Empathy and Emotional Contagion as a Link Between Recognized and Felt Emotions in Music Listening

Hauke Egermann
Technische Universität Berlin, Germany

Stephen McAdams
McGill University, Montreal, Canada

Previous studies have shown that there is a difference between recognized and induced emotion in music listening. In this study, empathy is tested as a possible moderator between recognition and induction that is, on its own, moderated via music preference evaluations and other individual and situational features. Preference was also tested to determine whether it had an effect on measures of emotion independently from emotional expression. A web-based experiment gathered from 3,164 music listeners emotion, empathy, and preference ratings in a between-subjects design embedded in a music-personality test. Stimuli were a sample of 23 musical excerpts (each 30 seconds long, five randomly assigned to each participant) from various musical styles chosen to represent different emotions and preferences. Listeners in the recognition rating condition rated measures of valence and arousal significantly differently than listeners in the felt rating condition. Empathy ratings were shown to modulate this relationship: when empathy was present, the difference between the two rating types was reduced. Furthermore, we confirmed preference as one major predictor of empathy ratings. Emotional contagion was tested and confirmed as an additional direct effect of emotional expression on induced emotions. This study is among the first to explicitly test empathy and emotional contagion during music listening, helping to explain the often-reported emotional response to music in everyday life.

Received: January 20, 2012, accepted March 6, 2013.

Key words: emotion, music, emotional contagion, empathy, preference

In one listener, a sad piece of music might be perceived as expressing a negative emotional state and at the same time induce an unpleasant feeling. In another listener that same piece might induce a pleasant feeling. Since Gabrielsson (2001) called for a differentiation between expression and induction of emotion, there have been several experiments comparing the two phenomena (Evans & Schubert, 2008; Hunter, Schellenberg & Schimack, 2010; Kallinen & Ravaja, 2006). However, these studies have often remained rather exploratory in showing that different relations between the different emotion types exist, but have failed to give a theoretically grounded explanation of how the two different phenomena can be linked. We propose and test a model explaining the difference between expressed and felt emotions in music and show that this differentiation might be linked to the degree of self-rated empathy with the heard musicians, and to unconscious emotional contagion through emotional expression. Furthermore, we attempt to show how an independent conscious evaluation of a piece of music with regard to the listener's preference might predict empathy and create an additional response on its own.

Theories and Models on Music and Emotions

Emotion is defined according to the component process model (Scherer, 2004, 2005): An emotion episode consists of synchronized changes in several major components: (a) cognitive appraisal, (b) physiological arousal, (c) motor expression, and (d) subjective feeling. Many studies focus on measuring the subjective feeling component during music listening (Zentner & Eerola, 2010). It can be measured on two feeling dimensions, arousal (from calm to excited) and valence (from unpleasant to pleasant), as originally described by Russell (1980). The heuristic value of the two-dimensional emotion model was subsequently confirmed measuring emotional expression and induction through music (e.g., Egermann, Nagel, Altenmüller, & Kopiez, 2009; Nagel, Kopiez, Grewe, & Altenmüller, 2007; Schubert, 1999, 2001).

Gabrielsson (2001) described different relations between recognized and induced emotions: a positive relation exists when an expressed emotion is also felt by the music listener, a negative relation when the felt emotion is opposite to the expressed emotion, a non-systematic relation when an expressed emotion induces no feeling response or an expressed emotion elicits feelings of different qualities. No relation exists when no

Music Perception, volume 31, issue 2, pp. 139–156, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2013.31.2.139

emotion is expressed but a feeling is still induced. Gabrielsson and Lindström (2001) showed that the use of specific musical features (such as tempo, mode, articulation, dynamics) is associated with differentiated emotional expressions. Furthermore, many of these features are thought to show similarities to expressive voice features (Juslin & Laukka, 2003) and some might be culturally independent universals (Fritz et al., 2009). Focusing on emotion induction, Scherer and Zentner (2001) presented a theoretical account, describing several emotion production routes, including appraisal, empathy, memory, and peripheral arousal. Juslin and Västfjäll (2008, 2010) described several similar psychological induction mechanisms: Cognitive appraisal, brainstem reflexes, visual imagery, violating or confirming listeners' expectations (Egermann, Pearce, Wiggins, & McAdams, 2013; Huron, 2006; Meyer, 1956), evaluative conditioning, episodic memory, rhythmic entrainment, and finally, emotional contagion. This last mechanism is described as the internal "mimicking" of emotional expression in the music. This mechanism is the only one to link musical emotional expressions with induced emotions. Several empirical studies have investigated the difference in the two emotion types during music listening.

RESEARCH COMPARING FELT AND RECOGNIZED EMOTION MUSIC

Swanwick (1973, 1975) conducted experiments using simple two-tone patterns to assess two different ratings by music listeners: markings of "M" for meaning (expressed emotion) and "L" for listening (felt emotions). He did not present many results and noticed that there seems to be a "very complex" (Swanwick, 1973, p.12) relationship between the two emotion types. Gembris (1982) also asked his participants to provide two separate ratings at the same time (expression and induction) and calculated difference scores. He only presented selected results that often showed medium-sized correlations between the two kinds of ratings. Gembris finally hypothesized that a positive relation might be moderated by empathy and identification of the listener with the music. Kallinen and Ravaja (2006) tried to explain individual differences between recognized and felt emotions and their possible correlations with several personality measures using a within-subject design. In general, felt valence was rated as stronger than expressed valence, but expressed arousal was rated as stronger than felt arousal. Furthermore, they acknowledged that the emotional response to the music might have been related to familiarity and preference, emphasizing the need to consider these possible moderating factors. Investigating how measures of emotion influence musical preferences, Schubert (2007a) showed that intensity ratings of emotion recognition were higher than those of emotion induction within participants. Likewise, Hunter et al. (2010) compared ratings of recognized and felt emotional responses to music within participants and found that both were correlated, but expressed emotions were often rated as being more intense than felt emotions. Evans and Schubert (2008) reported that in 60% of participants' ratings, a positive relation was observed. Interestingly, pieces with a positive relation were preferred more, indicating that preference might play a role in modulating the relationship between expression and induction.

Summarizing this research, the following conclusions can be drawn. Often, within-subjects designs were used that might lead to less differentiated responses (Gabrielsson, 2001; Schubert 2007b). Recognized emotions were rated as more intense than felt emotions. In this study, we hypothesize that significant differences can be found between the two rating types in a between-subjects design as well. In previous research, a positive relation was most often observed between the two types of emotion. However, negative or non-systematic relations have also been reported, indicating a possible multi-causal relationship between induction and expression and other interacting mechanisms (Juslin & Västfjäll, 2008). Thus, these results suggest that the differentiation discussed by Gabrielsson can be confirmed, but detailed theoretical accounts linking emotional expression and induction are still missing.

EMPATHY, EMOTIONAL CONTAGION, AND MUSIC LISTENING

One psychological mechanism relating expression and induction of emotions is empathy. It involves two different components, one cognitive, "perspective taking," and another emotional, described as "feeling with someone else" (Preston & de Waal, 2002). The latter component will be the focus of this investigation. A similar concept is emotional contagion, which is described as an unconscious automatic mimicking of others' expressions affecting one's own emotional state (Juslin & Västfjäll, 2008). Emotional contagion differs from empathy in its degree of self- and other-awareness (Decety & Jackson, 2004). In empathetic reactions, a conscious distinction between one's own and others' feelings is experienced, whereas emotional contagion lacks this differentiation. Thus, empathy might be conceived as a rather voluntary top-down process, whereas emotional contagion is most often discussed as a primarily bottom-up process (Singer & Lamm, 2009). Emotional responses to music might be based on both: a) empathy with a performer or composer to whom expressions might be attributed (Scherer & Zentner, 2001), and b) a rather automated unconscious contagion through mimicking of expressive cues in the music (Juslin & Västfjäll, 2008). However, both phenomena can be understood as very related, because they create emotional responses that match those being expressed (Preston & de Waal, 2002).

The idea of emotional contagion through the internal mimicking of emotional expressions in music already dates back to ancient Greek philosophy, where music was thought to imitate nature, which is in turn imitated by the listener ("mimesis," Gabrielsson, 2001). This idea was also proposed by Hausegger (1887), who described music-induced emotions as resulting from the internal mimicking of the expressive movements that were necessary to produce them. Lipps (1909) even claimed that this mechanism, termed *Einfühlung* (later translated into English as "empathy"), a theory of inner imitation, even of non-living objects, would be at the core of aesthetic experience. These ideas concerning the link between movement, expression, imitation, and induction inspired several other music scholars (Clynes, 1977; Livingstone & Thompson, 2009; Vickhoff & Malmgren, 2004). It has also been proposed that action and perception in music are realized through shared neural representations based on the mirror-neuron system (Molnar-Szakacs & Overy, 2006; Overy & Molnar-Szakacs, 2009). Molnar-Szakacs and Overy (2006) state "that humans may comprehend all communicative signals, whether visual or auditory, linguistic or musical, in terms of their understanding of the motor action behind that signal, and furthermore, in terms of the intention behind that motor action." (p. 238)

In the study presented here, it was hypothesized that if participants rated higher degrees of empathy, differences between recognition and felt ratings would decrease. Studies on emotional contagion sometimes access expressive facial muscle activation in listeners observing a performer (Livingstone, Thompson, & Russo, 2009). However, during music listening without visual display of performers' movements, measurement of internal mimicking of movements remains difficult. Therefore we tested for emotional contagion by way of indirect statistical tests for automated emotion induction by emotional expressions in the music. We predicted that if expressions have predictive value for ratings of induced emotions regardless of self-rated empathy, then emotional contagion had occurred.

MODULATION OF EMPATHY IN MUSIC AND INDIVIDUAL DIFFERENCES

Several individual differences in empathy have been described as personality traits in previous research (e.g., males were described as feeling less empathy than females in general, Mehrabian, Young, & Sato, 1988). Furthermore, empathizing was described as one of two personality styles of music listening (as opposed to systemizing), and again males reported less empathizing (Kreutz, Schubert, & Mitchell, 2008). Thus, one might predict that individual differences in empathy might be associated with different personality traits and gender. We also hypothesize that music preference ratings predict state empathy ratings. Schubert (2007a) reported that participants preferred pieces with which they were familiar and which induced strong emotions that matched recognized emotions (see also Evans & Schubert, 2008). Accordingly, there seems to be a correlation between music preference and the positive relation between expression and induction. Schubert concluded that music is preferred that induces and expresses the same emotions. However, from our point of view, it could also be the case that expressed emotions are felt only for preferred pieces, indicating a possible moderator of state empathy. Further, Scherer and Zentner (2001) note that: "the process of empathy requires sympathy" (p. 369), suggesting that liking (which is quite analogous to sympathy in the German language, Scherer, 1998) might moderate empathy. They also mention that listeners might identify with a performer, a response that is only likely to occur if he/she performs music that is preferred and fits that individual listener's preferences.

Additionally, preference might also be considered to create a positive emotional response on its own (Schubert, 1996, 2009). Musical preference evaluations might potentially reflect a whole set of individual differences in personality, resting arousal level, and social identity (Rentfrow & Gosling, 2003). The identity-building function of music might be involved in cognitive appraisals during music listening (Scherer, 1999). These evaluations might be directly linked to the music listener's social identity, e.g., a listener whose goal is to belong to a certain group of people (who fit with her/his identity) might evaluate the music that this group listens to as goal-directed and derive positive emotions from listening to it (Egermann, Kopiez, & Altenmüller, 2013).

If this preference response and the empathy response diverge, a dissociated state is created with mixed feelings (Garrido & Schubert, 2011; Hunter, Schellenberg, & Schimmack, 2008). For example, a lover of sad music might like the negative sadness that it induces through contagion, but at the same time might have an additional positive response to it. Music listeners might also aim for emotional stimulation in general regardless of the quality of emotion, and sensation seekers especially have been shown to prefer risky behavior and highly

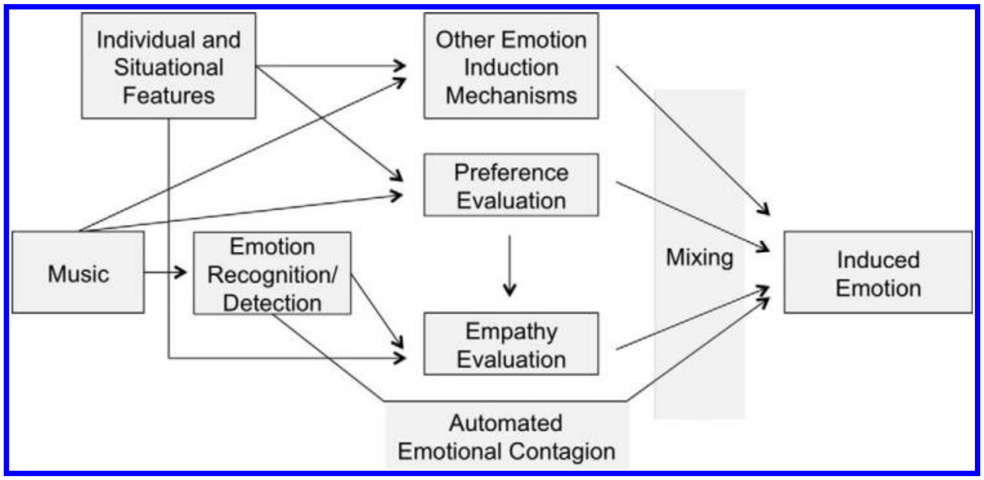

FIGURE 1. Model illustrating the hypothesized relationship between recognized and induced emotions in music listening.

stimulating music (Litle & Zuckermann, 1986). Listening to music in an aesthetic context could allow simulating virtual emotions and feelings that would be avoided in everyday life. For example, listening to scary music without any reason to be scared might be enjoyable to some people, but if the fear becomes real, we assume that scary music would be less likely to be preferred and pleasant.

Accordingly, we predict that empathy might be moderated by several individual differences such as gender and music preferences. However, this preference evaluation could also be understood as a second appraisal process that not only evaluates the emotion-inducing events, but also the resulting emotions themselves, modulating empathy and potentially creating an additional emotional response (with possibly mixed negative and positive feelings).

PROPOSED MODEL AND AIMS OF THE STUDY

We propose the following model to explain the relationship between expressed and felt emotion in music (Figure 1). First, on the left side, a piece of music with a recognized emotional expression is perceived by a listener and elicits several types of emotional responses, including a preference response due to a preference evaluation of that specific music and an empathetic resonance with, and contagion through, the emotional expressions. Preference and empathy are also interrelated: the preference evaluation mechanism can modulate (intensify or lower) the empathetic reaction. Both evaluations also have other inputs that are represented as individual and situational features. However, emotional contagion occurs automatically, is not mediated by preference and conscious empathy (indicated by the direct arrow between recognized/detected expressions and induced emotions). Furthermore, there are several other emotion-induction mechanisms (Juslin & Västfjäll, 2008), which will not be investigated here because of the focus on empathy, contagion and preference. Finally, all emotional outcomes are integrated via mixing (Hunter et al., 2008) into a consciously available induced emotional response. In the present study, we operationalized and measured several components of this model by assessing for every musical excerpt listened to (a) the overall valence (pleasantness) and arousal of expressed and induced emotions, (b) the self-rated empathy, (c) the preference for the excerpt, (d) several other individual and situational features that might influence empathy, and finally, (e) a statistical test for emotional contagion.

We hypothesized that arousal and valence ratings of induced emotions would be different from corresponding expressed emotion ratings in music in a between-subjects design (H1). Furthermore, we predicted that the higher the self-rated empathy, the smaller the differences between expressed and felt emotions would be (H2). As indicated by previous research, self-rated empathy with the performing musicians was thought to be predicted on the basis of preference evaluations of the music and other individual and situational factors like gender, attention, and musical expertise (H3). Furthermore, we evaluated whether ratings of felt valence and arousal can be predicted using emotional expression, empathy, and preference evaluations as in the model in Figure 1. A significant contribution of emotional expression will be interpreted as evidence for automated emotional contagion, whereas an interaction between expression and empathy ratings will be interpreted as representing the consciously available empathy evaluation (H4). Furthermore, we predict an independent main effect of preference on measures of felt emotions (H5), indicating a preference response to the music (Schubert, 1996, 2009).

Method

ONLINE MUSIC PERSONALITY TEST

The five hypotheses were tested in an online music listening setting. Web-based experimenting has many advantages compared to conventional lab experiments (Honing & Ladinig, 2008; Reips, 2002), such as bigger sample sizes, situating participants in a natural environment, and less researcher bias. There are also some disadvantages of this experimental approach, such as lack of direct control over the experimental setting, dropout of participants, and technical variability (e.g., soundcards, speakers, screens), potentially leading to higher random variation in dependent measures. However, this method can increase external validity and through thoughtful experimenting standards also allow investigations with high internal validity. We therefore included an instruction test as a high hurdle, a warm-up phase, and removed participants who rated themselves as not focused and serious (Reips, 2002). Furthermore, Egermann, Nagel, et al. (2009) have shown that online experiments provide valid results when accessing emotion induced by music and have applied that methodology to the measurement of social influences on musically induced emotions (Egermann, Grewe, Kopiez, & Altenmüller, 2009; Egermann, Kopiez et al., 2013).

As a cover story, the experiment was embedded in a German music personality test. Here, after listening to and rating the music, participants received their personalized test results describing links between their music preferences and personality (based on a study from Rentfrow & Gosling, 2003). However, at the end participants were presented with a disclaimer pointing out that personality descriptions were only based on weak correlations between music preferences and personality traits. They were not informed about the study's hypotheses and were encouraged to complete the test in order to reach the personalized results section. A visually appealing Flash-applet (Adobe Flash, Version 9.0) was programmed. It presented instructions, played back the music, displayed the questions, and collected answers from participants. It was connected via PHP (PHP: Hypertext Preprocessor, Version 4.4.2) to a MySQL-database (MySQL, Version 5.0), which stored all the data.

PARTICIPANTS

Participants were recruited by links to the study from several German websites. Three thousand one hundred and sixty-four participants (age in years: $M = 32$; $SD = 12$) completed the study. Fifty-eight percent were male and 42% female; 51% were not involved in music performance activities, 43% were involved as amateur musicians, and 5% as professional musicians by self-report.

STIMULI

In order to reduce participation time, all participants listened to a subset of five musical excerpts (in random order) randomly chosen from a total of 23 excerpts (30 s each). On average, each excerpt was rated by 687 participants ($SD = 19$). The excerpts were experimenter-selected to represent all four emotional quadrants resulting from the two-dimensional emotion model by Russell (1980), thus expressing negative or positive valence with low or high arousal. The excerpts were also selected to cover a wide range of possible preferences and styles, aiming to use music that was very liked and also very disliked. Half of the stimuli were instrumental ($n = 11$) and the other half a cappella ($n = 1$) or vocal with instrumental accompaniment ($n = 11$). In order to omit personal associations with the stimuli used (see Juslin & Västfjäll's, 2008, emotional conditioning and episodic memory mechanisms), which would make emotional responses more individual and difficult to predict, we tried to select mostly music that was unknown to an average music listener. A complete list of all stimuli used can be found in Appendix A. To play the music, participants used headphones (31%), internal speakers (28.5%), external speakers (31%), a stereo (9%), or other equipment (0.5%) as determined by self-report.

PROCEDURE AND STUDY DESIGN

Every site visitor was randomly directed to one of two versions of the music personality test. Participants directed to one version rated emotions recognized in the music ($n = 1{,}545$), whereas those in the other version rated their own felt emotions ($n = 1{,}619$). At the beginning of the study, all participants were asked to read instructions regarding their tasks. Participants were explicitly instructed to differentiate emotions expressed by the music from those felt by them. Subsequently, they could test the music playback capabilities of their computers. After that, they had to pass an instruction comprehension test that assured that they understood that their task was to rate either recognized or felt emotions and that they understood the definition of valence correctly. Subsequently, they could take part in a test trial as a warm-up. Following music listening and rating of five excerpts, participants answered questions concerning their socio-demographic background and their music preferences, including two questions about their ability to distinguish between felt and recognized emotions and general empathy during music listening in everyday life. Then they were presented with their personalized test results (based on general music style preference ratings, not the excerpt ratings) and had the opportunity to take part in a lottery. Here they could win one of five Amazon vouchers amounting to 10 EUR by pulling the lever of a virtual gambling machine. Completing the test took 10.6 minutes on average (*SD* = 2.9 min).

MUSIC RATINGS

After each excerpt, participants rated its recognized or induced emotions (depending on the rating condition) using the arousal and valence dimensions, plus felt empathy with the musicians, familiarity, and preference. The ratings were made by moving five sliders on five visual analog scales from negative (indicated by the minus sign) to positive (indicated by the plus sign). All music-rating scales were internally scored on a scale from 0 (negative) to 100 (positive). The empathy scale was entitled: "Did you empathize with the musicians you just heard?" (in German: "Haben Sie mit den so eben gehörten Musikern mitgefühlt?"). The familiarity and preference sliders were labeled with "Did you know that piece?" and "Did you like that piece?"

Following Russell (1980), negative valence was previously defined as "unpleasant" feeling (in German: "unangenehm") and positive valence as "pleasant" feeling (in German "angenehm"). We followed Russell's original definition of valence as pleasantness of the induced feeling, instead of the often used "emotional valence" for which participants have to rate the valence of this emotion as if it would occur in an everyday life context (Colombetti, 2005). Following the "emotional valence" definition, sadness induced by music would have to be rated as negatively valenced, whereas in our definition of valence, participants only rated the overall pleasantness of the induced feeling, thus allowing mixed feelings (Hunter et al., 2008) to be reported. Negative arousal was defined as "calming" (in German: "beruhigend") and positive arousal as "arousing" (in German: "erregend").

DATA ANALYSIS

To ensure data quality for every data set, we established a number of criteria that had to be fulfilled by the participants. Only participants completing the whole study were included. They were also asked to indicate on a visual analog scale how serious they were about taking part and, after participation, how concentrated they had been during the session. Scales ranged from 0 ("not serious/not concentrated") to 100 ("very serious/very concentrated"). Participants rating their concentration and seriousness lower than 50 were excluded from the data analysis. Participants who failed the comprehension test more than three times were not included in data analysis, as were those whose participation times were less than five minutes or more than 30 minutes. Following these criteria, 2,500 (44.1%) participants were excluded from the 5,664 participants beginning the experiment, representing an average dropout/exclusion rate comparable to other web experiments (Egermann, Nagel, et al., 2009).

All statistical analyses were based on a linear mixed modeling approach (West, Welch, & Galecki, 2007) that estimated significant coefficients controlling for random sources of variance. This approach analyzes the data in a way similar to a regression analysis testing predictor variables of outcome variables. We included partially crossed random effects for participants and items (musical excerpts), as suggested by Baayen, Davidson, and Bates (2008). Equation 1 illustrates the general model formulation by these authors:

$$y_{ij} = X_{ij}\beta + S_i s_i + W_j w_j + \varepsilon_{ij} \quad (1)$$

where, y_{ij} represents the responses of subject *i* to item *j*. X_{ij} is the experimental design matrix, consisting of an initial column of ones (representing the intercept) and followed by columns representing factor contrasts and covariates. This matrix is multiplied by the population coefficients vector β. The terms $S_i s_i$ and $W_j w_j$ help make the model's predictions more correct for the subjects and items (musical excerpts) used in the experiment, respectively. The S_i matrix (the random effects structure for subject) is a full copy of the X_{ij} matrix. It is multiplied with a vector specifying for subject *i* the adjustments required. The W_j matrix represents the item *j* random effect and is again a copy of the design matrix X_{ij}. The vector w_j contains adjustments made to the population intercept for each item *j*. The last term is a vector of residual errors ε_{ij}, including one error for each combination of subject and item.

As suggested by Baayen et al. (2008), all analyses were done using the software R (2.13) using the lmer function from the lme4 package (Bates, Maechler, & Bolker, 2011). Estimation of parameters was based on Restricted Maximum Likelihood (REML) and likelihood ratio tests were used to test for random effects. To automate the process of fitting, the fitLMER function was employed (Tremblay, 2011). It first back-fits fixed effects (by omitting fixed effects and their interactions with *t* values smaller than 2) and subsequently forward-fits random effects of lmer models. After identifying significant random effects, the procedure back-fits all fixed effects again. Finally, Markov chain Monte Carlo (MCMC) sampling of all significant model parameters produced their mean estimates, their associated confidence intervals, and *p* values (alpha errors). Modeling assumptions were finally tested by inspecting the model criticism plots produced by the associated mcp function (Tremblay, 2011). Measures for explained variance of the different models were derived by predicting outcome values with the corresponding values fitted by the model, using the simple regression function lm and reporting R^2 values. Because the distributions of empathy and preference ratings showed a clear bimodal distribution (with peaks in the acceptance and rejection regions of the corresponding rating scales), both variables were recoded in dichotomous format (with a split in the middle of the rating scale). This allowed for easier plotting and interpretation of the predicted interactions with these two variables.

Results

TREATMENT CHECK AND DESCRIPTIVE STATISTICS

First, we explored whether the different musical excerpts selected were rated as having different emotional expressions and inducing different emotions. As can be seen in Figure 2, the mean ratings of selected music pieces are quite spread out over the emotion space, covering every quadrant. For some excerpts, differences between felt and recognized ratings appear small (e.g., Nos. 1, 2, 7, 16, and 17), for some, recognized ratings represent opposite emotions (e.g., Nos. 3, 8, and 15), and for some, recognized emotion ratings appear to be more intense (i.e., they are farther from the origin) than felt emotion ratings (e.g., Nos. 4, 5, 11, 19, 23).

Furthermore, the 23 excerpts selected can be shown to induce different levels of empathy and preference (Figure 3). Some pieces were preferred/empathized with very strongly (e.g., excerpt No. 6), and others were disliked/not empathized with (e.g., excerpt No. 8). On average, most pieces were rather unfamiliar ($M = 12$, $SD = 27$), with one exception, excerpt No. 6, the familiarity rating of which was 93 on average ($SD = 18$). However, this outlier did not influence later modeling results, as models estimated without it were not different from those including it. Participants indicated that they were more or less able to distinguish between recognized and felt emotions with an average group rating of $M = 61$ ($SD = 26$) on a rating scale from “0” (“difficult to distinguish”) to “100” (“easy to distinguish”).

DIFFERENCE BETWEEN RECOGNIZED AND FELT EMOTION AND MODULATION VIA EMPATHY

The first hypothesis in this study predicted a significant difference between ratings of recognized and induced emotions. Furthermore, it was assumed that this difference would be smaller when self-rated empathy was present. Both assumptions were corroborated by the data in this experiment.

In general, ratings of recognized valence and arousal were different from those of induced feelings, and the distance between the two groups was smaller when participants indicated that they empathized with the musicians being heard (see Figure 4). In order to test whether these observations were significant, we built two linear mixed models for arousal and valence ratings (Table 1). These models included three significant fixed effects for rating type (being in the induced emotion rating group was represented with a negative coefficient), reporting high empathy (with a positive coefficient), and the interaction of the two (also with a positive coefficient). A positive coefficient represents a positive influence of the parameter on the dependent outcome variable, whereas a negative coefficient represents the opposite, negative influence. The coincidence of both being in the induced emotion rating group and feeling empathy led to a change of the groups’ associated negative coefficient for arousal and valence, indicating that there was a reduced difference between recognized and felt emotions if participants indicated that they felt empathy with the musicians. Furthermore, two significant random effects as intercepts for subjects and musical excerpts were observed, controlling for intraclass correlations within the dataset and generalizing the results beyond the participants and stimuli used.

MODERATORS OF EMPATHY

After identifying empathy as moderating between felt and recognized emotions, we tested which individual and situational factors influenced this empathy rating. Therefore, we constructed another linear mixed effects model predicting the rated empathy of each of the excerpts listened to using the ratings of preference and familiarity, as well as participants’ self-rated seriousness, concentration, ability to differentiate between recognized and felt emotions (ADRF), and their gender, age, education (highest school degree attained), musical expertise, and the participation duration (in minutes).

In order to reduce multicollinearity between those predictors, we applied a factor analysis to those items

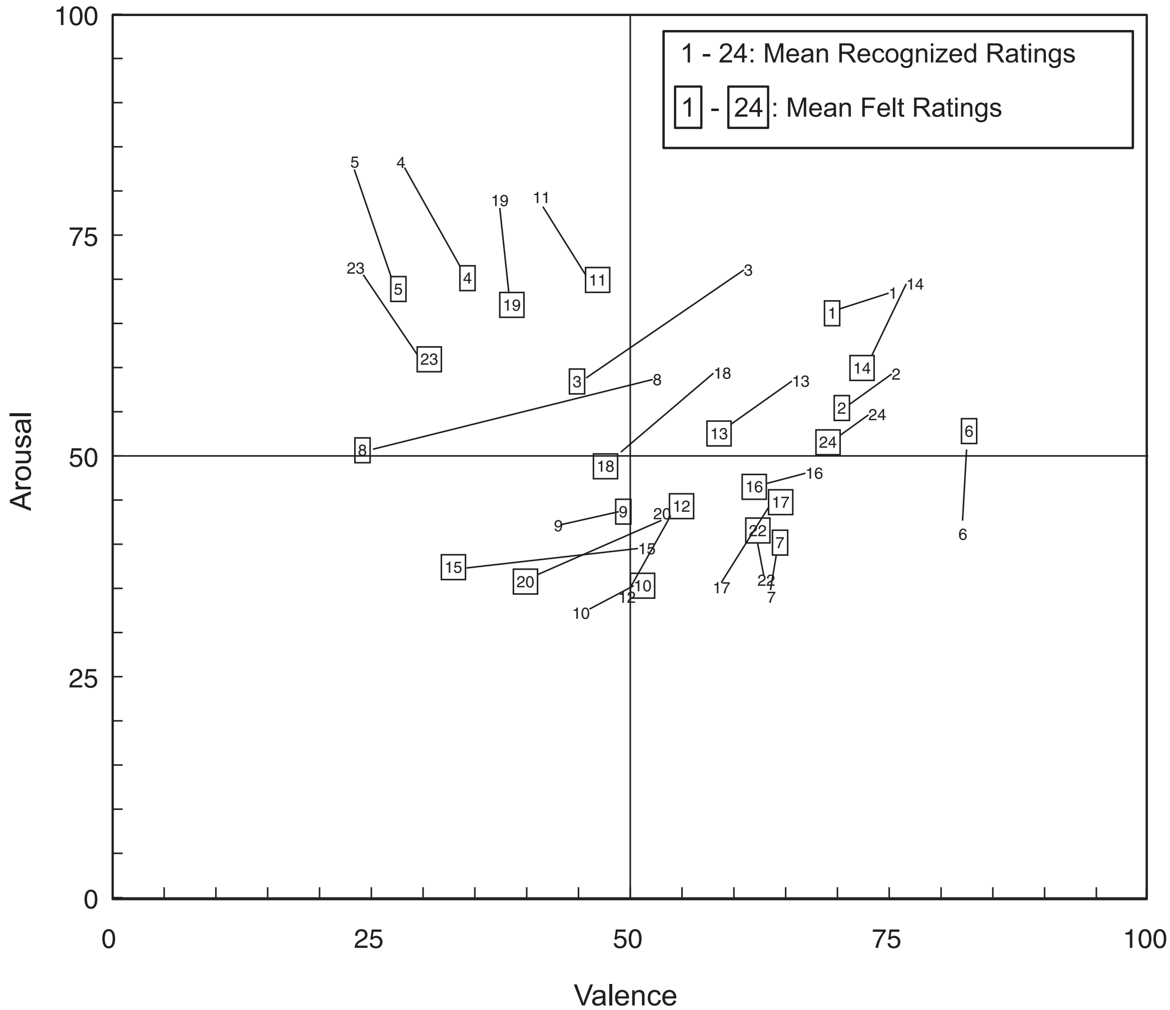

FIGURE 2. Emotion space with valence (pleasantness) and arousal dimensions and mean ratings of musical excerpts separated by rating type. Numbers in the space represent song numbers (see Appendix A). Mean ratings of participants in the felt emotion rating group are shown with boxes and those in the recognized emotion rating group are without boxes.

that were correlated and rated on continuous scales. The extraction was based on a principal component analysis using the varimax rotation method with Kaiser normalization. The rotation converged after three iterations. The factor solution was adequate for these data (Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .54, Bartlett's Test of Sphericity: approx. chi-square = 5722.74, $df = 10$, $p < .001$). The result was a structure consisting of two predictors: representing a) quality of participation (with high loadings of concentration, seriousness, and ADRF) and b) a combination of familiarity and preference ratings.

Automated backwards fitting of fixed effects produced five significant coefficients (see Table 2). Effects of age and education were not significant. Significant positive predictors (the higher their values were, the higher the empathy rating) included: the preference/familiarity factor (explaining most of the deviance), the quality of participation, and—with the smallest effect size—participation time. Furthermore, being male (as opposed to being female) and not being involved in music performance activities (as opposed to being an amateur or professional musician) were significant negative predictors of empathy. When participants were asked after the music listening about feeling empathy with musicians in general, those with musical expertise also provided higher ratings than those without (no musical performance activity: M = 58, SD = 28, vs. musicians: M = 68, SD = 25, $F(1, 3158) = 120.67$, $p < .001$, rating scale from 0 to 100).

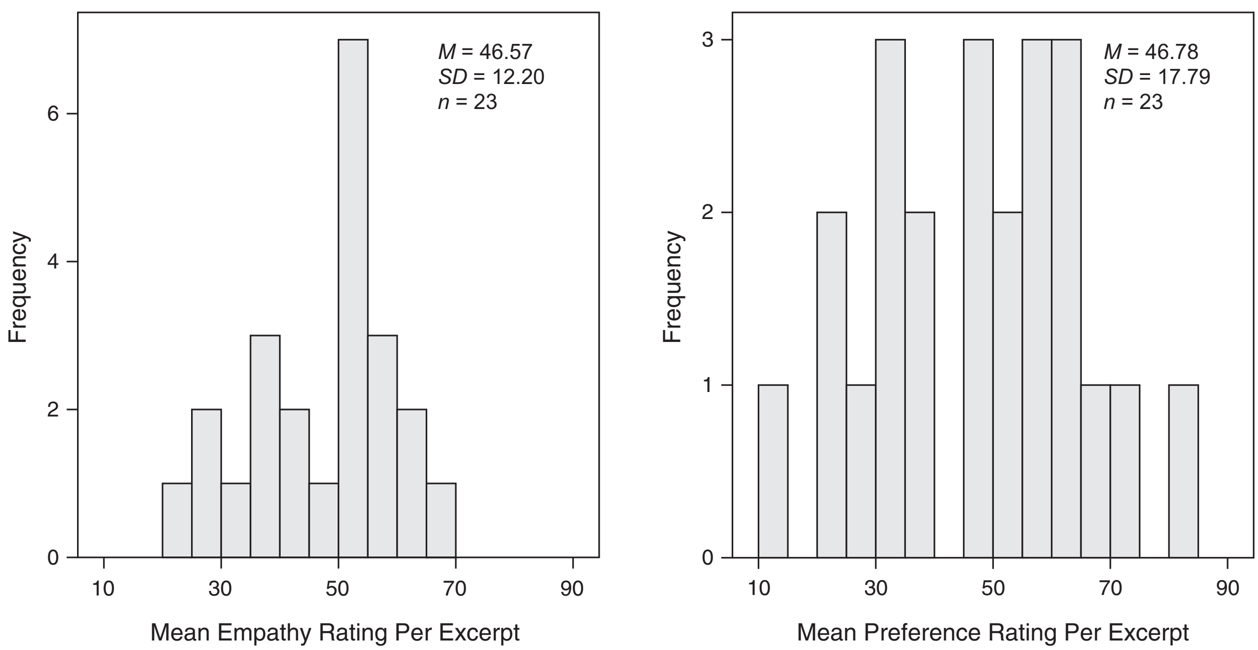

FIGURE 3. Histograms showing the distributions of mean preference and empathy ratings of all musical excerpts presented to participants.

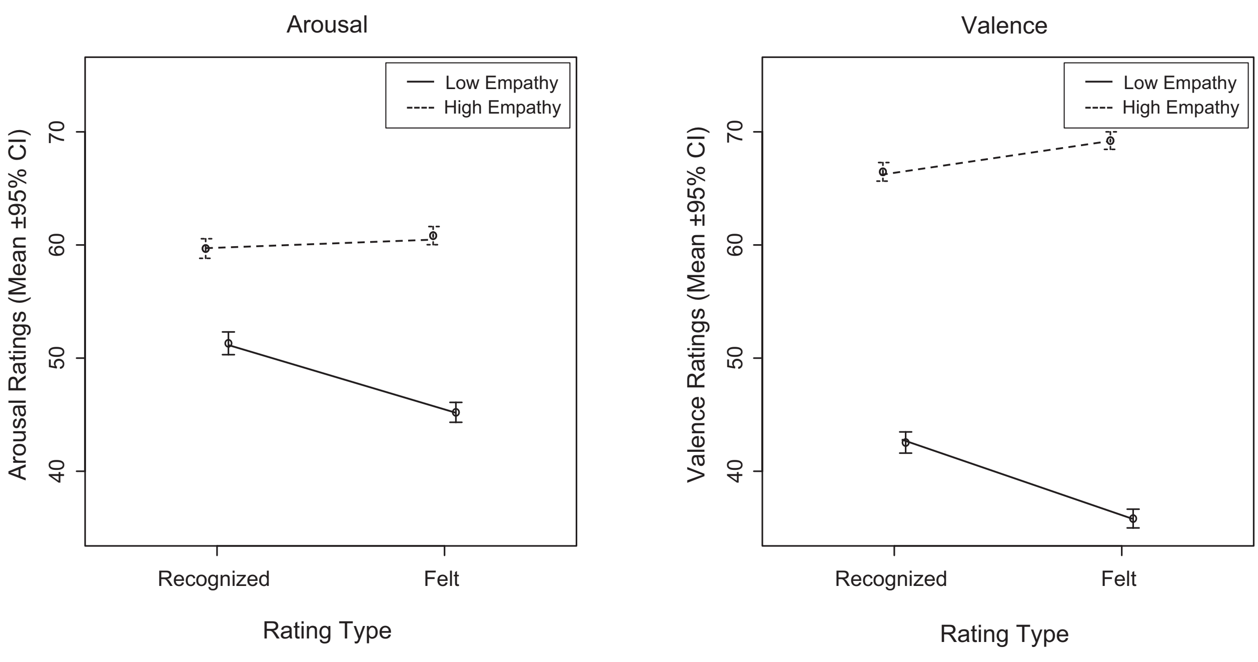

FIGURE 4. Mean arousal and valence ratings separated by rating type (group) and self-rated empathy. Dots represent mean values and bars 95% confidence intervals.

Because the role of preference as a moderator of empathy independently of familiarity was of interest for subsequent analyses, the modeling procedure described above was repeated with only preference ratings and all other significant predictors excluding familiarity (not shown here). The use of preference resulted in an increased explained variance and fit of the model (model with combined preference/familiarity factor, adj. $R^2 = .60$, AIC $= 144104$, AIC indicates goodness of model fit with smaller values, vs. model with only

TABLE 1. ***Mixed Effects Modeling Parameter Estimates for Arousal and Valence Ratings Predicted via Rating Type (Recognized vs. Felt Emotions) and Empathy Ratings.***

Fixed Effects	Arousal Mean Estimated Coefficients[1]	Valence Mean Estimated Coefficients[1]
Intercept	49	46
Group (being in felt rating condition)[2]	−6***	−7***
Empathy (high empathy)[2,3]	12***	17.3***
Group* Empathy[2]	6***	9.87***
Random Effects	**Variance (SD)**	
Subject Intercept[4]	72.45 (8.51)***	38.45 (6.20)***
Musical excerpt Intercept[4]	225.74 (15.02)***	167.22 (12.93)***
Residual	537.18 (23.1)	537.18 (23.18)
R^2	.435	.46

*Notes: $n = 15{,}800$; [1]MCMC Sampling ($n = 10{,}000$); [2]Dummy variables; [3]Empathy recoded: 0-49 = 0 (low empathy), 50-100 = 1 (high empathy); [4]Chi-Sq, Log like Test; ***$p < .001$.*

TABLE 2. ***Mixed Effects Modeling Parameter Estimates for Empathy Ratings Predicted via Situational and Personal Features.***

Fixed Effects	Mean Estimated Coefficients[1]
Intercept	45.7
Preference/Familiarity[3]	19.8***
Quality of Participation[3]	2.0***
Gender (being male)[2]	−2.1***
No Musicianship (being not active in music performance)[2]	−1.3**
Participation time (in min)	0.3**
Random Effects	**Variance (SD)**
Subject Intercept[4]	133.29 (11.545)***
Musical excerpt Intercept[4]	68.426 (8.272)***
Residual	442.018 (21.024)
R^2	.60

*Notes: $n = 15{,}800$; [1]MCMC Sampling ($n = 10{,}000$); [2]Dummy variables; [3]Extracted factor values; [4]Chi-Sq, Log like Test; **$p < .01$, ***$p < .001$.*

preference as predictor without familiarity, adj. $R^2 = .67$, AIC = 140875).

MODELING RESULTS: FELT EMOTIONS

Finally, we constructed two linear mixed models to predict only the ratings of felt arousal and valence as two final tests for the suggested model that links recognized and felt emotion in music listening (Figure 1). Here, we could differentiate the influence of emotional contagion and empathy with emotional expression in the music. Additionally, these models tested for an independent main effect of preference on emotion without any involvement of expressed emotions.

Similar to previous models, two significant random intercepts for subjects and musical excerpts were included. As we employed a between-subjects design, no measures for recognized emotions were taken from participants in the felt emotions rating condition. But because descriptors of emotional expression were required for these analyses, we estimated the expressed arousal and valence for each excerpt by computing its corresponding mean arousal and valence values and categorizing them as either positive or negative (below or above the middle position on the corresponding rating scale). The creation of these dummy variables facilitated graphical presentation and interpretation of interaction effects. Modeling results were the same when continuous predictors were used (not shown here). Furthermore, we tried to test for the influence of empathy independently of preference. Therefore, we used residual empathy values that were derived from predicting empathy using only preference ratings (using a mixed effects model similar to those in the previous sections but with only one fixed predictor). These

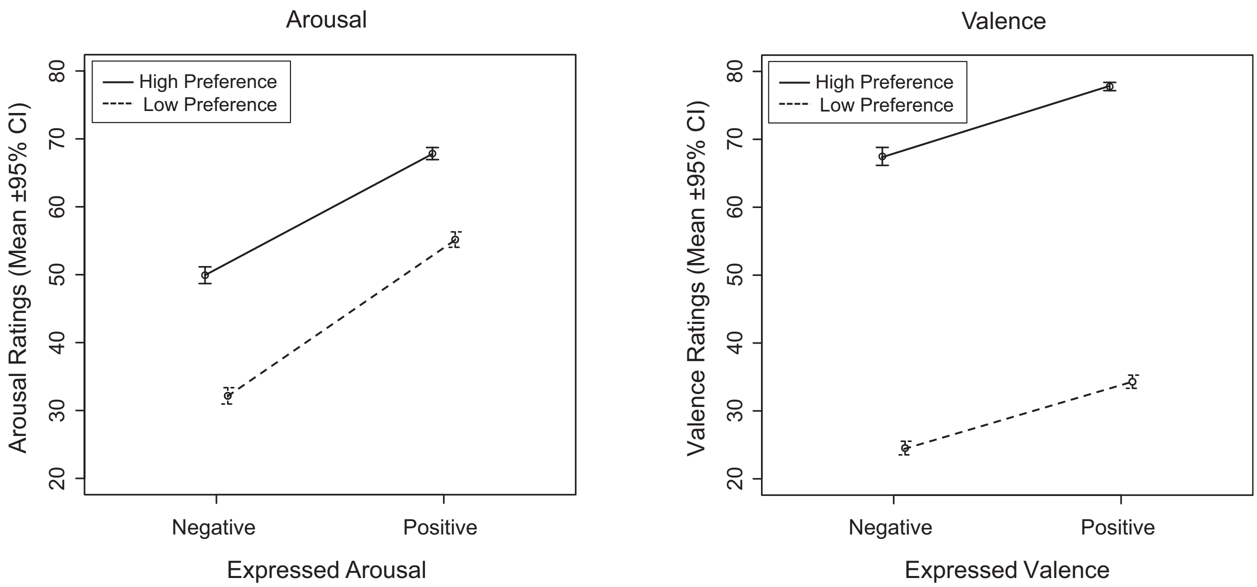

FIGURE 5. Mean arousal and valence ratings separated by expressed emotion and preference. Dots represent mean values and bars 95% confidence intervals.

residual empathy values were correlated to the real empathy ratings, $r(15798) = .67$, $p < .001$, but were independent of the preference ratings, $r(15798) = -.00007$, *ns*.

Figure 5 presents group mean values of arousal and valence ratings. If positive arousal or valence was expressed in the music, ratings on the corresponding dimensions were higher as compared to negative expressions. Furthermore, ratings were higher when the piece listened to was liked compared to disliked pieces on both emotion dimensions.

Felt valence ratings were modeled using the following fixed predictors: preference, residual empathy (after removing the effect of preference ratings), expressed valence, and their interactions. Significant main effects were found for preference, residual empathy, expressed valence, an interaction between preference and residual empathy, and an interaction between residual empathy and expressed valence. The three-way interaction was omitted from the automated model fitting process due to its lack of significant contribution. The three main effects were all positive, with preference having the biggest coefficient followed by expressed valence and residual empathy (the first two main effects are displayed in Figure 5). The coefficient related to the interaction between the latter two terms (the coincidence of residual empathy with positive expressed emotion) was also positive, whereas the interaction between preference and residual empathy (the coincidence of residual empathy with preference) was negative. This interaction between preference and residual empathy with a negative coefficient of −6.19 might be understood in the following way: when preference is present, the additional valence associated with residual empathy (with a coefficient of 7.35) is removed from the model.

Subsequently, we built a model to predict felt arousal ratings using measures of preference, residual empathy, expressed arousal and all possible interactions. The automated fitting procedure removed all interactions in the model, resulting in three significant positive main effects for preference, expressed arousal and empathy (Table 3, see Figure 5 for main effects of preference and expressed arousal).

Discussion

All five hypotheses were supported: There was a significant negative main effect of being in the induced-emotion rating group as opposed to the recognized-emotion rating group (induced-emotion ratings were more negative and calmer, H1). There was a significant positive interaction effect indicating that self-rated empathy modulates the positive relation between recognized and felt emotions in music listening (H2).

Self-rated empathy was shown to be moderated by a listener's preference for the musical excerpt and other individual and situational features (H3). The preference/familiarity factor explained most of the variance

TABLE 3. *Mixed Effects Modeling Parameter Estimates for Felt Valence and Arousal Ratings Predicted via Preference, Empathy, and Emotional Expressions.*

Fixed Effects	Arousal Mean Estimated Coefficients[1]	Valence Mean Estimated Coefficients[1]
Intercept	28.53	23.27
Preference (preferring excerpt)[2]	15.76***	40.91***
Residual Empathy (high empathy)[3]	8.77***	7.35***
Expressed Emotion (positive arousal or positive valence)[2]	20.82***	9.19**
Preference * Residual Empathy[2]	—	−6.19***
Residual Empathy *Expressed positive valence[2,3]	—	3.00**
Random Effects	Variance (SD)	
Subject Intercept[4]	76.89 (8.77)***	29.28 (5.41) ***
Musical excerpt Intercept[4]	56.18 (7.50) ***	34.56 (5.88) ***
Residual	525.22 (22.92)	363.30 (19.06)
R^2	.42	.64

Notes: $n = 8{,}075$; [1]MCMC Sampling ($n = 10{,}000$); [2]*Recoded dummy variables: 0-49 = 0 (no preference, negative valence, negative arousal), 50-100 = 1 (preference, positive valence, positive arousal);* [3]*Recoded dummy variable: residual empathy < 0 = 0 (low empathy), residual empathy > 0 = 1 (high empathy);* [4]*Chi-Sq, Log like Test; **p < .01, ***p < .001.*

in empathy, but further analysis showed that preference alone correlated even stronger with empathy. This finding might be due to the lack of variance in the familiarity ratings, because most pieces were unknown to the participants. Males rated empathy slightly lower than females, a finding that corresponds to previous research on empathy (Kreutz et al., 2008; Mehrabian et al., 1988). However, it remains to be seen whether males really felt less empathy compared to females, or if they only reported less due to sociocultural norms. If the former is the case, this might explain the previously reported gender differences in studies on emotional responses to music (Nater, Abbruzzese, Krebs, & Ehlert, 2006; Panksepp & Berntzki, 2002). Musicians also reported more empathy than those without musical activity, a finding that might be explained by the fact that musicians might have identified themselves more strongly with the performing musicians. Furthermore, they might have had more knowledge about performing the music and were more likely to internally imitate the movements associated with playing the music (Overy & Molnar-Szakacs, 2009). Among other reasons, this finding might also explain why musicians have been shown to respond more strongly to music emotionally than people with less musical expertise (Grewe, Kopiez, & Altenmüller, 2009). However, one would also need to test whether it was only the ratings of empathy that increased or whether empathetic reactions also increased within this subgroup. Predicting empathy, we also observed a significant contribution of the quality of participation factor: the more concentrated, serious, and capable of distinguishing between recognized and felt emotions (ADRF) participants were, the higher their ratings of empathy. The last variable (ADRF) might be interpreted on the one hand as a measure of focus on the instructions and on the other as a measure of general empathy with musicians. If someone is not able to distinguish between recognized and felt emotions, he/she might in general be very empathetic with musical expression. However, general musical empathy and ADRF were only very weakly correlated, $r(3158) = .19$, $p < .001$. Further, paying attention to the expressed emotions in music is associated with higher empathy, a correlation that could also explain the observed effect of self-rated attention on the number of strong emotions reported (so-called "chills") and arousal (Egermann, Nagel, et al., 2009).

Finally, we tested the predictability of induced valence and arousal ratings in the corresponding rating group (H4 and H5). Here, we identified positive main effects for preference, expressed emotions, and empathy ratings. Two interactions were significant for valence ratings, none for the arousal ratings. One was interpreted as a ceiling effect eliminating the effect of additional empathy when preference was stronger. The other interaction effect might represent the empathy-mediated modulation of the effect of emotional expression, similar to the findings discussed with respect to the second hypothesis of this study. Due to the two independent main effects of preference and expression on valence and arousal, mixing of congruent or non-congruent (opposite) emotions was also confirmed: if non-congruent emotions were present (e.g., listening to a preferred piece with negative emotional expression), overall ratings of subjective pleasantness or arousal were lower than when congruent emotions were present (e.g., listening to a preferred piece with positive emotional expression).

A MODEL LINKING EXPRESSION AND INDUCTION

The confirmation of all five hypotheses might be interpreted as evidence for the model proposed in Figure 1. We demonstrated the role of consciously available empathy in moderating whether expressed emotions are felt, and at the same time we identified an independent effect of emotional expression on induced emotions that might represent an automatic emotional contagion with recognized emotional expression (direct, non-moderated link between expressions and feelings, Juslin & Västfjäll, 2008). Thus, empathy and emotional contagion appear to be very similar during music listening, but still represent different mental processes.

We showed that empathy was moderated by different individual and situational features, and we identified music preference as one of the strongest of them. This preference evaluation might reflect a whole set of sub-appraisals that are similar to the appraisal dimensions of emotions (Juslin & Västfjäll, 2008; Scherer, 1999). Preference might moderate empathy and contagion through emotion regulation, which has been described as being achieved in several ways: selection and modification of the emotion-inducing situation, deployment of attention, cognitive change (reappraisal), and modulation and suppression of responses (Gross, 1998, 2002). Liking a piece of music might lead to increased attention to expressed emotions and thus increased empathy. Reappraisals have also been investigated as so-called *meta-appraisals* in order to describe the enjoyment of negative media content such as sad movies (Schramm & Wirth, 2010). Thus, in addition to the music, the induced feeling might also be evaluated on its own, and these evaluations could probably create an additional emotional response, as demonstrated here through the main effect of preference on valence/pleasantness and arousal (cf. preference response reported by Schubert, 2007a). However, no direct measures of all these appraisals have been collected here and conclusions remain speculative.

Rating the overall pleasantness and arousal of the music-induced feelings, we found evidence for a mixing of different emotional responses to music listening, because preference and expressed emotions independently influenced ratings. Models could only explain a part of the observed variance in emotion ratings. This emphasizes other emotion-induction mechanisms that were not investigated here (Juslin & Västfjäll, 2008).

We found evidence for an empathy-modulated effect of expression on induced emotion ratings. However, predicting only felt emotions, we only found evidence for this modulation in valence ratings, not in arousal ratings. This could mean that for arousal, empathy modulation of emotional expression is weaker than for valence or there is more automated contagion/mimicry along this dimension. But this might also be related to the rather simplified approach of estimating emotional expression as group means of corresponding valence and arousal ratings. However, some studies also show individual differences in emotion recognition in music (Vuoskoski & Eerola, 2011), which might also explain why no modulation of preference in the effect of estimated emotional expression was significant in both felt valence and arousal models.

Furthermore, we found positive main effects of empathy ratings on recognized and felt emotion ratings, a finding we did not expect. This also occurred when preference-independent residual empathy values were used. Thus, empathy might have to be interpreted as creating an additional emotional response on its own. Previous research has shown that people scoring high on empathizing personality scales are better at emotion recognition than those scoring lower on these measures (Besel & Yuille, 2010). However, an effect on the intensity/positivity of recognized emotions has not been reported and remains of interest for further investigation.

GENERAL IMPLICATIONS OF THESE FINDINGS

This study employed an innovative methodological approach of web experimenting. Although there are limitations due to the lack of direct control over participants and the experimental context, the data recorded from more than 3,000 participants provide quantitative evidence that is rarely reported in experimental psychological studies.

From these results, we might also learn something about the often discussed social function of music (e.g., Livingstone & Thompson, 2009). Expression of, recognition through, and contagion/empathizing with, emotions might be understood as a form of emotional communication (Egermann, 2010). These interactions might occur between musicians and listeners, but also among listeners. In social settings such as concerts or nightclubs, where groups of people are listening to music together, emotional responses might also be influenced by empathy and contagion with each other (Egermann et al., 2011).

Another possible function of music illustrated by these results might be to empathize with and experience emotions that would be avoided if they were real (e.g., sadness or anger). The creation of this possible dissociative state (Herbert, 2011; Schubert, 1996, 2009) might be one important goal leading listeners to engage with music. This might also represent the rather complex interaction of the different emotion-induction mechanisms (Juslin & Västfjäll, 2008).

The results of this study might also be used to improve music recommendation systems. They sometimes employ mechanisms that are capable of recognizing emotional expression in the music based on structural audio features (like the *miremotion* function in the MIR toolbox; Eerola, Lartillot, & Toiviainen, 2009). This information is then offered as a search criterion for users. However, as previous research has shown (Juslin et al., 2008), listeners often do not listen to music to recognize emotions in music; rather, they listen to feel the emotions recognized. Adding information about user's preferred musical styles to the search algorithm might then allow predictions of whether the expressed emotion will be felt in the listener.

LIMITATIONS AND FUTURE RESEARCH SUGGESTIONS

We were only able to provide correlative evidence for many of the factors of interest. For example, we assume that due to the observed effects, preference moderates empathy, but of course it could be the other way around: we only prefer pieces for which we feel empathy with the musician playing it (see also Schubert, 2007a), although that latter interpretation is rather counterintuitive from our point of view. Following Scherer and Zenter (2001), empathy requires sympathy, or as Preston and de Waal (2002) claim, moderators of empathy are familiarity, similarity, learning, past experience, and salience, all possibly related to preference. A third interpretation might be that other unknown underlying factors are determining the observed correlation between preference and empathy. However, it appears to be difficult to vary experimentally different degrees of preference or empathy, in order to investigate a direct causal relation between them. In a recent study, effects of empathy were tested by instructing participants to either empathize or inhibit empathy during music listening (Miu & Baltes, 2012). High empathy was shown to intensify emotional responses. However, there is still the possibility that participants in the low empathy condition inhibited not only empathy, but also emotional responses in general. Researching empathy in any way by asking participants to rate or respond with empathy might also lead to artificial responses or demand characteristics. One possible way to work around these conscious empathy manipulations and measurements could be to present additional personal information about musicians' emotional backgrounds while playing the music, and test for empathetic responses after emotion ratings have been made.

Furthermore, in light of the results presented here, it might be reasonable to include several other measures in future studies. The bipolarity of two-dimensional emotion rating scales can be questioned (Colombetti, 2005). Thus, future research might benefit from using independent assessments of positive and negative emotion. Focusing on general empathy with musicians, empathy was assessed by rating only one item here. Future studies could employ different rating scales differentiating empathy with performers, composers, or mere expressions in the music. Individual differences could be explained by including personality trait measures like the empathizing–systemizing scale (Kreutz et al., 2008). Here, emotional contagion was accessed by testing for direct, statistically nonmoderated, effects of emotional expressions on ratings of felt emotion. However, an alternative way would be to measure psychophysiological activations of music performers and observing listeners at the same time to identify moments of emotional contagion/empathy that would be indexed by concurrent activity (Jaimovich, Coghlan, & Knapp, 2010). Further interesting measures could be assessments of meta-appraisals (Schramm & Wirth, 2010) or reappraisals (Gross, 1998, 2002) to empathetic reactions with expressed emotions. This would give insights into the underlying evaluative processes during music listening that we only measured as preference in this study.

CONCLUSIONS

We interpret the results of this study as preliminary evidence for the model suggesting a link between recognized and induced emotions through empathy and emotional contagion in music listening. Replicating previous findings, ratings of recognized emotions did not equal those of felt emotion. Empathy was shown to affect the difference between the two emotion rating types and could be predicted by music preference ratings. By finally also modeling only felt emotion ratings, it was shown that preference, empathy, and expressed emotion (indicating automated emotional contagion) were significant predictors. Results also indicated that mixed emotions were induced in listeners, as felt arousal and valence ratings of excerpts with negative expressions were higher for preferred pieces than for non preferred ones (and vice versa). Taken together, these results might take us closer to understanding differences and commonalities between expressed and induced emotions in music, and at the same time inform us about empathy and emotional contagion in music listening.

Author Note

Part of this work was carried out at the Institute of Music Physiology and Musicians' Medicine, Hanover University of Music, Drama, and Media, Germany during the doctoral dissertation of the first author under the supervision of Prof. Dr. E. Altenmüller and Prof. Dr. R. Kopiez. We wish to thank both for their extraordinary support during that entire period. This work was partially funded by a German Research Foundation (DFG) grant (AL 269-6) to E. Altenmüller and R. Kopiez. The work in Canada was supported by a Social Sciences and Humanities Research Council grant (#410-2009-2201) and a Canada Research Chair awarded to Stephen McAdams. We would like to thank all participants and the team of graphic design students from the University of Applied Sciences and Arts Hannover for designing the web application used for the online music personality test.

Correspondence concerning this article should be addressed to Dr. Hauke Egermann, Fachgebiet Audiokommunikation, Technische Universität Berlin, Sekr. EN-8, Einsteinufer 17c, 10587 Berlin, Germany. E-mail: hauke@egermann.net

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes, (R package version 0.999375-39) [Computer software].

Besel, L. D. S., & Yuille, J. C. (2010). Individual differences in empathy: The role of facial expression recognition. *Personality and Individual Differences, 49*(2), 107-112.

Clynes, M. (1977). *Sentics - The touch of emotions*. New York: Doubleday.

Colombetti, G. (2005). Appraising valence. *Journal of Consciousness Studies, 12*(8), 103-126.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71-100.

Eerola, T., Lartillot, O., & Toiviainen, P. (2009, October). Prediction of multidimensional emotional ratings in music from audio using multivariate regression models. *Proceedings of the 10th International Conference on Music Information Retrieval*. Kobe, Japan.

Egermann, H. (2010). *Social influences on emotions experienced during music listening - An exploratory study using physiological and web-based psychological methods*. Berlin, Germany: Wissenschaftlicher Verlag Berlin.

Egermann, H., Grewe, O., Kopiez, R., & Altenmüller, E. (2009). Social feedback influences musically induced emotions. *The Neurosciences and Music III: Disorders and Plasticity: Annals of the New York Academy of Sciences, 1169*, 346-350.

Egermann, H., Kopiez, R., & Altenmüller, E. (2013). The influence of social normative and informational feedback on musically induced emotions in an online music listening setting. *Psychomusicology: Music, Mind, and Brain, 23, 21-32.*

Egermann, H., Nagel, F., Altenmüller, E., & Kopiez, R. (2009). Continuous measurement of musically-induced emotion: A web experiment. *International Journal of Internet Science, 4*(*1*), 4-20.

Egermann, H., Pearce, M., Wiggins, G. & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive Affective Behavioral Neurosciences, 13*, 533-553.

Egermann, H., Sutherland, M. E., Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2011). Does music listening in a social context alter experience? A physiological and psychological perspective on emotion. *Musicae Scientiae, 15*, 307-323.

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotion in music. *Musicae Scientiae, 12*(*1*), 75-99.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., et al. (2009). Universal recognition of three basic emotions in music. *Current Biology, 19*, 573-576.

Gabrielsson, A. (2001). Emotion perceived and emotion felt same or different? *Musicae Scientiae, Special Issue 2001-2002*, 123-147.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds), *Music and emotion: Theory and research* (pp. 223-248). New York: Oxford University Press.

Garrido, S., & Schubert, E. (2011). Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception, 28*, 279-296.

Gembris, H. (1982). Experimentelle Untersuchungen, Musik und Emotionen betreffend [Experimental investigations concerning music and emotion]. In K.-E. Behne (Ed.), *Musikpädagogische Forschung, Band 3, Gefühl als Erlebnis - Ausdruck als Sinn* (pp. 146-161). Laaber, Germany: Laaber.

Grewe, O., Kopiez, R., & Altenmüller, E. (2009). The chill parameter: Goose bumps and shivers as promising measures in emotion research. *Music Perception, 27*, 61-74.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271-299.

GROSS, J. J. (2002). Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology, 39*, 281-291.

HAUSEGGER, W. (1887). *Die Musik als Ausdruck* [Music as Expression]. Wien: Karl Konegen.

HONING, H., & LADINIG, O. (2008). The potential of the Internet for music perception research: A comment on lab-based versus Web-based studies. *Empirical Musicology Review, 3*(*1*), 4-7.

HERBERT, R. (2011). An empirical study of normative dissociation in musical and non-musical everyday life experiences. *Psychology of Music, 41, 372-394.*

HUNTER, P. G., SCHELLENBERG, G. E., & SCHIMMACK, U. (2008). Mixed affective responses to music with conflicting cues. *Cognition and Emotion, 22*(*2*), 1-26.

HUNTER, P. G., SCHELLENBERG, 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. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge, MA: MIT Press.

JAIMOVICH, J., COGHLAN, N. & KNAPP, R. B., (2010). Contagion of physiological correlates of emotion between performer and audience: An exploratory study. In J. Kim & P. Karjalainen (Eds.), *Proceedings of B-INTERFACE 2010, 1st International Workshop on Bio-inspired Human-Machine Interfaces and Healthcare Applications* (pp. 67-74). Valencia, Spain: INSTICC Press.

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

JUSLIN, P. N., LILJESTRÖM, S., VÄSTFJÄLL, D., BARRADAS, G., & SILVA, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion, 8*, 668-683.

JUSLIN, P. N., & VÄSTFJÄLL, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559-575; discussion 575-621.

JUSLIN, P. N., & VÄSTFJÄLL, D. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 223-253). Oxford, UK: Oxford University Press.

KALLINEN, K., & RAVAJA, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae, 10*(*2*), 191-213.

KREUTZ, G., SCHUBERT, E., & MITCHELL, L. A. (2008). Cognitive styles of music listening. *Music Perception, 26*, 57-73.

LIPPS, T. (1909). *Leitfaden der Psychologie* [Primer of psychology]. Leipzig, Germany: Engelmann.

LITLE, P., & ZUCKERMANN, M. (1986). Sensation seeking and music preferences. *Personality and Individual Differences, 7*, 575-578.

LIVINGSTONE, R. S., & THOMPSON, W. F. (2009). The emergence of music from the theory of mind. *Musicae Scientiae, Special Issue 2009-2010*, 83-115.

LIVINGSTONE, S. R., THOMPSON, W. F., & RUSSO, F. A. (2009). Facial expressions and emotional singing: A study of perception and production with motion capture and electromyography. *Music Perception, 26*, 475-488.

MEHRABIAN, A., YOUNG, A. L., & SATO, S. (1988). Emotional empathy and associated individual differences. *Current Psychology: Research and Reviews, 7*(*3*), 221-240.

MEYER, L. B. (1956). *Emotion and meaning in music.* Chicago, IL: University Of Chicago Press.

MIU, A. C., & BALTEŞ, F. R. (2012). Empathy manipulation impacts music-induced emotions: A psychophysiological study on opera. *PloS one, 7*, e30618.

MOLNAR-SZAKACS, I., & OVERY, K. (2006). Music and mirror neurons: From motion to 'e'motion. *Social Cognitive and Affective Neuroscience, 1*, 235-241.

NAGEL, F., KOPIEZ, R., GREWE, O., & ALTENMÜLLER, E. (2007). EMuJoy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods, 39*, 283-290.

NATER, U. M., ABBRUZZESE, E., KREBS, M., & EHLERT, U. (2006). Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology, 62*, 300-308.

OVERY, K., & MOLNAR-SZAKACS, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception, 26*, 489-504.

PANKSEPP, J., & BERNATZKY, G. (2002). Emotional sounds and the brain: The neuro-affective foundations of musical appreciation. *Behavioural Processes, 60*(*2*), 133-155.

PRESTON, S. D., & DE WAAL, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(*1*), 1-20; 20-71 (discussion).

REIPS, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology, 49*, 243-256.

RENTFROW, P. J., & GOSLING, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*, 1236-1256.

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

SCHERER, K. R. (1998). Emotionsprozesse im Medienkontext: Forschungsillustrationen und Zukunftsperspektiven [Emotion processes in the context of the media: Illustrative research and perspectives for the future]. *Medienpsychologie*, *10*, 276-293.

SCHERER, K. R. (1999). Appraisal theory. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 637-663). Chichester, UK: Wiley.

SCHERER, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33*(*3*), 239-251.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*, 695-729.

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 361-392). Oxford, UK: Oxford University Press.

Schramm, H., & Wirth, W. (2010). Exploring the paradox of sad-film enjoyment: The role of multiple appraisals and meta-appraisals. *Poetics, 38*, 319-335.

Schubert, E. (1996). Enjoyment of negative emotions in music: An associative network explanation. *Psychology of Music, 24*, 18-28.

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

Schubert, E. (2001). Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 393-414). Oxford, UK: Oxford University Press.

Schubert, E. (2007a). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music, 35*, 499-515.

Schubert, E. (2007b). Locus of emotion: The effect of task order and age on emotion perceived and emotion felt in response to music. *Journal of Music Therapy, 44*, 344-368.

Schubert, E. (2009). The fundamental function of music. *Musicae Scientiae*, *Special Issue 2009-2010*, 63-81.

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences, 1156*, 81-96.

Swanwick, K. (1973). Musical cognition and aesthetic response. *Psychology of Music, 1*, 7-13.

Swanwick, K. (1975). Can there be objectivity in listening to music? *Psychology of Music, 3*, 17-23.

Tremblay, A. (2011). *A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions.* R package version 1.6. [Computer software].

Vickhoff, B., & Malmgren, H. (2004). Why does music move us? *Philosophical Communications, Web Series, Department of Philosophy, Göteborg University, Sweden, 34*, 1-27. Retrieved October 7, 2013, from https://gupea.ub.gu.se/bitstream/2077/19475/1/gupea_2077_19475_1.pdf

Vuoskoski, J. K., & Eerola, T. (2011). The role of mood and personality in the perception of emotions represented by music. *Cortex*, *47*, 1099-1106.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: Chapman & Hall/CRC Press.

Zentner, M., & Eerola, T. (2010). Self-report measures and models of musical emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 185–222). Oxford, UK: Oxford University Press.

Appendix A

TABLE 1A. *Musical Stimuli*

No.	Performer	Composer	Title/Movement	Album	Lable/Source	Year	Style
8	4 Holterbuam	–	Jubiläums Jodler	Dila, dala. . . . (20 Jahre)	MCP Records	2005	German Folk
1	Africando All Stars	–	Doni Doni	Betece	Stern's Africa	2000	African Latin
4	At the Gates	–	World Of Lies	Slaughter oft the Soul	Earache Records	1995	Metal
6	BBC Scottish Symphony Orchestra	Edward Grieg	Peer Gynt-Suite I: Morning mood	Das ABC der Klassischen Musik	Naxos	1995	Romantic
5	Circle Takes the Square	–	Same Shade as Concrete	As the Roots Undo	Robotic Empire	2004	Emo rock
24	Danubius Ensemble	Antonio Vivaldi	Trio Sonata B-Major, op. 5,5, Corrente: Allegro	Das ABC der Klassischen Musik	Naxos	1999	Baroque
22	Erick Truffaz/ Nya	–	Siegfred	Bending New Corners	Kameleon Music	1994	Jazz Fusion
3	Horváth István	–	Nem kell nékem pogácsa	Mulatok, Mert Jó Kedvem Van	Hungaroton Classic Ltd.	1995	Hunga-rian Folk
14	Jeno Jando	Wolfgang A. Mozart	Piano Concerto. 21 C-Major, KV 467	Das ABC der Klassischen Musik	Naxos	1988	Classical
7	John Coltrane	Duke Ellington	In a Sentimental Mood	Duke Ellington and John Coltrane	Impulse!	1988	Jazz
10	Kronos Quartet	Philip Glass	String Quartet No. 4 (Buczak), 2. Movement	Kronos Quartet Performs Philip Glass	Nonesuch	1990	Minimal music
11	Kronos Quartet	Terry Riley	Salome Dances For Peace: Half-Wolf Dances Mad In Moonlight	Winter was Hard	Nonesuch	2004	Minimal music
9	Kronos Quartet	John Zorn	Forbidden Fruit	Winter was Hard	Nonesuch	2004	Avant-garde
12	Kronos Quartet	Aulis Sallinen	Winter Was Hard, Op. 20	Winter was Hard	Nonesuch	2000	Modern tonal
13	Muungano National Choir	–	Vanga Yohana	Missa Juba	Philips	2001	African Gospel
20	O. Markovic	–	Ti ne znas sta je ljubav	Starogradski biseri 1	Hi-Fi Centar	2002	Serbian Folk
18	Sigur Rós	–	Olsen Olsen	Agaetis byrjun	Fat Cat Records	2004	Indie Pop
16	Sigur Rós	–	Staralfur	Agaetis byrjun	Fat Cat Records	1996	Indie Pop
17	Sigur Rós	–	Agaetis byrjun	Agaetis byrjun	Fat Cat Records	1996	Indie Pop
19	Standstill	–	Un Gran Final	First Album	Defiance Records	1994	Indie Rock
15	Stefan Mross	–	Il silenzio	Von Herzen Alles Gute	Montana	1994	German folksy
2	The London Virtuosi	Tommaso Albinoni	Oboe Concerto C-Major, op. 9,5, Allegro	Das ABC der Klassischen Musik	Naxos	1962	Baroque
23	Village of Savoonga	Wolfgang Petters	My Mind - Your Mind	Philipp Schatz	Kollaps/ Haus-musik	1994	Post-Rock