimuli

Songs were selected per participant where the prediction was that they would direct his/her SCL relatively up,

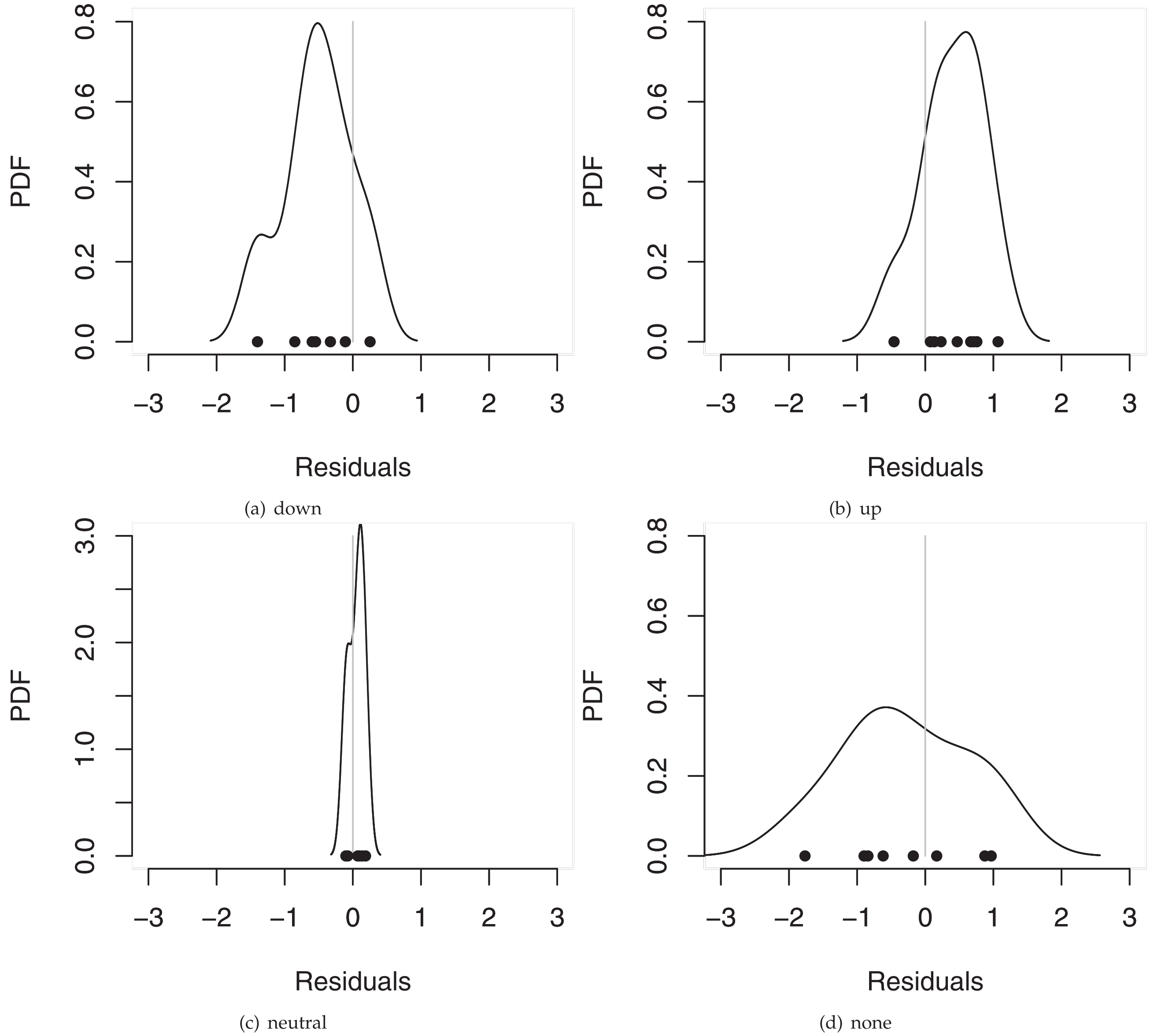


Fig. 4. The probability distribution of four songs of one participant. Each dot in the figures A-D represents the delta residual SCL value of a song when presented once. Panel A shows a song most likely to decrease SCL, panel B shows a song most likely to increase SCL, panel C shows a song most likely to not change the SCL, and panel D shows a song where reliability is low in terms of predicting how the song will influence SCL.

down, or not change the SCL at all. This was done based on the results from the training phase. The eight songs with the greatest probability of increasing or decreasing SCL were selected for the up and down conditions. In addition, eight songs were chosen with the highest probabilities that they would not change the SCL as the mid ranges. The total number of songs used in this phase was 24 (i.e., 3×8) songs. In line with the individual selection of the songs in the training phase, the songs in the evaluation phase were also different for different participants. Table 1 provides the average probability distributions of the eight songs per direction up, down, and neutral.

A repeated-measures MANOVA with Direction (up/neutral/down) was then applied to the valence and energy ratings of these selected songs to determine whether the song ratings (provided at the beginning of the experiment) were different for the different directions. The results showed a significant univariate effect of Direction for the valence ($F(2, 18) = 3.58,\ p = 0.025,\ \eta^2 = 0.285$) and energy ratings ($F(2, 18) = 4.09,\ p = 0.017,\ \eta^2 = 0.312$). Pairwise comparisons show that the ratings in the up condition were significantly higher in energy ratings ($p = 0.015$) and marginally higher in valence ratings ($p = 0.063$) compared to the down ratings, while these scores were marginally higher

TABLE 1
Confusion Matrix of the SCL Direction Probabilities of the Eight Selected Songs per Participant per State for the Evaluation Phase

	Probability directing SCL to interval		
Range	$< \infty, -.25]$	$< -.25, .25 >$	$[.25, \infty >$
Down	**.47**	.38	.15
Neutral	.20	**.57**	.22
Up	.15	.42	**.42**

The values represent the average distributions of the residual values for all participants.

compared to the neutral songs (energy $p = 0.095$; valence $p = 0.080$). The neutral songs did not differ from the down songs. The average valence and energy ratings of the subset of songs selected for the evaluation phase were follows: Mean (Standard Error), Valence Up $M = 4.75(0.24)$, Neutral $M = 4.08(0.23)$, Down $M = 4.02(0.24)$; Energy Up $M = 4.67(0.16)$, Neutral $M = 3.98(0.14)$, Down $M = 3.97(0.265)$.

3.1.3 Materials

The UWIST Mood Adjective Checklist (UMACL) was used to obtain a subjective indication of valence and energy levels [53]. The two mood dimensions concerned were valence (ranging from unpleasant to pleasant) and energy (ranging from tired/without energy to awake/full of energy). The UMACL contains eight unipolar items for each mood dimension, starting with "Right now I am feeling..," and with answers ranging from 1: "not at all" to 7: "very much." Skin conductance and skin temperature measurements were recorded with the Nexus-10 (Mind-Media B.V., Roermond, the Netherlands). Skin temperature was added in this phase to show a possible relationship with mood valence [7]. The skin conductance sensor was attached to the fingers in a similar way as in the training phase. A thermistor skin temperature sensor was attached to the lowest phalanx of the little finger of the nondominant hand using adhesive tape (sample frequency 128 Hz). The temperature sensor records temperature with a resolution of 0.001°C in a range of 10°C to 40°C.

3.1.4 Procedure

Similar to the training phase of the experiment this evaluation phase took place at the participant's own desk, and the participant was asked to listen to music while doing normal office work during several sessions. The participants were unaware of the aim of the experiment; they were solely told that the aim was to collect physiological response measurements when listening to music while working. The participants could start a session whenever they felt like it, with a maximum of two sessions a day. Before each session, the participants attached the skin conductance and skin temperature sensors to their nondominant hand. The participants then started the experiment by pressing the "start session" icon on the desktop of their experiment PC.

Each session consisted of four different stages:

1. a baseline period,
2. a UMACL questionnaire,
3. and SCL direction (relatively up, down, or dynamic neutral), and
4. a UMACL questionnaire.

The baseline period consisted of four of the eight neutral songs, which were counterbalanced over the sessions. The UMACL questionnaires were presented on the computer also used for the music selection and presentation. Each session lasted a maximum of 48 minutes: 16 minutes for the baseline period and 32 minutes for the SCL direction part. On average, the participants took about four weeks to finish this evaluation part of the experiment, having about one session a day.

During the musical SCL direction period, SCL was influenced to change into one of three directions: relatively up, down, or dynamically to a neutral state. In the up or down sessions, the eight songs that were most likely to increase or decrease SCL were presented. The order in which the eight songs were presented followed an 8×8 diagram-balanced Latin square [48]; the last two rows of the Latin square were not used as there were six sessions per mood state. In the dynamic condition, the playlist of eight songs was adjusted in real-time based on the current SCL. The aim of this condition was to direct the SCL to a neutral state. The neutral state for a session was defined as the normalized mean SCL, using z-transformation, of the SCL obtained during the baseline period. The songs in the dynamic condition were equal to the songs of the up and down SCL conditions. The song to be played next was automatically selected based on the average normalized SCL of the last minute of the preceding song (i.e., if the SCL was higher than the neutral state a song for decreasing SCL was chosen, otherwise a song for increasing SCL was selected). In order to start this condition in the same way each time, the first song was always an increasing SCL song; this was an arbitrary choice and could equally have been a down song.

3.2 Data Processing

3.2.1 Subjective Data

Chronbach's alpha values for both the valence and energy items of the UMACL questionnaire were larger than 0.8, which indicated good reliability of the two mood dimension scores. Next, mood reaction scores were calculated by subtracting the ratings obtained after the baseline period of a session from the ratings obtained after the mood induction of the same session.

3.2.2 SCL and ST Data

As with the data processing of the training phase data, the average SCL of the last minute of each song was calculated. These values were then normalized with z-transformations using the mean and standard deviation of the SCL obtained during the entire duration of the baseline period. The average SCL was calculated per Direction (up/down/dynamic neutral) over the six sessions of each mood direction per participant. The data processing of the ST data was similar to that of the SCL data. There was missing data for two participants, due to equipment failure: For one session of a participant, the SCL for one song in the up and dynamic conditions was missing, and for one session of another participant, data from one song in the down and dynamic conditions were missing. The missing data points were replaced by average residuals of the song.

3.3 Results

3.3.1 Subjective Data

The average values during the baseline period were 4.64 ($SE = 0.069$) for valence and 3.97 ($SE = 0.067$) for energy out of a possible score of 7. To investigate whether the mood direction influenced the subjective mood ratings, a one-tailed repeated-measures MANOVA with Direction (up, down, dynamic) as the within-subject variable was

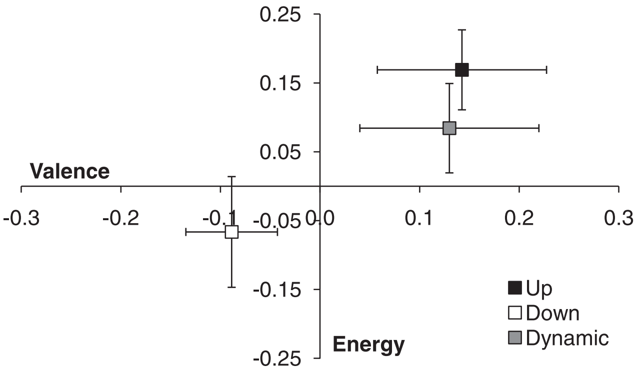


Fig. 5. The average energy and valence reaction scores over all participants and sessions in the up, down, and dynamic neutral conditions. Error bars represent standard errors.

applied to the valence and energy reaction scores. Data were analyzed at the subject level (and not at the session level), so they were averaged over sessions. We used one-tailed tests because we had clear expectations on the direction of the effects. Significant main effects were found for Direction on valence ($F(2,14) = 3.145$, $p = 0.037$, $\eta^2 = 0.310$) and a marginally significant effect on energy ($F(2,14) = 2.72$, $p = 0.056$, $\eta^2 = 0.280$). Pairwise comparisons of the main effect showed significantly higher energy levels for the up condition compared to the down condition ($p = 0.009$). The energy levels of the dynamic condition were not significantly different from the down or up conditions (all $p > 0.10$). Furthermore, the valence levels of the down condition were significantly lower compared to the up ($p = 0.041$) and dynamic ($p = 0.047$) conditions. The valence levels of the up and dynamic conditions did not significantly differ ($p > 0.10$). Means and standard errors are given in Fig. 5 and in Table 2.

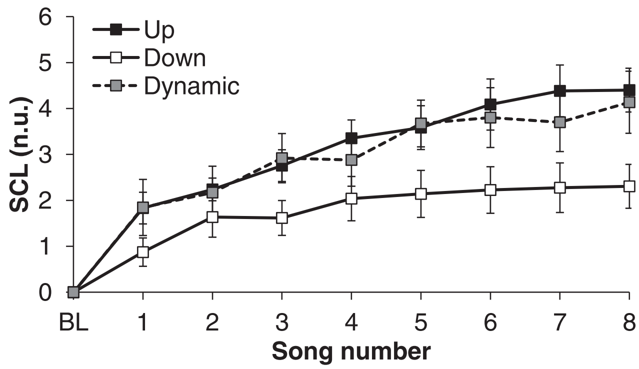
3.3.2 SCL and ST Data

To verify whether the SCL differed for the three mood directions and to see how this pattern evolved over time, a one-tailed repeated-measures ANOVA was applied to SCL data with Song number (SN) (1, 2, 3, 4, 5, 6, 7, 8) and Direction (up/down/dynamic) as the within-subject factors. Data were analyzed at the subject (and not at the session) level, so they were averaged over sessions. The results showed significant univariate main effects of Direction ($F(2,14) = 6.39$, $p = 0.006$, $\eta^2 = 0.48$), Song number ($F(7,49) = 21.32$, $p < 0.001$, $\eta^2 = 0.75$), and an interaction between Direction and Song number ($F(14,98) = 1.95$, $p = 0.015$, $\eta^2 = 0.22$). Pairwise comparisons of the Direction with Song number interaction showed that the

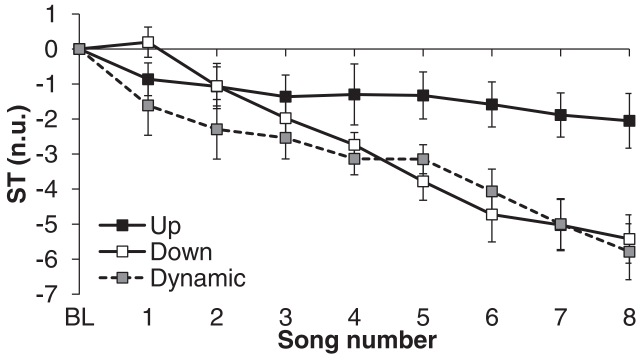
TABLE 2
The Mean Valence and Energy Reaction Scores, and the Means of the Eight Songs of the Normalized Skin Conductance Levels and Skin Temperature for Each Condition (Up, Down, Dynamic Neutral)

Condition	Valence	Energy	SCL	ST
Down	-.09 (.05)	-.07 (.08)	2.31 (.48)	-5.43 (.69)
Dynamic	.13 (.09)	.08 (.07)	4.14 (.68)	-5.79 (.80)
Up	.14 (.09)	.17 (.06)	4.40 (.48)	-2.05 (.78)

Standard errors are indicated between brackets.



(a) SCL



(b) ST

Fig. 6. The average SCL and ST values for each song number for the three mood directions (up, down, dynamic neutral). Error bars represent standard errors, BL = baseline.

SCL in the up condition was significantly higher compared to the down condition from the first song onward (all $p < 0.05$), with the exception of the second song. During the dynamic condition, the SCL was marginally significantly higher than in the down condition from the fifth song onward (all $p < 0.05$). The SCL did not differ significantly between the up and the dynamic conditions. Fig. 6 and Table 2 show the mean SCL averaged over the participants and sessions for each mood direction.

In the dynamic session, on average, 20.5 percent of the song choices included a song directing the SCL upward. The first song choice was not included in this estimate as the choice of the first song was set by default to increase the SCL. Table 3 shows the percentages of the song choices to increase or decrease the SCL for each decision moment averaged over all participants.

To investigate whether the ST differed between the three conditions and to see how this pattern evolved over time, a one-tailed repeated-measures ANOVA was applied to the ST data with Song number (1, 2, 3, 4, 5, 6, 7, 8) and Direction (up/down/dynamic) as the within-subject factors. The results showed significant univariate main effects of Direction ($F(2,14) = 6.96$, $p = 0.004$, $\eta^2 = 0.50$), Song number ($F(7,49) = 13.26$, $p < 0.001$, $\eta^2 = 0.65$), and an interaction between Direction and Song number ($F(14,98) = 4.13$, $p < 0.001$, $\eta^2 = 0.37$). Pairwise comparisons of the Direction with Song number interaction showed that the up condition has higher ST compared to the down condition from the fifth song onward (all $p < 0.010$). The ST was significantly higher in the up condition compared to the

TABLE 3
The Real-Time Song Choices per Song Number Show How Often (as a Percentage) a Song Was Chosen to Either Increase or Decrease the SCL to a Dynamic Neutral State

SN	1	2	3	4	5	6	7	8	average
Up	100	25.00	25.00	16.67	20.83	18.75	16.67	20.83	20.54
Down	0	75.00	75.00	83.33	79.17	81.25	83.33	79.17	79.46

dynamic condition from the second minute onward (all $p < 0.05$). The ST did not differ significantly between the down and the dynamic conditions. Fig. 6 and Table 2 show the ST values for each condition.

To obtain the relationship between the subjective mood ratings of the three conditions and the physiological responses, Pearsons' correlations (r) were calculated. The average SCL and ST data obtained during the eighth song were used as input ($N = 18$). The results showed a significant correlation between SCL and energy ratings (energy $r = 0.362$, $p = 0.041$; valence $r = 0.174$, $n.s.$). The ST shows marginally significant correlations with both the valence and energy ratings (energy $r = 0.330$, $p = 0.057$; valence $r = 0.292$, $p = 0.083$). The valence and energy ratings were furthermore positively correlated with each other ($r = 0.488$, $p = 0.008$). All in all, SCL was mostly related to energy, and ST to both valence and energy.

4 Discussion

This user test evaluated the feasibility of the affective music player [7] in an actual office environment. Participants first listened to a variety of songs taken from their own music databases during normal office activities while skin conductance (SCL) was continuously measured. Based on the output of this training phase, a prediction of the influence of each song on the SCL was made.

The results indicated that songs can predictably influence skin conductance level even in a noisy environment when music is listened to in the background. The evaluation phase showed that SCL was higher when aimed at relatively higher SCL levels and lower when aimed at relatively lower SCL levels. In addition, the current results indicate that the difference between the two directions lasts for over 30 minutes. This is a new finding because this long a duration of musical induction (over 30 minutes) has not been reported before. The subjective mood ratings validated that higher energy levels coincide with higher SCLs and lower energy levels with lower SCLs, which is also reflected in the significant correlation between SCL and and energy ratings. In the dynamic neutral situation, we aimed to keep the SCL measure in between the lower and higher SCL conditions. The results showed that the SCL mostly increased in this condition, and therefore songs to decrease the SCL were automatically selected most often.

The effects on ST are largely congruent with those on SCL. Songs selected for lower SCL also decreased ST more than songs selected for higher SCL. Since the former tended to also decrease the mood valence more than the latter, this result is in line with the part of literature that attributes lower skin temperatures to negatively valenced music [20], [26], [28], [30]. And as a consequence, it does not comply with the results of others claiming the opposite [7], [25]. According to Lundqvist et al. [25], such a decrease in skin temperature is linked to vasoconstriction of blood vessels as a consequence of sympathethic arousal. The present energetic arousal measurements, however, do not point in that direction, since the music selected for the down condition was found to be even slightly lower in energetic arousal (see Fig. 5). Indeed, the correlations we found between skin temperature on the one hand and energy or valence mood ratings on the other are both only marginally significant. Possibly yet other mood dimensions, such as the dimension of tense arousal as described by Matthews et al. [53], play a role here.

The dynamic condition in the evaluation phase shows a different pattern for SCL and ST. In this dynamic condition, the SCL followed the pattern of the up condition, whereas for the ST, it followed the pattern of the down condition, decreasing considerably. This different result of directing physiology in the dynamic condition might be explained by the playlist in this condition. In this playlist, songs with different characteristics were alternated, i.e., with extreme differences from each other. This might have led to some stress, as reflected in higher SCL, as well as to lower skin temperature because of stress-driven vasoconstriction of the skin blood vessels. These results do not necessarily indicate that it is impossible to direct SCL to a neutral state. That might still be possible when only neutral songs are presented, as opposed to alternating songs in a dynamic neutral state. This might then lead to a neutral state, but that was not tested in the current experiment.

4.1 Limitations and Further Research

We have to point out that although the mood difference scores in the evaluation phase did significantly differ between the up and down conditions, their absolute differences were only small. Therefore, it can be questioned whether the mood changes found are relevant. The results of the present study indicated, however, that the changes in mood did correlate with changes in SCL. Hence, even though changes in mood were maybe not always conscious to the user, the bodily changes validated that music had an effect on the user. Future research could investigate under which conditions also the mood as experienced by the listener is influenced to a greater extent. This might be expected in a setting in which the music listened to as a primary task or in the background of a less involving task than office work.

The valence and energy ratings of the songs selected for the evaluation phase were closer to each other than the ratings from the initial song set in the training phase. This implies that the initial ratings provided by the participants themselves were not necessarily the same as the extent to which they induce moods. This is in accordance with the literature which indicates that what is mostly rated is the affect expressed by the music, which is not necessarily the same as the mood induced by that music [1], [18]. This underlines the added value of the affective music player; song ratings and people's own reflections on the influence of a song on their mood are not always accurate. This gap can, to a certain extent, be closed by the AMP.

The evaluation phase of the experiment tested the influence of a playlist of songs that aimed to either relatively increase or decrease SCL. First of all, the SCL increased in every session, including the down session, compared to the baseline. This effect could most possibly be assigned to a time-on-task effect. The time-on-task effect is a well-known effect showing that physiological signals, mostly in cardiovascular studies, tend to change over the course of time irrespective of the experimental conditions [51], [52]. From the SCL data obtained during the baseline period (four neutral songs), it also appeared that the SCL gradually increased over time during this period. Nevertheless, irrespective of this effect, the results showed that after two songs the SCL was significantly higher in the up compared to the down condition. The results also show that with these same playlists the mood, as measured by valence and energy, was also directed toward slightly, but significantly different states. Increased energy levels when SCL was higher are consistent with previous research which has shown that SCL is generally related to arousal [54]. In a similar vein, the skin temperature decreased during the course of the evaluation phase irrespective of the mood direction. This effect can be attributed to the small amount of movement of the fingers during desk work, which cools the lateral parts of the body including the fingers.

The skin temperature showed differentiation between the two induced moods from the fifth song onward, and thus reacts more slowly to music mood induction than SCL. Janssen et al. [7] showed that the skin temperature differentiated between up and down from the fourth song onwards. Although as discussed before, their differentiation is in the opposite direction, it appears that at least on the timing of the skin temperature reactions to music, their results are more or less in line with the present ones.

Future research can further extend our findings in several ways. First, by integrating SCL (mostly sensitive to arousal) and ST (also sensitive to valence) the extent to which the affective music player can direct mood toward each quadrant of the valence-energy model could be expanded and tested. Although it might not seem very useful to be able to direct toward a negative mood in an office context, we believe it should be left up to the user to choose the target mood. As such, it is probably a good idea if the AMP has the ability to direct to different moods as much as possible. It might very well be that users want the target mood to match (to an extent) their own moods. Furthermore, it is also known that negative moods might be useful in certain workplace contexts, because in some situations they can motivate to work harder [5]. Hence, we think it can also be useful for the AMP to be able to direct to more negative states instead of only to positive states. Finally, it remains to be seen to what extent valence and energy can actually be distinguished in musical mood induction, and if there are songs that direct to all four quadrants of the valence-energy model.

Second, the scope of the context could be extended to environments that have been shown to benefit from the influence of music on mood, for example, while driving a car [55], [56], [57], or to reduce patient anxiety [58], [59].

Third, while the affective music player is designed to improve its performance each time a song is played, a bootstrap is needed for new listeners or for new songs. This bootstrap could be explored but could, for example, also be obtained from music features that predict a song's effect [24], [35]: If song B is new and similarity algorithms based on audio features indicate that it is similar to song A which is already known to the user, then song B can be initiated with the same prediction as song A. Furthermore, we believe that affective responses to music can change over time. With the model we propose, it is possible to keep updating the model every time a song is listened to. By also incorporating a mechanism that discards data points that were gathered further into the past, or including a weighing based on the time from moment of listening to now, this could be incorporated in the system.

4.2 Conclusion

To sum up, the present experiment tested the concept of the affective music player. In the training phase, a variety of songs were presented so that an individual prediction of the influence of a song on the SCL could be made per participant. Successively, the results were validated in the evaluation phase. The results showed that both the SCL and the subjective mood ratings were directed relatively up or down as expected. All in all, the current results suggest the feasibility of the affective music player in a highly ecologically valid setting of office work. As such, our study provides a promising outlook for the feasibility of other physiologically driven affect-oriented systems. Undoubtedly, there will be some challenges to overcome in the development of the affective music player, such as the need for unobtrusive physiological measuring equipment. However, as unobtrusive affect monitoring devices are developing rapidly [60], [61], it is possible that an affective music player will become commercially available in the future.

To conclude, music is an important means of regulating mood in various everyday situations. It is readily available to everyone and can be listened to almost anywhere. It is often used to direct mood and there is, therefore, extensive application potential for music players that support this ability to direct mood. Such devices need to incorporate unobtrusive affect measurements and personalize music selection accordingly to be able to direct mood to a target state. The extensive user study described in this paper validated the concept of the affective music player and shows that music can be used to reliably direct mood toward a desired state during office work.

Acknowledgments

The work of this study was supported by the European REFLECT project (FP7-215893). Furthermore, they would like to thank Karel Brookhuis, Dick de Waard, Ben Mulder, and Egon van den Broek for sharing their thoughts on the study, and Gert-Jan de Vries and Marten Pijl for their the assistance with the preprocessing of the data.

References

[1] P.N. Juslin and J. Sloboda, *Handbook of Music and Emotion: Theory, Research, Applications.* Oxford Univ. Press, 2010.

[2] J.J. Gross, *Handbook of Emotion Regulation.* Guilford Press, 2007.

[3] A.C. North and D.J. Hargreaves, *The Social and Applied Psychology of Music.* Oxford Univ. Press, 2008.

[4] A.M. Isen, "Positive Affect and Decision Making," *Handbook of Emotions,* second ed., C.R. Snyder and S. Lopez, eds., pp. 417-435, Oxford Univ. Press, 2000.

[5] G.H.E. Gendolla, "On the Impact of Mood on Behavior: An Integrative Theory and a Review," *Rev. of General Psychology,* vol. 4, pp. 378-408, 2000.

[6] T. Lesiuk, "The Effect of Preferred Music on Mood and Performance in a High-Cognitive Demand Occupation," *J. Music Therapy,* vol. 47, no. 2, pp. 137-154, 2010.

[7] J.H. Janssen, E.L. Van den Broek, and J.H.D.M. Westerink, "Tune in to Your Emotions: A Robust Personalized Affective Music Player," *User Modeling and User Adapted Interaction,* vol. 22, pp. 25-279, 2012.

[8] N. Oliver and L. Kreger-Stickles, "Papa: Physiology and Purpose-Aware Automatic Playlist Generation," *Proc. Seventh Int'l Conf. Music Information Retrieval,* pp. 250-253, 2006.

[9] G.G. Cassidy and R.A.R. MacDonald, "The Effects of Music Choice on Task Performance: A Study of the Impact of Self-Selected and Experimenter-Selected Music on Driving Game Performance and Experience," *Musicae Scientiae,* vol. 13, pp. 357-386, 2009.

[10] P. Wilhelm and D. Schoebi, "Assessing Mood in Daily Life," *European J. Psychological Assessment,* vol. 23, no. 3, pp. 258-267, 2007.

[11] V. Konečni, "Social Interaction and Musical Preference," *The Psychology of Music,* D. Deutsch, ed., pp. 497-516, Academic Press, 1982.

[12] R.M. Yerkes and J.D. Dodson, "The Relation of Strength of Stimulus to Rapidity of Habit-Formation," *J. Comparative Neurology and Psychology,* vol. 18, pp. 459-482, 1908.

[13] D.E. Berlyne, *Aesthetics and Psychobiology.* Appleton-Century-Crofts, 1971.

[14] D. Hodges, "Psycho-Physiological Measures," *Handbook of Music and Emotions: Theory, Research, Applications,* second ed., P.N. Juslin and J.A. Sloboda, eds., chapter 11, pp. 279-311, Oxford Univ. Press, 2010.

[15] R.A. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," *IEEE Trans. Affective Computing,* vol. 1, no. 1, pp. 18-37, Jan. 2010.

[16] J.H. Janssen, J.N. Bailenson, W.A. IJsselsteijn, and J.H.D.M. Westerink, "Intimate Heartbeats: Opportunities for Affective Communication Technology," *IEEE Trans. Affective Computing,* vol. 1, no. 2, pp. 72-80, July-Dec. 2010.

[17] R.W. Picard, *Affective Computing.* MIT Press, 1997.

[18] M.D. Van der Zwaag, J.H.D.M. Westerink, L.J.M. Mulder, D. de Waard, and K.A. Brookhuis, "Sustaining Moods with Background Music: Physiological Patterns and Subjective Experience," submitted for publication.

[19] J.T. Cacioppo, L. Tassinary, and G.G. Berntson, *Handbook of Psychophysiology.* Cambridge Univ. Press, 2007.

[20] C.L. Krumhansl, "An Exploratory Study of Musical Emotions and Psychophysiology," *Canadian J. Experimental Psychology,* vol. 51, pp. 336-353, 1997.

[21] I. Nyklíček, J.F. Thayer, and L.J.P. Van Doornen, "Cardiorespiratory Differentiation of Musically-Induced Emotions." *J. Psychophysiology,* vol. 11, pp. 304-321, 1997.

[22] L. Venables and S.H. Fairclough, "The Influence of Performance Feedback on Goal-Setting and Mental Effort Regulation," *Motivation and Emotion,* vol. 33, pp. 63-74, 2009.

[23] M.D. Van der Zwaag and J.H.D.M. Westerink, "Physiological Patterns during Music Mood Induction," submitted for publication.

[24] M.D. Van der Zwaag, J.H.D.M. Westerink, and E.L. Van den Broek, "Emotional and Psychophysiological Responses to Tempo, Mode, and Percussiveness," *Musicae Scientiae,* vol. 15, no. 2, pp. 250-269, 2011.

[25] L.-O. Lundqvist, F. Carlsson, P. Hilmersson, and P.N. Juslin, "Emotional Responses to Music: Experience, Expression, and Physiology," *Psychology of Music,* vol. 37, pp. 61-90, 2009.

[26] R.A. McFarland, "Relationship of Skin Temperature Changes to the Emotions Accompanying Music," *Biofeedback and Self-Regulation,* vol. 10, pp. 255-267, 1985.

[27] S. Rimm-Kaufman and J. Kagan, "The Psychological Significance of Changes in Skin Temperature," *Motivation and Emotion,* vol. 20, pp. 64-78, 1996.

[28] R.A. McFarland and R. Kadish, "Sex Differences in Finger Temperature Response to Music," *Int'l J. Psychophysiology,* vol. 11, pp. 295-298, 1991.

[29] B. Mittelmann and H. Wolff, "Emotions and Skin Temperature: Observations on Patients during Psychotherapeutic (Psychoanalytic) Interviews," *Psychosomatic Medicine,* vol. 5, pp. 211-231, 1943.

[30] U.M. Nater, E. Abbruzzese, M. Krebs, and U. Ehlert, "Sex Differences in Emotional and Psychophysiological Responses to Musical Stimuli," *Int'l J. Psychophysiology,* vol. 62, pp. 300-308, 2006.

[31] E.N. Marieb and K. Hoehn, *Human Anatomy & Physiology,* seventh ed. Pearson Education, 2007.

[32] J.-W. Chung and G.S. Vercoe, "The Affective Remixer: Personalized Music Arranging," *Proc. Extended Abstracts on Human Factors in Computing Systems (CHI '06),* pp. 393-398, 2006.

[33] J. Kim and E. André, "Emotion Recognition Based on Physiological Changes in Music Listening," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 30, no. 12, pp. 2067-2083, Dec. 2008.

[34] J. Skowronek, M.F. McKinney, and S. Van de Par, "Ground Truth for Automatic Music Mood Classification," *Proc. Seventh Int'l Conf. Music Information Retrieval (ISMIR),* pp. 395-396, 2006.

[35] J. Skowronek, M.F. McKinney, and S. Van de Par, "A Demonstrator for Automatic Music Mood Estimation," *Proc. Eighth Int'l Conf. Music Information Retrieval (ISMIR),* pp. 345-346, 2007.

[36] M.D. Van der Zwaag, J.H.D.M. Westerink, and E.L. Van den Broek, "Deploying Music Characteristics for an Affective Music Player," *Proc. IEEE Int'l Conf. Affective Computing & Intelligent Interaction,* J. Cohn, A. Nijholt, and M. Pantic, eds., pp. 459-465, 2009.

[37] G. Wijnalda, S. Pauws, F. Vignoli, and H. Stuckenschmidt, "A Personalized Music System for Motivation in Sport Performance," *IEEE Pervasive Computing,* vol. 4, no. 3, pp. 26-32, July/Sept. 2005.

[38] S. Zhang, Q. Tian, S. Jiang, Q. Huang, and W. Gao, "Affective MTB Analysis Based on Arousal and Valence Features," *Proc. IEEE Int'l Conf. Multimedia and Expo (ICME '08),* pp. 1369-1372, 2008.

[39] N. Oliver and F. Flores-Mangas, "MPTrain: A Mobile, Music and Physiology-Based Personal Trainer," *Proc. Eighth Conf. Human-Computer Interaction with Mobile Devices and Services,* pp. 21-28, 2006.

[40] J.A. Healey, R.W. Picard, and F. Dabek, "A New Affect-Perceiving Interface and Its Application to Personalized Music Selection," *Proc. Workshop Perceptual User Interfaces (PUI),* M. Turk, ed., pp. 2-5, 1998.

[41] J.H. Janssen, E.L. Van den Broek, and J.H.D.M. Westerink, "Personalized Affective Music Player," *Proc. IEEE Int'l Conf. Affective Computing and Intelligent Interaction (ACII),* pp. 472-477, 2009.

[42] K. Höök, "Affective Loop Experiences: What Are They?" *Proc. Third Int'l Conf. Persuasive Technology,* pp. 1-12, 2008.

[43] K. Höök, "Affective Loop Experiences: Designing for Interactional Embodiment," *Philosophical Trans. Royal Soc. B,* vol. 364, pp. 3585-3595, 2009.

[44] R. Geenen and F.J.R. Van de Vijver, "A Simple Test of the Law of Initial Values," *Psychophysiology,* vol. 30, pp. 525-530, 1993.

[45] A. Schroeder, M.D. Van der Zwaag, and M. Hammer, "A Middleware Architecture for Human-Centred Pervasive Adaptive Applications," *Proc. IEEE Int'l Conf. Self-Adaptive and Self-Organizing Systems,* vol. 1, pp. 138-143, 2008.

[46] A. Haarmann, W. Boucsein, and F. Schaefer, "Combining Electrodermal Responses and Cardiovascular Measures for Probing Adaptive Automation during Simulated Flight," *Applied Ergonomics,* vol. 40, no. 6, pp. 1026-1040, 2009.

[47] D. Novak, M. Mihelj, J. Ziherl, A. Olensek, and M. Munih, "Measuring Motor Actions and Psychophysiology for Task Difficulty Estimation in Human-Robot Interaction," *Proc. Seventh Int'l Conf. Methods and Techniques in Behavioral Research (MB '10),* pp. 1-4, 2010.

[48] W.A. Wagenaar, "Note on the Construction of Digram-Balanced Latin Squares," *Psychological Bull.,* vol. 72, no. 6, pp. 384-386, 1969.

[49] J. Wilder, *Stimulus and Response: The Law of Initial Values.* Wright, 1967.

[50] B.W. Silverman, *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, 1986.

[51] S.H. Fairclough and L. Venables, "Prediction of Subjective States from Psychophysiology: A Multivariate Approach," *Biological Psychology,* vol. 71, pp. 100-110, 2006.

[52] L.J.M. Mulder, C. Dijksterhuis, A. Stuiver, and D. De Waard, "Cardiovascular State Changes during Performance of a Simulated Ambulance Dispatchers' Task: Potential Use for Adaptive Support," *Applied Ergonomics,* vol. 40, no. 6, pp. 965-77, 2009.

[53] G. Matthews, D.M. Jones, and A.G. Chamberlain, "Refining the Measurement of Mood: The UWIST Mood Adjective Checklist," *British J. Psychology,* vol. 81, no. 1, pp. 17-42, 1990.

[54] W. Boucsein, *Electrodermal Activity,* first ed. Plenum Press, 1992.

[55] M.D. Van der Zwaag, S.H. Fairclough, E. Spiridon, and J.H.D.M. Westerink, "The Impact of Music on Affect during Anger Inducing Drives," *Proc. Fourth Int'l Conf. Affective Computing and Intelligent Interaction,* 2011.

[56] M.D. Van der Zwaag, C. Dijksterhuis, D. De Waard, L.J.M. Mulder, J.H.D.M. Westerink, and K.A. Brookhuis, "The Influence of Music on Mood and Performance While Driving," *Ergonomics,* vol. 55, no. 1, pp. 12-22, 2012.

[57] M.D. Van der Zwaag, J.H. Janssen, C. Nass, J.H.D.M. Westerink, S. Chowdhury, and D. De Waard, "Using Music to Regulate Mood While Driving," submitted for publication.

[58] M.D. Van der Zwaag, T. Tijs, J.H.D.M. Westerink, and R. Molegraaf, "Personalized Relax Music to Reduce Patient Anxiety," *Proc. 32nd Int'l Conf. Stress and Anxiety Research Soc. (STAR),* P. Buchwald, K. Moore, and T. Ringeisen, eds., 2011.

[59] M.D. Van der Zwaag, T. Tijs, and J.H.D.M. Westerink, "Personalized Calming Music to Reduce Patient Anxiety," *Proc. Int'l Conf. Music Health and Wellbeing, Current Developments in Research and Practice,* M. Shipton and E. Himonides, eds., pp. 211-212, 2011.

[60] M. Ouwerkerk, "Unobtrusive Emotions Sensing in Daily Life," *Sensing Emotions. The Impact of Context on Experience Measurements. Philips Research Book Series,* 12th ed., J.H.D.M. Westerink, M. Krans, and M. Ouwerkerk, eds., pp. 21-39, Springer, 2011.

[61] M.-Z. Poh, N.C. Swensen, and R.W. Picard, "A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity," *IEEE Trans. Biomedical Eng.,* vol. 57, no. 5, pp. 1243-1252, 2010.

Marjolein D. van der Zwaag received the MSc degree in artificial intelligence in 2007 from the Radboud University Nijmegen, The Netherlands, and the PhD degree from the University of Groningen in April 2012. She is currently a researcher in the Brain, Body, and Behavior Department of Philips Research, Eindhoven, The Netherlands.

Joris H. Janssen received the BSc and MSc (cum laude) degrees in artificial intelligence from the Radboud University Nijmegen, The Netherlands, and the PhD (cum laude) degree in human technology interaction from Eindhoven University of Technology, The Netherlands. He was a visiting scholar at Stanford University for a year and is currently a researcher at Philips Research. His work appeared in about 30 peer-reviewed conference and journal articles as well as book chapters and five patent applications.

Joyce H.D.M. Westerink studied physics at Utrecht University (MSc degree, 1985) and received the PhD degree from Eindhoven University of Technology in 1991 as part of her work at Philips Research, and has remained active in that field ever since. She then joined Philips Research in Eindhoven, The Netherlands, and specialized in human perception and interaction with technology.