Article

Musicae Scientiae
2015, Vol. 19(4) 412–432

DOI: 10.1177/1029864915606796
msx.sagepub.com

Investigating emotional responses to self-selected sad music via self-report and automated facial analysis

Karim Weth
University of Salzburg, Austria and University of Bamberg, Germany

Marius H. Raab
University of Bamberg, Germany and Forschungsgruppe EPÆG, Germany

Claus-Christian Carbon
University of Bamberg, Germany and Forschungsgruppe EPÆG, Germany

Abstract

People often listen to sad music in spite of its seemingly negative qualities. Sad music, and especially sad music with a personal significance, has been shown to evoke a wide span of emotions with both positive and negative qualities. We compared emotional responses to familiar self-selected sad music (SSSM) with both unfamiliar sad and unfamiliar happy music. Alongside self-reports, a commercial, continuous measure of discrete facial expressions was applied, promising an in-depth assessment of both the quality and strength of experienced affective states at any given point in time. Results of the facial analysis showed that SSSM evoked more mixed affective states than unfamiliar sad music. Also, listeners reacted with consistent facial expressions to distinct musical events, e.g. the introduction of a lead voice. SSSM evoked more self-reported feelings of nostalgia, reminiscence, being moved, and chills and tears than unfamiliar sad and happy music. Furthermore, SSSM resulted in more self-reported happiness and a similar trend with happy facial expressions compared to unfamiliar sad music. These results point to the emotional diversity and the strong involvement of positive affective states elicited by SSSM, even when compared with music of similar quality, such as unfamiliar sad music. Automated facial analysis allows us to observe emotions on a more detailed level in terms of time resolution, onset, intensity and concurrence of discrete affective states. This technique is promising for future research, particularly when investigating mixed emotions and the social aspect of emotions in response to music.

Keywords

affective responses, basic emotion, continuous evaluation procedure (CEP), emotions, FaceReader, FACS, facial expressions, mixed emotions, nostalgia, sad music, self-report

Corresponding author:
Karim Weth, Department of Psychology, University of Salzburg, Hellbrunnerstraße 34, Salzburg, Austria.
Email: karim.weth@gmail.com

Sadness is an emotional state characterized by negative valence and low physiological arousal, often accompanied by changes in experience, physiology, cognition, and behavior (Clore & Huntsinger, 2007; Lench, Flores, & Bench, 2011). It is often evoked by the loss of social bonds, or by failure in achieving important life goals. The relationship between everyday and music-induced sadness is unclear, as, for example, in an everyday context, sadness is regarded as an aversive state, whereas music-induced sadness attracts many listeners (Huron, 2011; Zentner, Grandjean, & Scherer, 2008).

Several studies investigated listeners' reactions to their preferred sad music, often labeled as *self-identified sad music* (SISM) or *self-selected sad music* (SSSM). Vuoskoski and Eerola (2012) let participants listen to both unfamiliar and self-selected sad music, and concluded that both are capable of inducing genuine sadness – as measured by indirect techniques, namely word recall and judgement tasks. In contrast, other studies (e.g., Taruffi & Koelsch, 2014; Van den Tol & Edwards, 2013, 2015) used retrospective approaches; that is, participants did not listen to their music selection during testing, but reported their experiences retrospectively. These studies showed that emotional responses are not limited to mere sadness, but are actually more complex. For example, in an international online survey (n = 772), Taruffi and Koelsch (2014) found that listeners' emotional experiences with sad music consist of both positive and negative emotions, or emotions with mixed qualities: In fact, the four most common reactions, as measured with the *Geneva Emotional Music Scale* (GEMS, see Zentner et al., 2008) were nostalgia, peacefulness, tenderness, and sadness. Nostalgia, the most frequently experienced emotion, is defined by both positive and negative emotionality and might best be described as a *bittersweet* emotion (Wildschut, Sedikides, Arndt, & Routledge, 2006). Nostalgia is often evoked by the retrieval of memories related to an individual's personal biography.

Sad music often evokes such emotions with opposing valences, even with highly controlled musical stimuli. For example, Hunter, Schellenberg, and Schimmack (2010) manipulated tempo and mode of short computer-generated excerpts and found that the sad versions (slow tempo, minor mode) elicited the strongest ambivalent emotions, composed of both happiness *and* sadness. The same effect was shown when the stimuli were presented for merely three musical events; that is, with a duration of 0.5 s to 1 s, depending on tempo (Weth & Kickinger, 2013). In our opinion, grasping the phenomenon of the attraction of sad music would require measures that consider mixed affective states – to discern the joy within the sadness. Furthermore, we can assume that such ambivalent emotional experiences are not cognitively accessible in full, so self-reports should be accompanied by objective methods to detect non-conscious affective shifts.

Furthermore, sad music has been shown to be particularly effective in eliciting chills (Panksepp, 1995). The sensation of chills (or goose bumps, frisson, etc.) is regarded as a strong and meaningful emotional reaction by many researchers (Grewe, Katzur, Kopiez, & Altenmuller, 2011), although its psychological significance has also been questioned (e.g., Konečni, Wanic, & Brown, 2007).

Self-reports

Using self-reports to measure emotion in response to music is popular and regarded as the best and most straightforward way to investigate emotion by some researchers (e.g., Gabrielsson, 2002). Similarly, Scherer and Coutinho (2013) expressed that self-reports offer easy access to internal processes of listeners. However, there is a number of problems researchers face when employing self-reports. One major one is that most self-reports are gathered at the end of an event or an entire piece of music – although such a measure might capture the overall impression of experiences, it is also clearly susceptible to cognitive biases. Furthermore, such measures also impede continuous inspection of affective changes and therefore make the analysis of dynamics impossible. A new solution to this problem was recently developed by Muth, Raab, and Carbon (2015), who called their measure targeting the temporal dynamics of aesthetic experience *Continuous Evaluation Procedure* (CEP). By employing a slider operated by the participant within a standardized experimental procedure, CEP aims at dynamic changes of thoughts and feelings; for example, while watching a movie or listening to music. Although this procedure enables capture of affective transitions and complex dynamics of aesthetic appreciation with a high temporal resolution (e.g., for movies at 30 Hz), CEP still struggles with potential cognitive penetration of the evaluations; that is, conscious processing of the experienced affect. Scherer and Coutinho (2013) also put forward that any kind of explicit measurement faces the problem that not all aspects of an emotion might be captured by such routines, as certain thresholds of becoming conscious must be exceeded to enable the participant to explicate his or her inner states.

Furthermore, self-reports can be subject to social biases, that is, participants might answer according to what is socially accepted and/or they believe is expected from them.

Another relevant issue in this context is the ongoing debate whether music is capable of directly eliciting genuine, "garden-variety" emotions in listeners (the "emotivist position"; see, e.g., Juslin & Västfjäll, 2008) or whether listeners only perceive the emotions portrayed by the music (the "cognitivist position"; see, e.g., Kivy, 1990). According to Kivy (1990), subjects might confuse perceived and felt emotions when reporting their responses to music. This can be particularly problematic when investigating sad music, as perceived and felt emotion often do not correspond: felt emotions tend to be richer and range from positive to negative emotions, while in terms of perceived emotion, sadness is predominant for most pieces of sad music (Kawakami, Furukawa, Katahira, & Okanoya, 2013; Weth & Kickinger, 2013). As a solution to these limitations, Scherer and Coutinho (2013) suggest to complement self-reports with objective and preferably continuous measures, such as motor expressions or neurophysiological responses. This strategy also offers the advantage of aligning these data with musical structures. In general, more objective methods are needed in order to clarify the question on the "genuineness" of sadness in response to music.

Another implication of post-performance self-reports is that participants are forced to collapse their emotional response to an entire song into a single mean value, merely revealing information about a musical piece as a whole – consequently, any kind of such measures should essentially be interpreted as a cognitive evaluation, as opposed to an immediate emotional reaction. By using time-sensitive self-reports, these restrictions are partly removed, and such measures are becoming more popular (Grewe, Nagel, Kopiez, & Altenmüller, 2007; Schubert, 1999, 2010), not least due to technical innovations and tools. For example, the tool *EmuJoy* (Nagel, Kopiez, Grewe, & Altenmüller, 2007) is a continuous self-report measure with which participants can report their current emotional state in the two-dimensional emotion system (valence and arousal), mainly used in response to musical stimuli.

Facial expressions

Among non-verbal signals, facial expressions arguably play the most prominent role in the expression of emotions in social interaction and social communication (Frith, 2009). In fact, they are efficient and differentiated to an extent that they can even be processed below the level of consciousness, when faces are presented with the backward-masking technique (e.g., Dimberg, Thunberg, & Elmehed, 2000; Whalen et al., 1998). For instance, when seeing a face for only 200 ms, our perceptual processing is already capable of capturing the main amount of emotional expression conveyed by the face (compared to seeing the face without time limit; Derntl, Seidel, Kainz, & Carbon, 2009).

Moreover, the production of facial expressions is spontaneous and rapid, and can occur without attention or awareness (Schneider & Shiffrin, 1977). Ekman (1999) proposed a small number of discrete emotions that are represented by distinct facial expressions. These "basic emotional expressions" are assumed to be expressed and recognized universally across cultures (Ekman & Friesen, 1971; see also Brown, 1991; Carlson & Hatfield, 1992). These expressions are manifested by patterns of contractions of more than 40 independent facial muscles (so-called *Facial Action Units*). Despite its popularity, the model has received substantial criticism regarding the universality of both expression (Jack, Garrod, Yu, Cladara, & Schyns, 2012; Krumhuber & Scherer, 2011) and recognition (Marsh, Elfenbein, & Ambady, 2003; Moriguchi et al., 2005) of facial expressions. For example, people are more accurate in recognizing facial expressions of members of their ethnic in-group (Elfenbein & Ambady, 2002); and in general display differing abilities in decoding facial expressions based on their cultural background (Beaupré & Hess, 2005). For an extensive critical view and a reflection on the shortcomings of these classic findings of emotional psychology, we like to refer to Russell (1994) and Nelson and Russell (2013).

In an experimental context, facial expressions are commonly measured with *surface Electromyography* (sEMG). This technique utilizes two muscles regions as indicators for either positive or negative valence (Cacioppo, Petty, Losch, & Kim, 1986): the *Zygomaticus major* (involved in smiling/represents positive valence) and the *Corrugator supercilii* (involved in frowning/represents negative valence). These two regions have been successfully linked to positively and negatively valenced stimuli in several (aesthetic) domains, such as advertisement videos (Hazlett & Hazlett, 1999), facial expressions (Dimberg et al., 2000), images (Lang, Greenwald, Bradley, & Hamm, 1993), and vocal speech (Hietanen, Surakka, & Linnankoski, 1998). In the musical domain, happy/positively valenced music commonly elicited increased zygomaticus activity (Juslin, Harmat, & Eerola, 2014; Kallinen, 2004; Lundqvist, Carlsson, Hilmersson, & Juslin, 2008; Witvliet & Vrana, 1996). Corrugator activity has been linked to sad/negatively valenced music, however, it is not as convincing as the zygomaticus–happy link. We assume this is due to the mixed quality of sad-music experiences, which defies a distinct mapping of single muscular activity onto explicit affective states. Due to their focus on valence, all sEMG studies are restricted to dimensional models of emotion (e.g., Russell, 1980) and therefore cannot shed light on specific qualities of the experienced emotion. Furthermore, one is restricted to a small number of muscle regions, if nothing else because of the burden for the participant (and therefore because of the consequences for the aesthetic experience) when multiple wires and electrodes are attached to the face.

State-of-the-art software, for example, *FaceReader* by Noldus Information Technology (Wageningen, The Netherlands), has made an automated and continuous measurement of discrete facial expressions possible. FaceReader is based on a trained Artificial Neural Network (ANN), which is operated by a statistical learning algorithm modeled after biological neural networks – an approach often used in cognitive science and machine learning. FaceReader applications include consumer research (Danner, Sidorkina, Joechl, & Duerrschmid, 2013), clinical research (Cohen, Morrison, & Callaway, 2013), personality psychology (Chentsova-Dutton & Tsai, 2010), and user experience research (Goldberg, 2014). They have utilized this software and demonstrated its possibilities. Some advantages that this technique features are: (1) it measures discrete emotions continuously and objectively (see "Measures" section), (2) participants are not confronted with cognitive load, (3) a set of six discrete emotions and their respective strengths can be analyzed simultaneously, and (4) participants are not attached to any measurement gear and hence have a naturalistic listening experience.

Aims and hypotheses

We aimed to investigate emotional responses to self-selected sad music (SSSM) by using self-reports and continuous measurement of facial expressions. Since a personal history with a certain piece of music (and consequently familiarity) might intensify the emotional experience, we also contrasted SSSM with unfamiliar sad and happy music. Besides merely investigating the quality and momentary strength of the emotions, we also considered cumulative strength (i.e., momentary strength integrated over time) and the co-activation of individual emotions.

Main hypotheses

We expected that SSSM, compared to both unfamiliar sad and happy music, would generate:

i. more sadness (self-reports and facial expression)
ii. more happiness than unfamiliar sad music, but less than unfamiliar happy music (self-reports and facial expressions)
iii. stronger overall emotions (self-reports and facial expressions)
iv. stronger mixed emotions (co-activation of multiple discrete facial expressions, see "Calculation of variables" section)

Additional hypotheses

v. SSSM (compared to both unfamiliar sad and happy music) evokes more nostalgia, reminiscence, feelings of being moved, chills, and tears (self-reports)
vi. We explored if emotionally significant musical events would evoke consistent patterns of facial expressions across listeners.

Methods

Participants

A total of 24 participants (15 female) between 23 and 30 years of age (*M* = 26.33, *SD* = 1.97) participated in the study. No professional musicians were included in the sample. Twelve of the 17 remaining participants (see section Data analysis for exclusion criteria) did not play an instrument at all or merely received general musical training as children for a maximum of three years. The remaining five participants played an instrument on a regular basis (as amateurs). On average, the participants had 4.59 years of instrumental experience. None of the participants reported hearing impairments or particular problems with auditory processing.

Material

For the SSSM, we asked every participant during the initial contact (via email) to select one song that matches the following criteria: "*Please bring along a* sad *piece of music that has a strong emotional impact on you. With 'sad music', we mean a song that other people would also identify as sad, although they do not necessarily share the same experience with the song as you do.*" The

Table 1. Technical parameters of the experimenter-selected excerpts.

Musical condition	Tempo	Bars	Time signature
Albinoni	Adagio	3–29 (da capo) → 1–21	3/4
Mozart	Allegro con spirito	1–61 (da capo) → 1–61	6/8

selections consisted of music from several genres, including classical, soul, alternative, and pop. All but one excerpt included vocals.

Furthermore, two additional pieces of music were selected for the study. We aimed to contrast two musical properties: emotional quality (sad vs. happy) and familiarity (unfamiliar vs. personally significant songs). The two unfamiliar songs were selected to portray happy or sad emotion. For sad music, we chose Albinoni's *Adagio in G minor*, which has repeatedly been used to evoke sadness in earlier research (e.g., Thompson, Schellenberg, & Husain, 2001). Based on our own assessment, we selected an excerpt from Mozart's *29th symphony in A major* (KV 201/186a) as happy music, namely the *Allegro con spirito* movement. Table 1 displays more information for the respective musical excerpts.

Procedure

Upon arrival, participants were asked to complete a paper-and-pen questionnaire that captured socio-demographic variables and years of musical experience ("For how many years did you play/have you been playing an instrument on a regular basis?"). Informed consent was given for the video recording; however, the instructions were expressed in a way to make the participants believe that the video recording was secondary to the study's purpose. The focus on facial expressions was also not mentioned. Every participant was then led into a soundproof music studio and sat in a comfortable armchair in front of a desk. The music was played back on high-quality studio monitors (Neumann KH 120A) at an appropriate level. The sequence of the three musical excerpts was counterbalanced. Videos of the participants were recorded with the built-in HD camera of an Apple MacBook (to reduce awareness to the video recording). Before leaving, participants were debriefed about the purpose of the study.

Measures

Self-reports of emotion. Following every song, participants rated the emotions they had experienced on unipolar visual analog scales of 100 mm length (coded as 1 to 100). The instructions made a clear distinction between perceived and felt emotions, and participants were explicitly briefed to report the latter, as we were interested in listeners' emotional experiences, as opposed to the music's expressiveness (Konečni, 2008). The emotions in question were happiness, sadness, surprise, anger, fear, and disgust, corresponding to the six discrete emotions included in the analysis of facial expressions (based on the concept by Ekman & Friesen, 1971). We did not expect all emotions to result in clear patterns of activation; for example, disgust is rarely reported in response to music. However, we included it in the subjective rating scales in order to compare results with the analysis of facial expressions.

Feelings of nostalgia, reminiscence, and being moved were rated likewise on visual analog scales and 5-point Likert scales were used to determine how much the listeners liked the songs (1 = *dislike* up to 5 = *like strongly*), as well as how familiar they were to them (1 = *completely*

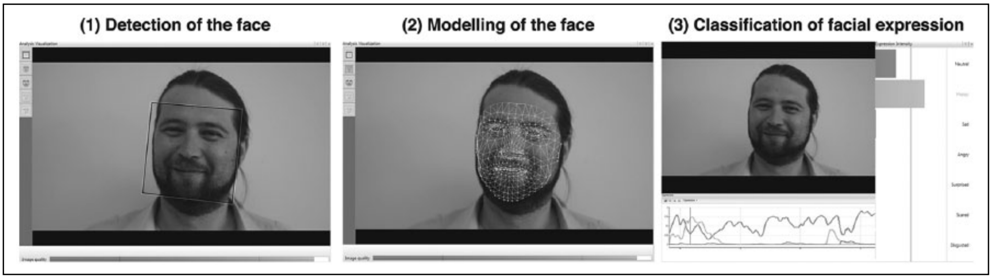

Figure 1. FaceReader's three steps of analysis.

unfamiliar up to 5 = *very familiar*). Lastly, participants indicated on dichotomous scales if they experienced any chills and tears (0 = *absent*, 1 = *present*).

Facial expressions. Facial expressions were coded using the FaceReader software by Noldus. FaceReader implements automated routines for analyzing facial expressions. The routines use an Artificial Neural Network (ANN), which is able to grasp non-linear relations between input (in our case: participants' videos) and output (here: detected emotion), and is robust concerning noisy data. According to the developer, FaceReader's ANN has been trained with more than 10,000 images and is based on the *Facial Action Coding System* (FACS, Ekman & Friesen, 1978; Hager, Ekman, & Friesen, 2002). The classification of facial expressions takes place on a frame-by-frame basis; that is, 30 frames per second in our case, assigning a decimal value between 0 and 1 to each of the following six discrete emotions: happiness, sadness, surprise, anger, fear, and disgust. Compared to ratings by human FACS coders, the software claims an accuracy of 85 to 96% for the distinct emotions (Noldus Information Technology, 2008; see also den Uyl & van Kuilenburg, 2005; van Kuilenburg, Wiering, & den Uyl, 2005). A validation with facial EMG has also shown good results (D'Arcey, 2013).

The software operates in three processing steps: (1) detection of a face by means of the Viola–Jones algorithm (Viola & Jones, 2001), (2) modeling of the face with 500 facial key points based on the Active Appearance method (Cootes & Taylor, 2000), and (3) classification of the facial expression based on the trained ANN (see Figure 1). We used the resulting time-series data to calculate new variables for testing our hypotheses. Although FaceReader supports a live analysis, we decided to record the videos during the testing situation, and to analyze them post-hoc to have control over optimization as well as analysis parameters offered by the software. This kind of analysis allowed us to run a calibration on a per-subject basis, where the whole video is considered once by the software, determining an individual's "emotional range". This way, people showing very salient expressions in general (e.g., if they are accustomed to convey emotions via facial expressions, which might be true for actors or experienced stage musicians) are comparable to people showing subtle expressions in general.

Data analysis

The data of seven participants were excluded from further analyses for a number of reasons: (1) incorrect facial recognition by the software; (2) signal loss for facial recognition due to individual movements, for example, burying one's face in one's hand, turning head sideways, etc., (3) familiarity with musical stimuli (Mozart and/or Albinoni).

In light of our focus on a naturalistic music listening experience, we regarded the rather high exclusion rate (n = 4) based on reason (2) as a confirmation of the ecologically valid situation.

Both the Mozart and Albinoni excerpts were edited to a length of 2:50 minutes. For the statistical analysis of the SSSM, we merely considered the first 2:50 minutes in order to keep the length constant across conditions. However, during testing the participants listened to their song in full length. A fade in/out of 0.5 seconds was applied to every excerpt. The loudness of each excerpt was equaled in accordance with the EBU R128 norm, while the speakers' output volume was held constant for all participants.

Calculation of variables. We used FaceReader's time-series data, which assigns a decimal value between 0 and 1 (0 = *emotion absent*, 1 = *emotion fully present*) to every individual emotion on a frame-by-frame basis (30 fps), as a starting point to calculate customized variables.

To assess the intensity of the individual facial expressions for Hypotheses 1 and 2, we calculated a variable named *emotional peaks* to suit a definition of emotions as brief, intensive and rapidly changing events (Juslin, 2011). First, average means and standard deviations for every emotion were extracted for every single video. As we were mainly interested in emotional peak events of every emotion, we calculated means of all data points which were above the second standard deviation of the raw mean values, resulting in a global indicator for those passages that had induced very strong emotions (compared to the rest of the individual listening experience). In order to account for the time-sensitivity of the measure, we multiplied the resulting mean value with the *amount of time* that the listeners displayed facial expressions exceeding the second standard deviation regarding their intensity. In our case, time was measured in *frames* (i.e., 1/30th of a second), since every data point represented one frame. The resulting variable thus incorporates information on the amplitude of the respective strong emotion (i.e., the peaks) as well as its extent over the course of the entire song in a single, global indicator.

All values were normalized prior to further analyses. The z-scores were calculated across the three musical conditions (SSSM, Albinoni, Mozart), across all participants, but separately for all six basic emotions. Thus, the z-scores for the individual emotions (Figure 4) cannot be compared across emotions in terms of their intensity. They rather mirror the differences between the three musical conditions for every emotion, since our hypotheses focus on the emotional differences of the individual musical excerpts.

As formulated in Hypothesis 4, we were also interested in the co-activation of multiple discrete emotions and hence defined *mixed emotions* as episodes in which at least two emotions were active simultaneously. To test for such reactions, we calculated an overall value for mixed emotions based on FaceReader's time-series data. Facial expressions are coded as a value between 0 and 1. We applied the following procedure: (a) in order to exclude baseline noise, we determined a cut-off value of .05 (that is, minor activations under 5% intensity), only considering values higher than .05 in the following; (b) then, all data points were identified in which any possible pairwise combination of the six emotions (yielding 15 possible combinations) were co-activated (i.e., when the value for both emotions was over .05 at the same point in time). In a next step (c), a mean value for every co-active emotion pair (i.e., the mean of both values) was calculated (momentary strength), and (d) subsequently multiplied with the *amount of time* (in *frames*, as mentioned above) the co-activation lasted (cumulative strength). As a last step (e), the resulting values for every emotion pair were added, resulting in an overall mixed emotion value, that is, a single value for each person and musical piece.

Table 2. Repeated-measure ANOVAs for emotional self-reports and facial expressions.

	$F(2, 32)$	P	η_p^2
Happiness			
Musical Excerpt (SR)	34.01	< .001	.680
Sadness			
Musical Excerpt (SR)	38.48	< .001	.706
Musical Excerpt (FE)	3.19	.055	.166
Nostalgia			
Musical Excerpt (SR)	42.49	< .001	.726
Reminiscence			
Musical Excerpt (SR)	34.41	< .001	.683
Being moved			
Musical Excerpt (SR)	15.49	< .001	.492

Note. Marginal significance values are presented as well ($p < .07$); SR = Self-report, FE = Facial Expression.

Results

Familiarity and liking

As a prerequisite for investigating our hypotheses, we compared whether the participants were more familiar with the SSSM excerpts than with the experimenter-selected sad and happy music. To do so, we used the ratings of the 5-point Likert scale (1 = *completely unfamiliar* to 5 = *very familiar*). Participants with a rating of 3 or higher for either the Albinoni or Mozart excerpt were excluded from the dataset ($n = 2$). As expected, among the remaining participants, SSSM was more familiar ($M = 4.82$, $SD = 0.39$) than both the Albinoni ($M = 1.12$, $SD = 0.33$) and Mozart ($M = 1.06$, $SD = 0.24$), as revealed by a repeated-measures ANOVA with musical excerpt (SSSM, Albinoni, Mozart) as the within-subject variable, $F(2,32) = 658.45$, $p < .001$, $\eta_p^2 = .976$. In fact, 15 out of 17 participants claimed having never heard Albinoni's Adagio before, and 16 out of the same 17 participants had no previous experience with the Mozart excerpt.

Liking for the individual musical pieces was captured on 5-point Likert scales. We compared liking for the individual musical excerpts and found a significant influence, $F(2,32) = 11.01$, $p < .001$, $\eta_p^2 = .41$. Post-hoc comparisons revealed that both SSSM ($M = 4.12$, $SD = .78$) and the Mozart excerpt ($M = 3.76$, $SD = 0.83$) were liked significantly more than the Albinoni piece ($M = 2.82$, $SD = 1.02$), $p = .003$ and $p = .015$, respectively.

Hypothesis 1: SSSM evokes more sadness than unfamiliar happy and sad music

In order to analyze the emotional self-reports for *sadness*, we conducted a repeated-measure ANOVA with *musical excerpt* as the within-subject variable (SSSM, Albinoni, Mozart). As can be seen in Table 2, sadness produced a significant effect. To evaluate the differences between the individual conditions, we conducted planned comparisons (dependent *t*-tests). As expected, SSSM received the highest ratings, however, differing significantly from the Mozart excerpt only (see Figure 2).

For facial expressions, we utilized the *emotional peak* value (see "Calculation of variables" section) for sadness and conducted a repeated-measure ANOVA with musical excerpt as the within-subject variable. The individual excerpts had a marginally significant effect on sadness

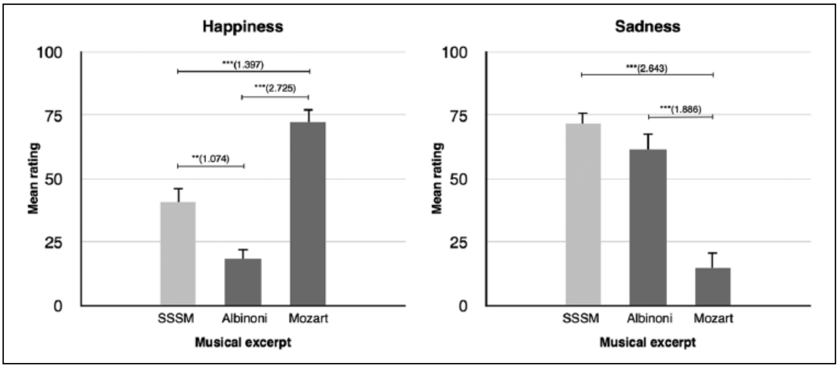

Figure 2. Means of listeners' emotion scores based on self-reports of basic emotions as a function of musical excerpt. Error bars indicate +1 standard error of the mean (*p < .05, **p < .01, ***p < .001). Values in parentheses show effect sizes (Cohen's d).

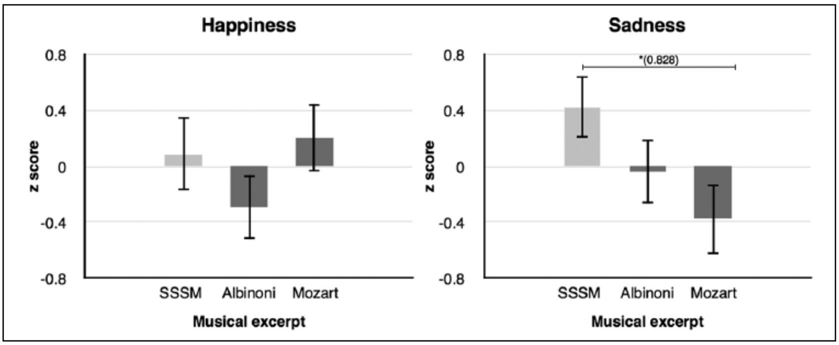

Figure 3. Means of listeners' emotion scores based on facial expressions (z-scores) as a function of musical condition. Error bars indicate +1 standard error of the mean (*p < .05, **p < .01, ***p < .001). Values in parentheses are effect sizes (Cohen's d).

(see Table 2). Planned comparisons (dependent t-tests) revealed that SSSM evoked the strongest sad facial expressions, followed by unfamiliar sad and happy music (see Figure 3). The difference between SSSM and Mozart was significant.

For good measure, we conducted similar ANOVAs and planned comparisons for the remaining discrete emotions *surprise*, *fear*, *anger*, and *disgust*. However, these results will not be further discussed since they do not directly relate to our hypotheses. For significant results of the ANOVAs, as well as means and standard errors of the individual scales, see Table S3, Figure S7 and Figure S8 in the Supplemental Material Online section.

Hypothesis 2: SSSM evokes more happiness than unfamiliar sad music, but less than unfamiliar happy music

To test our hypothesis, we conducted a repeated-measure ANOVA with *musical excerpt* as the within-subject variable to test if self-reported happiness differed in the individual musical conditions. This was the case as shown in Table 2. Planned comparisons (dependent t-tests) are

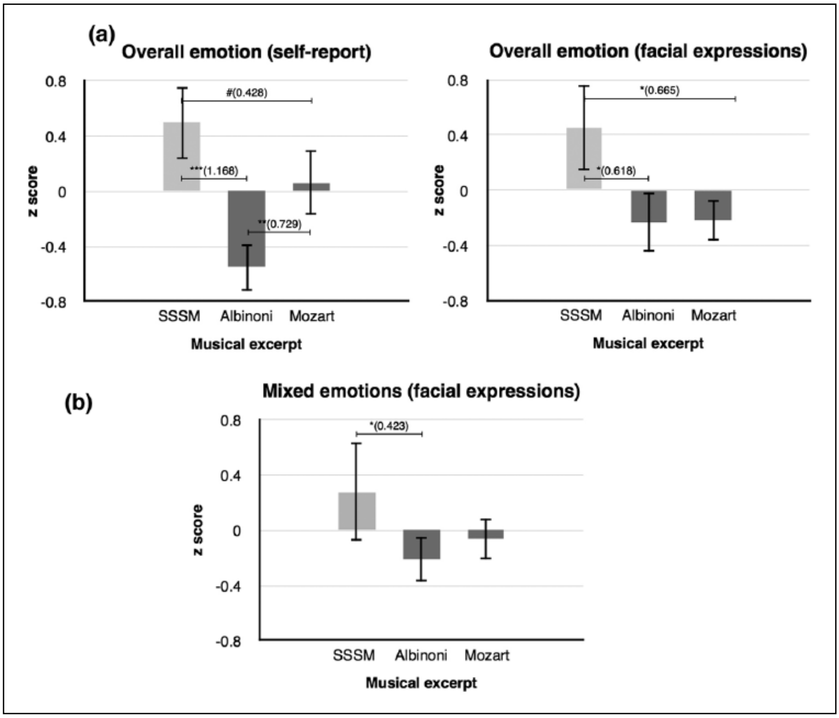

Figure 4. Means and standard errors of (a) overall and (b) mixed emotions (*p < .05, **p < .01, ***p < .001). Values in parentheses are effect sizes (Cohen's *d*).

presented in Figure 2, and show that Mozart produced higher ratings of *happiness* than SSSM and Albinoni. In line with our hypothesis, ratings for SSSM were also higher than for Albinoni.

On the basis of the emotional peak value (see "Calculation of variables" section), we tested happy facial expressions in an equivalent fashion. Contrary to our predictions, the ANOVA did not yield a significant result. Yet, we observed the expected trend (see Figure 3); namely, that Mozart evoked the strongest happy expressions, followed by SSSM, and lastly, the Albinoni excerpt.

Hypothesis 3: SSSM evokes stronger overall emotions than unfamiliar happy and sad music

Regarding the self-reported emotions, a factor representing the *overall emotional response* was established by simply adding the values of the six discrete emotion scales happiness, sadness, surprise, fear, anger, and disgust for a given song. A repeated-measure ANOVA with musical excerpt as a within-subject variable demonstrated a significant influence of musical excerpt on overall emotions, $F(2,32) = 16.02$, $p < .001$, $\eta_p^2 = .50$. Planned comparisons (dependent *t*-tests) revealed that SSSM resulted in the stronger overall emotions than the Mozart and the Albinoni excerpt (see Figure 4a).

Accordingly, we added the *emotional peak* values (see "Calculation of variables" section) of the six discrete emotions to assess the *overall emotional response* based on facial expressions. We conducted a repeated-measure ANOVA with musical excerpt as the within-subject variable

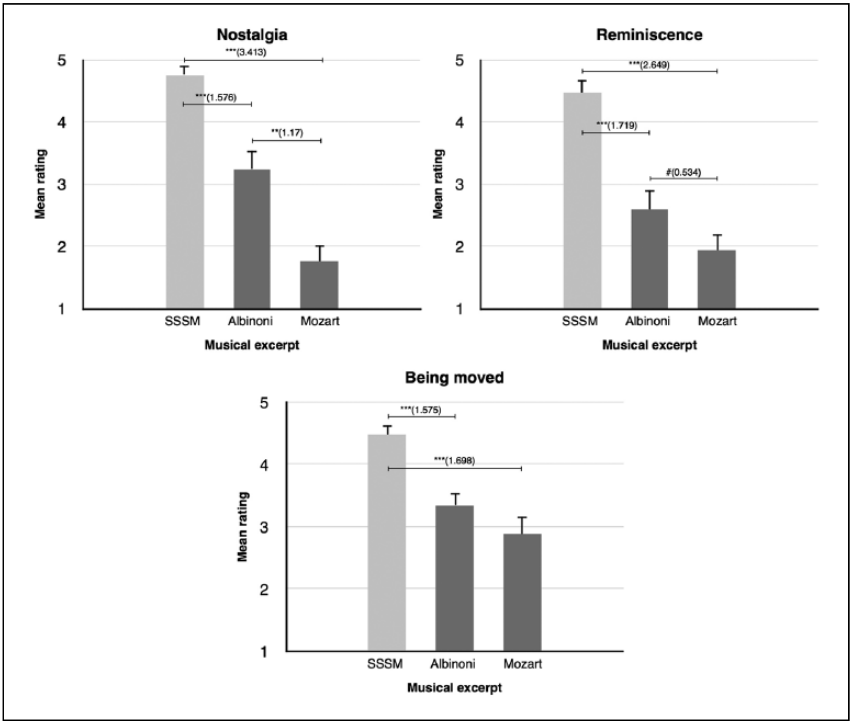

Figure 5. Means of listeners' emotion scores based on self-reports of additional emotions as a function of musical excerpt. Error bars indicate +1 standard error of the mean (**p* < .05, ***p* < .01, ****p* < .001). Values in parentheses show effect sizes (Cohen's *d*). Marginal significance (*p* < .08) is symbolized by #.

(SSSM, Albinoni, Mozart), results suggesting a significant influence, $F(2,32) = 4.57$, $p = .018$, $\eta_p^2 = .22$. As hypothesized, SSSM elicited stronger overall facial expressions than Mozart and Albinoni (see Figure 4a).

Hypothesis 4: SSSM evokes more mixed emotions than unfamiliar happy and sad music

We found a significant difference for mixed emotions in response to the individual musical excerpts as revealed by a repeated-measure ANOVA, $F(2,32) = 3.36$, $p = .047$, $\eta_p^2 = .173$. As expected, listeners displayed the strongest mixed emotions in response to SSSM, compared to the Mozart and the Albinoni excerpt (see Figure 4b).

Hypothesis 5: SSSM evokes more nostalgia, reminiscence, feelings of being moved, chills and tears, than unfamiliar happy and sad music

As revealed by repeated-measure ANOVAs and presented in Table 2, the additional emotional scales *nostalgia*, *reminiscence*, and *being moved*, were significantly influenced by the musical condition. Planned comparisons (dependent *t*-tests) showed that the experience of nostalgia, reminiscence, and being moved was significantly stronger for SSSM than for the Albinoni and Mozart excerpts (see Figure 5).

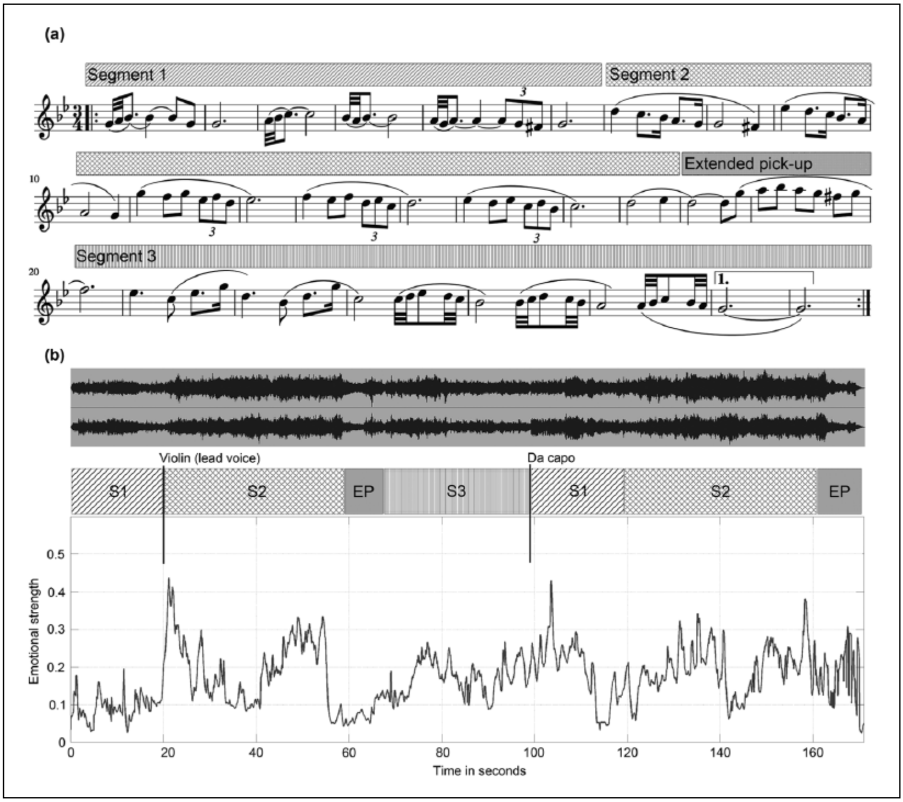

Figure 6. (a) Notation of the lead violin of the Albinoni excerpt with the identified segments. (b) Median curve of sad facial expressions for Albinoni's Adagio (*N* = 17). Segments are marked in boxes (S1–S3 = Segment 1–3; EP = Extended pick-up).

Participants reported their experience of *chills* and *tears* on dichotomous scales (yes/no). Eight of 17 participants (47%) experienced one or more sensations of chills in response to SSSM. Albinoni evoked chills in three cases (18%), while Mozart evoked no chills at all. A Friedman ANOVA confirmed that the number of reported chills differed significantly across the three musical excerpts, $\chi^2(2) = 9.80$, $p = .007$, *Kendall's W* = .29.

Similarly, SSSM elicited tears in four of 17 participants (24%), as compared with one for Albinoni (6%) and 0 in response to Mozart, this difference also being significant, $\chi^2(2) = 6.5$, $p = .039$, *Kendall's W* = .19.

Hypothesis 6: Consistency of facial expressions in response to musical events

Another issue was the consistency of listeners' facial expressions to specific musical events. Two musicologists were asked to conduct a segmentation task on our excerpt of Albinoni's Adagio; that is, to define the beginnings and endings of new musical parts. Figure 6a shows the segments that were identified by both musicologists. Segment 1 represents the introduction to the piece, consisting of the basso continuo and low violins that play short and stagnating melodic phrases. Segment 2 introduces the lead melody, in which the violins perform a constantly descending melodic line in several variations. Segment 3 combines attributes of the preceding segments, the violins both performing descending (bars 23–24), as well as stagnating (bars 25–27) melodic lines, which function as a transition to the repetition of segment 1 (da capo). Regarding the beginning of segment 3, our musicologists were discordant about its beginning. While one expert marked its beginning at bar 21, the other classified the same bar as an independent event, since it acts as an "extended pick-up", and lies outside of the mathematical counting of the piece.

We compared the results of the musical analysis with the median graph for sad facial expressions for all participants in response to the Albinoni excerpt. Figure 6b shows that the two highest peaks of sad facial expressions corresponded to two musical events defined in the musical analysis. The first (and highest) peak at 0:21 minutes represents the beginning of segment 2, which is defined by the introduction of the lead voice (the violin in this case). The second highest peak, occurring at 1:41 minutes, constitutes the da capo mark; that is, segment 1 being repeated for the first time.

Discussion

The present study explored emotional responses to familiar, self-selected sad music (SSSM), and compared these with responses to unfamiliar sad and happy music. The listeners' subjective experience of emotions was captured via self-reports. To measure facial expressions, we introduced a software-based approach to automatically assess facial expressions in a continuous fashion. The results in general suggest that music evokes measurable and distinct patterns of facial activity, depending on the respective musical excerpt, which are too complex to be understood when only two muscular regions are considered (as with facial EMG).

We found that SSSM evoked the highest amounts of mixed facial expressions (Hypothesis 4), compared to unfamiliar sad (significant) and happy music (not significant). This activity indicates that music in general, and especially SSSM, evokes a complex blend of emotions – emotions that, with other methodological approaches, would not be observable; or under other circumstances, would be perceived as discrete or even contradicting.

Regarding self-reports, SSSM scored significantly higher for happiness compared to unfamiliar sad music (Hypothesis 2); yet, four participants reported tears listening to SSSM. We also found that listeners experienced significantly more chills while listening to SSSM compared to unfamiliar sad as well as happy music (Hypothesis 5). In general, chills are regarded as intensely pleasurable emotional peak experiences by listeners (Grewe, Kopiez, & Altenmüller, 2009), and are accompanied by simultaneous sympathetic activity (Salimpoor, Benovoy, Longo, Cooperstock, & Zatorre, 2009), as well as by activation in the brain's reward circuitry and a release of dopamine (Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011). The above-mentioned results emphasize the role of positively valenced emotions with SSSM (Van den Tol & Edwards, 2015), which are regarded as a main motive for listening to sad music (Taruffi & Koelsch, 2014). Unfortunately, the analysis of happy facial expressions failed to reach significance; however, it showed the same trend as the corresponding self-reports.

Earlier studies (e.g., Sloboda, 1991) have identified specific musical passages that are particularly potent in evoking strong physiological responses in listeners. Among these musical passages, the entry of the lead voice has been found to induce synchronized responses in the subjective and physiological domain (Grewe et al., 2007). Our results extend these findings to the area of facial expressions, as we found a major increase in sad facial expressions across participants when the lead violin is introduced for the first time in Albinoni's Adagio (Hypothesis 6). Juslin and Laukka (2003, p. 803) described the violin as a "super-expressive" version of the human voice, as "it sounds a lot like the human voice while going far beyond what the human voice can do". In fact, this similarity and the violin's capability to exaggerate acoustic characteristics that listeners relate with sad emotion in speech might be a reason for its ability to evoke sad facial expressions, "mirroring" the sadness of a hypothetical individual. In contrast to its first occurrence, we merely found a local peak in sad facial expressions for the second time the lead voice enters, suggesting a habituation effect. Our findings extend previous results by Grewe et al. (2007), who found a similar effect for skin conductance response. In contrast to this habituation effect, we observed a global peak of sad facial expressions for the repetition of the introduction, while the first introduction only resulted in a local peak. At this point, it is difficult to interpret these inverted results – however we assume that these differences might be based on the different saliences of the respective musical events.

Consistent with Hypothesis 1, SSSM evoked significantly more sad facial expressions than the happy excerpt, which extends the somewhat thin body of evidence that sad music actually evokes sad facial expressions. Other studies found no difference in corrugator activity between (unfamiliar) sad and happy music (Lundqvist et al., 2008), perhaps due to the restriction to merely one muscle region. Similarly, in this study, sad facial expressions did not differ significantly between unfamiliar sad and happy music, although the former evoked stronger expressions. SSSM also induced the strongest subjective self-reported feelings of *sadness* (Hypothesis 1), followed by unfamiliar sad and happy music; both sad excerpts being significantly higher than the happy excerpt. Both facial analysis and self-reports revealed no significant difference in sadness between SSSM and unfamiliar sad music. This is somewhat surprising, however this result points to the complexity of emotions evoked by SSSM. In other words, sadness is merely one emotional facet of SSSM, which is evident in (1) the involvement of mixed emotions, and (2) the crucial role of positively valenced emotions with SSSM.

In terms of overall facial expressions, SSSM elicited significantly stronger activity than the other excerpts, which was confirmed by participants' self-reports (Hypothesis 3). *Familiarity* is a factor that is likely to contribute to the strong overall responses to the SSSM excerpts. It intensifies the emotional experience (Pereira et al., 2011) as well as the degree of *liking* for a piece of music (North & Hargreaves, 1995; Peretz, Gagnon, & Bouchard, 1998). Indeed, in the present study listeners liked the familiar SSSM excerpts most, but showed a similar degree of liking for the unfamiliar happy music. The tendency of liking unfamiliar happy over sad music has also been previously reported (Hunter et al., 2010; Thompson et al., 2001).

The importance of familiarity, combined with the observed blend of joyful and sad emotions, points us to *nostalgia*, an affective state classified as a complex emotion (Johnson-Laird & Oakley, 1989), which differs from a basic emotion due to its reflection of high-level cognitive appraisals linked to an individual's biography. Nostalgia is mainly triggered by negative affect, social interactions, and sensory input such as music (Barrett et al., 2010). Accordingly, we found a relationship with sad musical excerpts, especially SSSM, which evoked significantly stronger ratings of nostalgia (Hypothesis 5) than the other two excerpts. Unfamiliar sad music also elicited significantly stronger nostalgia than happy music. Nostalgia has been repeatedly found to play a role with SSSM. In fact, several studies with nearly 1,500 participants combined (Taruffi & Koelsch, 2014; Weth & Parncutt, 2014) showed that nostalgia is the most common reaction to SSSM. Our results confirm the strong involvement of nostalgia in SSSM, and furthermore suggest that nostalgia is facilitated by music with sad properties in general. Sedikides, Wildschut, and Baden (2004) noted that people often engage in nostalgia in order to repair negative moods, to boost self-esteem, and to feel social connectedness. These tendencies are also important motives of listening to sad music (Van den Tol & Edwards, 2015). Similarly to the results for nostalgia, we found higher ratings for reminiscence with SSSM and unfamiliar sad music compared to unfamiliar happy music. Listeners often reminisce and ruminate about past events when listening to SSSM (Garrido, 2009). The high ratings for nostalgia and reminiscence with SSSM propose a strong involvement of episodic memories. Future research could investigate if episodic memory retrieval is predominant in evoking emotions when listening to SSSM. Thereby, the BRECVEMA framework (which includes the eight mechanisms Brain Stem Reflex, Rhythmic Entrainment, Evaluative Conditioning, Contagion, Visual Imagery, Episodic Memory, Musical Expctancy, and Aesthetic Judgment) of musical emotions (Juslin, 2013; Juslin & Västfjäll, 2008) could function as a theoretical basis, according to which, *episodic memory* is one of seven proposed mechanisms that trigger emotions in response to music.

Limitations and outlook

A fact worth considering is that 94% of the SSSM excerpts contained lyrics. Especially in the case of sad music, lyrics may be crucial as they enhance the experience of sadness (Ali & Peynircioğlu, 2006; Brattico et al., 2011). However, Ali and Peynircioğlu (2006) also noted that melodic content in general has a bigger impact on listeners' emotions than lyrics. In this study, we cannot rule out the fact that lyrics might have contributed to the general effects of SSSM. However, it was important for us to pose no restrictions on the participants regarding their self-selected music. It might even be the case that our participants favored music with lyrics *because* the emotional impact is stronger; in other words, because lyrics contribute to the quality of the mixed emotion we're aiming at. In future studies, it would be interesting to compare if instrumental SSSM has equivalent effects on facial expressions as SSSM with lyrics. However, following Juslin and Laukka (2003), violins might be perceived similarly to a human voice, making a clear distinction *with vs. without vocals* questionable, at least regarding the emotional impact.

In general, facial expressions are automatic and spontaneous, although under certain circumstances individuals voluntarily utilize them to express a desired emotion; for example, to act in a socially desirable way. The display of facial expressions increases when other individuals are present, and particularly when the communicator is aware that his behavior is being observed (Frith, 2009). We cannot exclude the possibility for such effects in the current study, however, there are several reasons to assume the predominance of involuntary facial expressions: (1) participants were not informed about the relevance of facial expressions and (2) there was no social context (in which facial expressions would be more prominent and hence, susceptible for biases). We must consider, however, that we might fail to capture all occurring emotions, as there are situations in which emotions are experienced subjectively, however no facial expressions are evident (Ekman & Friesen, 1971). Nevertheless, we were able to record observable facial expressions although participants listened to music alone and were not aware of their facial expressions being recorded.

Regarding the software's performance, we occasionally observed unexplainable impairments in terms of technical reliability. For example, in a few instances, we encountered sudden signal losses in facial recognition for no apparent reason. At this point, it is next to impossible to deal with these flaws, as the software's algorithm is not transparent for the researcher. This may be based on the fact that the software was created for non-academic, commercial use, such as practical market research. Despite these limitations in transparency, we believe that FaceReader offers great possibilities for future research, if we are able to better understand its exact operations and therefore evaluate its results with more certainty.

This study was an initial step towards the use of a new method to measure facial expressions in music. For purposes of validation, we focused on straightforward hypotheses and included self-reports to evaluate and validate the findings. In our opinion, future research should emphasize two key points to maximize the benefits of FaceReader: (1) its time-sensitive aspects, for example, when considering specific musical events and their relationship with facial expressions, and (2) the impact of facial expressions on social behavior. The latter point bears interesting possibilities for future research on the communication of emotions in music performance, dance, and when listening to music in a social context. These aspects have been neglected in research so far, yet, they may have major implications regarding the question of why we listen to music in general. In fact, social bonding and social cohesion are regarded by some as the most likely reason why music has survived throughout human evolution (Huron, 2001). To our knowledge, this study is the first in the field of music and emotion to measure facial activity based on the FACS system (Ekman & Friesen, 1977), and consequently the basic emotion model (see, e.g., Ekman, 1992; Izard, 1992). This possibility provides major advantages for future research, for example, to investigate mixed emotions, a concept that is becoming increasingly popular (Hunter et al., 2010; Larsen & McGraw, 2011). Mixed emotions, and especially emotions composed of positive and negative qualities, are crucial for research on sad music (see above, the mixed emotional quality of nostalgia). A better understanding might help us to gain insights into the paradox of why we enjoy listening to sad music. Nostalgia – including a reflection of ups and downs, encounters and farewells we have experienced so far – might help us to better understand our actual self and our current social relations.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Supplemental material

Supplemental online material is available from http://msx.sagepub.com/content/by/supplemental-data

References

Ali, S. O., & Peynircioğlu, Z. F. (2006). Songs and emotions: Are lyrics and melodies equal partners? *Psychology of Music*, *34*(4), 511–534.

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C., & Janata, P. (2010). Music-evoked nostalgia: Affect, memory, and personality. *Emotion*, *10*(3), 390–403.

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, *36*(3), 355–370.

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S. K., & Tervaniemi, M. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology*, *2*, 308.

Brown, D. E. (1991). *Human universals*. New York, NY: McGraw-Hill.

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*(2), 260–268.

Carlson, J. G., & Hatfield, E. (1992). *Psychology of emotion*. San Diego, CA: Harcourt Brace Jovanovich.

Chentsova-Dutton, Y. E., & Tsai, J. L. (2010). Self-focused attention and emotional reactivity: The role of culture. *Journal of Personality and Social Psychology*, *98*(3), 507–519.

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, *11*(9), 393–399.

Cohen, A. S., Morrison, S. C., & Callaway, D. A. (2013). Computerized facial analysis for understanding constricted/blunted affect: Initial feasibility, reliability, and validity data. *Schizophrenia Research*, *148*(1–3), 111–116.

Cootes, T. F., & Taylor, C. J. (2000). *Statistical models of appearance for computer vision* [Technical report]. Manchester, UK: University of Manchester.

Danner, L., Sidorkina, L., Joechl, M., & Duerrschmid, K. (2013). Make a face! Implicit and explicit measurement of facial expressions elicited by orange juices using face reading technology. *Food Quality and Preference*, *32*, 167–172.

D'Arcey, J. T. (2013). *Assessing the validity of FaceReader using facial EMG* (Unpublished master's thesis). California State University, Chico.

den Uyl, M. J., & van Kuilenburg, H. (2005). The FaceReader: Online facial expression recognition. In L. P. J. J. Noldus, F. Grieco, L.W. S. Loijens, & P.H. Zimmerman (Eds.), *Proceedings of Measuring Behavior, 5th International Conference on Methods and Techniques in Behavioral Research* (pp. 589–590). Wageningen, The Netherlands: Noldus Information Technology.

Derntl, B., Seidel, E.-M., Kainz, E., & Carbon, C.-C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception*, *38*(12), 1849–1862.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3), 169–200.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & T. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). New York, NY: Wiley.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological Bulletin*, *128*(2), 243–249.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, *5*(1 Suppl., Fall 2001–2002), 123–147.

Garrido, S. (2009). Rumination and sad music: A review of the literature and a future direction. In C. Stevens, E. Schubert, B. Kruithof, K. Buckley, & S. Fazio (Eds.), *Proceedings of the 2nd International Conference on Music Communication Science (ICoMCS2)*. Sydney, Australia: HCSNet, University of Western Sydney.

Goldberg, J. H. (2014). Measuring software screen complexity: Relating eye tracking, emotional valence, and subjective ratings. *International Journal of Human-Computer Interaction*, *30*(7), 518–532.

Grewe, O., Katzur, B., Kopiez, R., & Altenmüller, E. (2011). Chills in different sensory domains: Frisson elicited by acoustical, visual, tactile and gustatory stimuli. *Psychology of Music*, *39*(2), 220–239.

Grewe, O., Kopiez, R., & Altenmüller, E. (2009). Chills as an indicator of individual emotional peaks. *Annals of the New York Academy of Sciences*, *1169*(1), 351–354.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion*, *7*(4), 774–788.

Hager, J. C., Ekman, P., & Friesen, W. V. (2002). *Facial action coding system* (2nd ed.). Salt Lake City, UT: A Human Face.

Hazlett, R. L., & Hazlett, S. Y. (1999). Emotional response to television commercials: Facial EMG vs. self-report. *Journal of Advertising Research*, *39*, 7–24.

Hietanen, J. K., Surakka, V., & Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology*, *35*(5), 530–536.

Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(1), 47–56.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, *930*(1), 43–61.

Huron, D. (2011). Why is sad music pleasurable? A possible role for prolactin. *Musicae Scientiae*, *15*(2), 146–158.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, *99*(3), 561–565.

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7241–7244.

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and Emotion*, *3*(2), 81–123.

Juslin, P. N. (2011). Music and emotion: Seven questions, seven answers. In I. Deliège & J. W. Davidson (Eds.), *Music and the mind: Essays in honour of John Sloboda* (pp. 113–138). New York, NY: Oxford University Press.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Toward a unified theory of musical emotions. *Physics of Life Reviews*, 10, 235–266.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 1–63.

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, *42*(4), 599–623.

Kallinen, K. (2004). Emotion related psychophysiological responses to listening music with eyes open versus eyes-closed: electrodermal (EDA), electrocardiac (ECG), and electromyographic (EMG) measures. In S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the 8th International Conference on Music Perception and Cognition* (pp. 299–301). Rundle Mall, Australia: Causal Productions.

Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, *4*, 311 (article 311).

Kivy, P. (1990). *Music alone: Philosophical reflections on the purely musical experience*. Ithaca, NY: Cornell University Press.

Konečni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 115–129.

Konečni, V. J., Wanic, R. A., & Brown, A. (2007). Emotional and aesthetic antecedents and consequences of music-induced thrills. *The American Journal of Psychology*, *120*(4), 619–643.

Krumhuber, E. G., & Scherer, K. R. (2011). Affect bursts: Dynamic patterns of facial expression. *Emotion*, *11*(4), 825–841.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273.

Larsen, J. T., & McGraw, A. P. (2011). Further evidence for mixed emotions. *Journal of Personality and Social Psychology*, *100*(6), 1095–1110.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*(5), 834–855.

Lundqvist, L. O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2008). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music*, *37*(1), 61–90.

Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal "accents": Cultural differences in facial expressions of emotion. *Psychological Science*, *14*(4), 373–376.

Moriguchi, Y., Ohnishi, T., Kawachi, T., Mori, T., Hirakata, M., Yamada, M., & Komaki, G. (2005). Specific brain activation in Japanese and Caucasian people to fearful faces. *Neuroreport*, *16*(2), 133–136.

Muth, C., Raab, M. H., & Carbon, C.-C. (2015). The stream of experience when watching artistic movies: Dynamic aesthetic effects revealed by the continuous evaluation procedure (CEP). *Frontiers in Psychology*, *6*, 365.

Nagel, F., Kopiez, R., Grewe, O., & Altenmüller, E. (2007). EMuJoy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods*, *39*(2), 283–290.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review*, *5*(1), 8–15.

Noldus Information Technology. (2008). *FaceReader reference manual (Version 5)*. Wageningen, The Netherlands: Author.

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, *14*, 77–93.

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception*, *13*(2), 171–207.

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: Familiarity matters. *PLoS ONE*, *6*(11), e27241–9.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*(2), 111–141.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141.

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, *14*(2), 257–262.

Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS ONE*, *4*(10), e7487.

Scherer, K. R., & Coutinho, E. (2013). How music creates emotion: A multifactorial process approach. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The emotional power of music: Multidisciplinary perspectives on musical expression, arousal, and social control* (pp. 121–146). Oxford, UK: Oxford University Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Schubert, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, *51*(3), 154–165.

Schubert, E. (2010). Continuous self-report methods. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 223–253). Oxford, UK: Oxford University Press.

Sedikides, C., Wildschut, T., & Baden, D. (2004). Nostalgia: Conceptual issues and existential functions. In J. Greenberg, S. Koole, & T. Pyszczynski (Eds.), *Handbook of experimental existential psychology* (pp. 200–214). New York, NY: Guilford.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, *19*(2), 110–120.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS ONE*, *9*(10), e110490.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*(3), 248–251.

Van den Tol, A. J. M., & Edwards, J. (2013). Exploring a rationale for choosing to listen to sad music when feeling sad. *Psychology of Music*, *41*(4), 440–465.

Van den Tol, A. J. M., & Edwards, J. (2015). Listening to sad music in adverse situations: How music selection strategies relate to self-regulatory goals, listening effects, and mood enhancement. *Psychology of Music*, *43*, 473–494.

Van Kuilenburg, H., Wiering, M., & den Uyl, M. (2005). A model based method for automatic facial expression recognition. In J. Gama, R. Camacho, P. Brazdil, A. M. Jorge, & L. Torgo (Eds.), *Proceedings of the 16th European Conference on Machine Learning* (pp. 194–205). Berlin, Germany: Springer Verlag.

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, 1, 511–518.

Vuoskoski, J. K., & Eerola, T. (2012). Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(3), 204–213.

Weth, K., & Kickinger, M. (2013, June). *Ambivalent emotions in music: We like sad music when it makes us happy*. Paper presented at the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland. Retrieved from https://jyx.jyu.fi/dspace/handle/123456789/41597

Weth, K., & Parncutt, R. (2014, September). *Functions of sad music in daily life: An explorative study on the influence of exposure, coping strategies and personality*. Paper presented at the International Conference of Students of Systematic Musicology, London, UK. Retrieved from http://ojs.gold.ac.uk/index.php/sysmus14/article/view/261/276

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience*, *18*(1), 411–418.

Wildschut, T., Sedikides, C., Arndt, J., & Routledge, C. (2006). Nostalgia: Content, triggers, functions. *Journal of Personality and Social Psychology*, *91*(5), 975–993.

Witvliet, C. V., & Vrana, S. R. (1996). The emotional impact of instrumental music on affect ratings, facial EMG, autonomic measures, and the startle reflex: Effects of valence and arousal. *Psychophysiology Supplement, 91*.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521.