Assessing Emotional Experiences of Opera Spectators in Situ

Klaus R. Scherer\textsuperscript{1,2}, Stéphanie Trznadel\textsuperscript{1}, Bernardino Fantini\textsuperscript{1}, Eduardo Coutinho\textsuperscript{3}

\textsuperscript{1}Swiss Center for Affective Sciences, University of Geneva
\textsuperscript{2}University of Munich
\textsuperscript{3}University of Liverpool

Corresponding address: Klaus R. Scherer, Department of Psychology, University of Geneva, Boulevard du Pont-d'Arve, 40, CH-1211 Geneva, Switzerland. E-mail: Klaus.Scherer@unige.ch.
Acknowledgements

The series of studies reported here were conducted by members of the Music and Emotion Focus of the Swiss Center for Affective Sciences (Klaus Scherer, Bernardino Fantini, Eduardo Coutinho, and their collaborators) in collaboration with the Geneva Opera House (Daniel Dollé and his collaborators). The research was funded by an ERC Advanced Grant in the European Community's 7th Framework Programme under grant agreement 230331-PROPEREMO (Production and perception of emotion: an affective sciences approach) to Klaus Scherer and by the National Center of Competence in Research (NCCR) Affective Sciences financed by the Swiss National Science Foundation (51NF40-104897) and hosted by the University of Geneva.

K.S. conceived, financed, and supervised this research. He designed the studies, analyzed large parts of the data, and wrote the final manuscript. S.T. performed much of the data analysis and wrote the first draft of the manuscript. B.F. contributed to the design and the realization of the field studies and made major contributions to the manuscript. E.C. co-developed the GEMIAC rating scale and analyzed parts of the data. K.S. and S.T. share first authorship, names appearing in alphabetic order.
Abstract

Opera performances elicit strong emotional reactions in listeners. Yet, empirical demonstrations of these effects in situ are rare. Here we report a series of studies examining the emotional reactions of participants invited to the dress rehearsal of three different operas at the Geneva opera house before large audiences. Using a new affect checklist developed specifically for in situ studies of music performances, we asked participants to record 1) the intensity of 12 different types of affective reactions they experienced during selected scenes or 2) the frequency with which they experienced these emotions during specific acts or the opera as a whole. Results showed a high degree of specificity regarding the emotional impact of the operas as a whole and of individual scenes/acts. For one opera, Verdi’s Macbeth, we also asked participants to rate both the emotions they actually felt during the opera scenes and the emotions expressed by the music or the singers’ interpretation. Results confirm that spectators are clearly able to separate their own affective responses from what they perceive to be the emotions portrayed by the orchestra or on the stage. In addition, we evaluated the effect of different types of pre-performance information sessions (on plot or music) as well as of participant personality and prior mood. Overall, the results demonstrate the feasibility of measuring highly differentiated emotional audience reactions to an opera performance with a brief validated checklist during actual performances in the opera house and the validity of self-reported emotions.

Keywords: music and emotion, opera performances, affective reactions, felt and perceived emotions, mood and personality effects
Introduction

In recent years, the capacity of music to elicit emotions in the listener has become a major topic for scientific inquiry, both theoretical and experimental (see contributions in Bhatara, Laukka, & Levitin, 2014; Cochrane, Fantini, & Scherer, 2013, and Juslin & Sloboda, 2011). Much of this work has focused on instrumental music, probably partly due to a philosophical tradition that holds that the essential question is to what extent “absolute music” (i.e., instrumental music without any programmatic content) can elicit emotional reactions (Kivy, 1990; Konečni, 2008). As amply documented in the literature, theorists and researchers disagree with this claim and consider music as a performance art, of which there are many different genres, all of which are able to induce affective states. In fact, singing and especially opera have been traditionally considered as the most emotional types of music (Cohen, 2016).

Specifically, operatic performances consist of a dramatic plot as developed in the libretto, the singing and acting of the singers, the music played by the orchestra, and the (often elaborate) scenery and costumes, all of which are likely to involve many of the mechanisms by which emotions are induced in listeners. In opera, the vocal characters and the orchestral music that supports them interact with many other elements such as literature, dance, visual art, scenic movements, and lighting to create a dramatic context, which are all conducive to the production of strong emotions in the spectator. The causes of the emotional effects in an opera performance are therefore multifaceted, and it is important to examine the extent to which each of these major factors influence emotion elicitation. Scherer, Zentner, & Schacht (2001/2002) asked a large group of participants in a music science conference to rate the importance of the following
determinants for their affective reaction to music: musical structure, acoustic features, interpretation, technical quality, listener’s mood, affective involvement, context factors, or listener personality. While musical structure was given the highest rating of the list of determinants followed by technical, acoustic, and interpretational features, listener and situational context factors also received high ratings.

Scherer and Zentner (2001) called the underlying mechanisms for these factors *routes* and provided a number of examples: (1) specific types of appraisal (such as novelty, unexpectedness, pleasantness), (2) music-related memory associations, (3) contagion, and empathy, (4) entrainment and proprioceptive feedback, and (5) facilitation of pre-existing emotions (disinhibition). They suggested a *route model* to highlight the major pathways through the cognitive and emotional systems involved, with each of these recruiting a large number of neural and somatic structures and many different mechanisms in terms of the underlying machinery. Juslin and Västfjäll (2008) have suggested a highly similar list of mechanisms in the form of a theoretical framework called BRECSEM (brainstem reflex, rhythmic entrainment, evaluative conditioning, contagion, visual imagery, episodic memory, and musical expectancy). Scherer and Coutinho (2013) have pointed out that three of the BRECSEM factors can be subsumed under the "appraisal" route: brainstem reflex as a low-level novelty and intrinsic pleasantness check; evaluative conditioning as the precondition for creating valenced associations; and musical expectancy as part of the discrepancy of expectation check. Episodic memory is covered under the memory mechanism, as is visual imagery in the form of memory-dependent constructive imagination. Empathy and emotional contagion cover the BRECSEM contagion mechanism. Rhythmic entrainment corresponds to proprioceptive feedback, which also contains emotion facilitation.
Specifically, the route model assumes that these mechanisms operate via complex, multiple, nonlinear interactions between different determinants that can be broken down into four classes, to elicit emotions in the listener: Structural features (segmental and suprasegmental aspects of music), Performance features (Performer skills, intentions, moods), Listener features (musical expertise, stable dispositions, current motivational/mood state), and Contextual features (location, event).

This model was originally intended for instrumental music and needs to be extended by additional factors to cover the essential elements of opera as described above. In particular, the Structural features need to be complemented by three important factors: Plot, Lyrics, and Staging. Particular attention also has to be given to the vocal dimension, as the singing voice is one important specificity that characterizes opera in relation to other performing arts. Given the difficulty of measuring and experimentally manipulating these factors in ecologically valid settings, such as concert halls and opera houses, these theoretical suggestions have rarely been investigated empirically (with some exceptions: e.g., for opera, Balteș & Miu, 2014; Vaitl, Vehrs, & Sternagel, 1993; see also http://www.transforming-musicology.org/news/2014-11-13_hearing-wagner-at-being-human/; for Lieder, Coutinho & Scherer, 2016; for electronic music, McAdams, Vines, Vieillard, Smith, & Reynolds, 2004). One of the central questions with which this research is confronted, is the validity of self-reported emotions induced by listening to music: Can listeners really report a highly differentiated spectrum of music-elicited emotions or do they mostly report what they perceive or believe the music is designed (and/or the interpreter intends) to express? This issue is particularly salient for vocal music, since the dramatic plot or the text generally also express emotional content -- often much more explicitly than "pure music".
The aim of this article is threefold:

1) Examine the feasibility of measuring felt emotional reactions to opera performances in situ, immediately after the performance of selected scenes or acts.

2) Evaluate the difference between perceived affect (as shown by the singers on stage) and actually felt affect (by the audience). It is often suggested in the literature that self-reported emotional responses are affected by the perceived “emotional character” of the music. Earlier studies investigating the relationship between felt and perceived emotions for instrumental music suggest that although the emotions experienced by listeners may sometimes reflect those they perceive in a piece of music, there are major differences between the two, mainly in the emotional intensity and valence (Evans & Schubert, 2008; Hunter, Schellenberg, & Schimmack, 2010; Kallinen & Ravaja, 2006; Konečni, Brown, & Wanic, 2008; Zentner et al., 2008, Study 2). In particular, very often the feeling evoked by a sad piece of music is rated as highly pleasant by the listener, a paradox which has found a lot of attention in the recent literature (Kawakami, Furukawa, Katahira, & Okanoya, 2013; Taruffi & Koelsch, 2014; Vuoskoski & Eerola, 2017). However, it is possible that the emotions acted out on an opera stage in both singing and acting has a stronger impact and is more likely to be directly mirrored in the spectator’s reported feelings.

3) Investigate the possibility of empirically assessing the role of some of the features outlined above in an operatic setting, in particular the effect of Listener features, as described in the routes model, namely a) prior information/knowledge, b) personality, and c) mood state. As to 1), many opera houses offer their public
introductory talks by experts before the performance, generally assuming that listening
to these talks will affect the nature of the audience reaction, resulting in greater
appreciation. Also, an extensive literature on the effect of musical knowledge, training,
or capacity on music appreciation (see, for example, Bigand, & Poulin-Charronnat,
2006) suggests that these factors may also influence the listener’s emotional reaction.
Here, we experimentally varied part of the expertise/prior knowledge feature by
providing an introduction either to the music or to the plot and the staging to subgroups
of the participants. The important role of listener background factors like personality
disposition and preferences, as well as mood before the performance on emotional
reactions had been empirically confirmed in the study by Scherer et al. (2001/2002) and
there is evidence for such effects in the literature (see, for example, Liljeström, Juslin, &
Västfjäll, 2013; Hunter, Schellenberg, & Griffith, 2011). In order to examine such
effects in an operatic context both personality and mood before and after the
performance were assessed by questionnaire.

We describe a series of studies conducted during dress rehearsals of different
operas in a major opera house to study the nature and intensity of the emotions invoked
in groups of participants in the audience. We chose highly different works from the
operatic repertoire to represent different types of music and plots with different
characters and actions. In addition, we obtained separate data for different acts of an
opera to allow comparison between different plot structures.

In all cases, the major dependent variable—emotions elicited by the operatic
performance—was measured by the beta version of a new emotion assessment
an extended, fuzzy-set version of the Geneva Emotional Music Scale (GEMS; Zentner et al., 2008): the GEneva Music-Induced Affect Checklist (GEMIAC), which was developed specifically for use in field studies, including performances in concert halls and opera houses (Coutinho & Scherer, 2017). The development of the checklist was based on an empirical investigation of the semantic structure of the relevant terms through a series of hierarchical cluster analyses. The beta version of the GEMIAC (abbreviated βGEMIAC here) differs only slightly from the final published version. We used this short, 12-item, fuzzy-set checklist, specifically developed for in situ studies, rather than the GEMS, because there would not have been enough time between acts without an intermission to rate the large number of items. We used the same checklist to obtain ratings of both felt and perceived emotions.

Obviously, the naturalistic setting of the studies, a public opera performance in an opera house, did not allow for systematic separation of the different factors outlined above. Thus, there is always a specific combination of structural factors—plot, lyrics, segmental and suprasegmental aspects of music, staging, performer skill, and expressive intention—the respective effects of which cannot be disentangled. However, given the present scarcity of empirical research in an ecologically valid music setting, it seems useful to examine the following questions:

1. Can we identify specific “feeling profiles” for the affective reactions to different operas and for specific acts or scenes of an opera?
2. Is it possible to differentiate the emotion elicited in the spectator by an operatic performance and the emotion the spectator perceives as being expressed by the performers?
3. How does concrete prior information about certain aspects of operatic performances (music vs. plot and staging) affect the spectators' emotional responses?

4. What are the effects of individual differences between spectators including personality and mood state on emotion elicitation?

Although one might have hunches on what to expect regarding some of these questions from plausibility and past research, we resisted the temptation to formulate specific hypotheses or even expectations, given the nascent state of theorizing and empirical research in the area of emotional effects of operatic performances. We hope that the findings of this series of studies can establish a more solid basis for hypothesis-guided future research.

**Method**

**General Procedure**

We invited groups of participants via advertisements placed in various public spaces of the University of Geneva to attend the general dress rehearsal of selected operas presented at the Grand Théâtre de Genève. We specified that prior experience with opera and intrinsic interest were required because we wanted participants willing to devote their full attention to the operas and the rating task rather than be motivated only by the payment. In addition, to be able to generalize results to opera-attending audiences, participants needed to have a minimum of understanding of the nature of the operatic event under study. Many, but not all, participants were students in different fields.
Choosing from the seasonal offerings of the Geneva Opera House, we selected operas from different historical periods: a romantic opera (Giuseppe Verdi’s *Macbeth*), a 20th-century opera (Richard Strauss’ *Der Rosenkavalier* - The Knight of the Rose or The Rose-Bearer), and a contemporary opera (Giorgio Battistelli’s *Richard III*). In a preliminary study, we used a performance of Richard Strauss’ *Elektra* (not reported here). The musical and dramatic textures of these operas are highly divergent and were thus expected to express and induce different emotions.

To manipulate expertise/prior knowledge (at least to some extent) we provided prior information about the operas by dividing the groups of participants into subgroups. Each of these received a 30 minute oral introduction to either the plot of the opera and the staging by the director (staging group), or the musical structure as written by the composer (music group). In the case of two operas, there was also a control group that did not receive any prior information. Participants were able to ask questions.

In all cases, participants attended the dress rehearsal, that is, the final performance before opening night, without any interruptions or interventions, in front a full house of invited spectators. In each case, the participant group sat in a special section of the first balcony, apart from the general public. They were asked to perform their ratings as soon as possible after the end of the selected scene or after the curtain fell after an act. There was sufficient time for the ratings to be completed. Participants were asked not to interact with each other during the performance and the rating periods. Two members of our group made sure that such interactions did not occur.

**Questionnaires**
Emotional experience felt during the performance. We assessed the participants’ emotional experience during the operas with the βGEMIAC. This emotional experience questionnaire comprises 12 fuzzy emotion categories with two to three items each (see Supplemental Material Online (SOM) Section A for a list of the categories and summary labels as well as an example of the original questionnaire). The same questionnaire was used during the selected parts and at the end of the opera, the only difference being that the participants were asked to rate the intensity of the emotions felt during brief scenes (on a 5-point scale ranging from 1 = not at all to 5 = very much) and the frequency at which they felt these emotions during acts or the whole opera (on a 5-point scale ranging from 1 = not at all to 5 = frequently). The reason for using these different formats is to get at different types of information - the ratings about the feelings of emotion experienced during a specific scene are measured in terms of relative intensity of the reaction and are made as close in time as possible to the occurrence of the reaction (in order to be clearly remembered and not superseded by new impressions and reactions). This is why, in the case of operas in which we focused on individual scenes (Macbeth, Richard III) we asked participants to provide intensity ratings as soon as possible after the chosen scenes. In some cases we wanted to get an overall, summative evaluation about the prominent feelings experienced throughout different acts (Rosenkavalier) or the opera as whole (Richard III). It is unreasonable to assume that participants can recall the precise intensity of the many different affective experiences during the 2-4 hours of watching the operas. In consequence, we asked them to rate the frequency with which they had experienced specific emotions during the act or whole opera after the final curtain fell, giving them sufficient time to recall their affective experiences during the evening. Both procedures are commonly used in music research, including studies of symphonies or
Lieder recitals. While the two ratings formats are not directly comparable, they both give an indication of the relative importance or preponderance of different types of affective experiences.

We also investigated the influence of various individual difference and context factors (routes) on participants’ emotional experience as follows:


**Mood before and after the opera.** Participants were asked to rate their current mood before and after the opera on six 5-point, bipolar scales: Bad-good, Sad-happy, Dissatisfied-content, Calm-excited, Tired-energetic, Lethargic-awake (adapted from the 6-item mood scale validated by Wilhelm & Schöbi, 2007).

**Description of Analyses and Results**

The above description summarizes the general approach across all studies. As the issues examined and methodological procedures used varied slightly across studies, we describe the specific purpose and the specific procedures in a brief paragraph at the beginning of the report of each study. In line with the questions enumerated at the end of the Introduction, analyses and results are described in four separate parts: 1) within-opera analyses (participants’ judgments of the intensity of emotional reactions for each act/scene and frequency of specific during the opera as a whole), and 2) between-opera analyses (comparison of emotion profiles across operas). 3) Information effect: the effect of the manipulation of the introductory information session (plot/staging vs
music), and 4) Individual differences: the effects of individual differences (age, sex,
personality, mood) on the emotional reactions reported.

Statistical Analyses

Data were transferred from the paper questionnaires used on-site to IBM SPSS
(version 22) for analyses. To control for response style differences between participants,
we standardized the raw data (z-transforms) for each person. Even though ratings on the
intensity of specific reactions to scenes/acts and the overall frequency of these reactions
in the opera as a whole are not directly comparable, the use of standardized scores (with
a mean of 0 and standard deviation of 1) allows a reasonable estimate of the relative
importance of specific emotions in the reactions to scenes/acts versus the opera as a
whole. To correct significance levels for multiple comparisons (given that the 12 fuzzy
items are individually tested) we used the Benjamini-Hochberg (BH) procedure
(Benjamini and Hochberg, 1995), setting the critical value for a false discovery rate
(FDR) to 0.10. Only effects for which the p values were inferior to the critical level of
the BH procedure are reported as significant.

Results

Within-Opera Analyses

In this section, we describe the emotional experience profile for each opera and
its selected parts from the ratings on the 12 fuzzy categories of the emotional experience
questionnaire. The central purpose of this approach was to investigate the extent to
which different parts of an opera generate unique emotional profiles (compared to the
integrated overall experience at the end of the performance) and the capacity of the new
GEMIAC affect checklist to capture such differences. Repeated-measures analyses of variance (ANOVAs) were performed for each of the data sets, according to the variables measured.

**Study 1 - Macbeth (Verdi)**

Verdi's operas are strongly marked by his musical characterization of the protagonists. The composer’s lyrical inspiration and inventiveness express and represent different feelings by using specific musical figures. In Verdi’s operas, there is a clear architectural construction of the whole drama, which creates the global environment in which the protagonists live, act, and feel. In the case of *Macbeth*, the structure of the plot is perfectly symmetrical, with two parts each composed of two acts and beginning with a witchcraft scene. Each part has a specific emotional atmosphere: the first part is characterized by love, sexual impulses, pride, and desire of power, and the second is dominated by shame, guilt, sorrow, and despair. This creates a powerful contrast between the stability of the general symmetry and the melodic fragmentation of the music and singing. In the different contexts, the individuality of the characters is represented by specific melodic lines and figures, especially by the expressive strength of the voices (the orchestra instruments often simulate human vocal utterances). These figures strictly adhere to the psychological reality of the protagonists, whose emotions are highly variable and specific to each dramatic moment.

We investigated two issues in the case of this opera: (1) potential differences in emotional response to different scenes of the opera (in each of three acts), and (2) potential differences between the participants’ *perception or inference* of the emotions of the characters as intended by the composer and interpreted by the protagonists (as expressed in the music/singing) versus the participants' *perception of their own*
subjective experience while listening (a factor in the analysis we will refer to as Modality).

Procedure. Thirty-nine participants (72% female, mean age 27 ± 9.5 years) were recruited via advertisements at the University of Geneva. Participants were invited to come to a performance of Macbeth and were asked to complete the βGEMIAC during three particular scenes (from Act I scene 7, starting with the words: “Or tutti sorgete, ministri infernali”; Act II scene 5, starting with the words: “Si colmi il calice di vino”; and Act IV scene 4, starting with the words: “Una macchia è qui tuttora!”). For each scene, participants were first asked to rate, on a scale from 1 to 5, the intensity of emotions they perceived on the stage, that is, the emotions they thought the singer was trying to convey, and then the intensity of emotions they themselves actually felt while watching and listening. We did not randomize the order of the two questions given that the expression-impression sequence seems more natural and because randomization may produce artifacts unless there is a large N and a large number of trials. Participants were asked to record their ratings immediately after the end of the scene.

Results. The mean ratings of the intensity of both perceived and felt emotions on the 12 βGEMIAC items for the three scenes in Act I, Act II, and Act IV are shown in Figure 1. A repeated-measures ANOVA was carried out on the z-transformed data for the felt emotions only to determine differences between the three scenes. The results are shown in Table 1 (tests violating the sphericity assumption had the degrees of freedom
corrected as indicated in the table) together a comparison of significant differences between the means (multiple comparisons corrected using the BH criterion). As suggested by the patterns documented in Table 1 and Figure 1, there are significant differences (with strong effect sizes) for the felt emotions of Sadness, Power, Animation, Nostalgia, Interest, Enchantment, and Tenseness between the three focal scenes in Acts I, II, and IV.

Scene 7/Act I generates stronger feelings of Power, Animation, and Interest than do the scenes in Acts II and IV. In direct contrast, the focal scenes in Acts II (Scene 5) and IV (Scene 4) produce stronger feelings of the opposite emotion classes of Sadness and Nostalgia. This is easily explained by the great differences in the emotional content of the scenes and the corresponding music composed by Verdi to illustrate these affects. Scene 7 in Act I reveals Lady Macbeth’s quest for power and royal status when she urges Macbeth to kill the king during his visit to the castle. In contrast, the focal scenes in Acts II and IV are both centered on emotional crises in the face of impending doom. In Scene 5/Act II, Lady Macbeth, adopting a dominant stance, sings a sprightly drinking song but Macbeth’s developing madness is becoming noticeable. Finally, in Scene 4/Act IV, the Lady also realizes the full extent of evil she has brought about and the impossibility of undoing the consequences.

The scenes in Acts I and IV produce stronger feelings of Enchantment. This can be in large part explained by the large number of unusual and mysterious events occurring in these acts. The opera starts in Act 1 with the powerful witch scene with the enigmatic prophecy and the whole first act is dominated by mystery, both dramatically and musically. Act IV is equally full of mysterious puzzling events, from the moving forest to the sleep walking of Lady Macbeth.
Finally, the scenes in Acts I and II elicit stronger Tenseness than Scene 4 in Act IV, most likely due to the suspense involved in the unfolding action and the aggressive actions being planned and committed. In contrast, in Act IV, dominated by Lady Macbeth’s desolation and resignation, the dire end is easily predictable.

To understand the relationship between perceived and felt emotions, we computed a repeated-measures ANOVA with both the Scene/Act and Modality factors (perceived vs. felt) as within-subject factors. The results (with appropriate adjustments for lack of sphericity) are shown in Table 2. The main effects for Scene/Act (Table 2a) are not discussed here, as the mean of the perceived and felt feelings is not of interest. Strong main effects for Modality (Table 2b) are found for Power, Animation, Sadness, and Nostalgia. As discernible from the patterns in Figure 1, more Power and Animation is attributed to the actions and expressions of the singers, whereas more Sadness and Nostalgia is felt by the participants. It is interesting that these are exactly the emotion classes that showed strong main effects for scene/act differences. The interaction effects yielded by the ANOVA (Table 2c) help explain this result. Again we find highly significant effects with strong effect sizes for Power, Animation, Sadness, and Nostalgia, as well as for Enchantment and Tenseness. In Figure 2, we plot the estimated marginal means for these interactions for ease of interpretation. The graphs for Power and Sadness (Figure 2, left side) clearly show that the difference between perceived and felt emotions occurs in Act II. The protagonists on the stage still express Animation and...
Power (in particular Lady Macbeth’s toast is full of energy and strength), but the spectator already sees what is coming and reacts with an increase in sadness.

The graphs for Enchantment and Tenseness (Figure 2, right side) show that it is again in Act II where we find the greatest discrepancy between perceived and felt emotions, most likely because the spectators can now predict how things will turn out in the end. In consequence, while the protagonists, especially an apparently relaxed Lady Macbeth in the enchanting drinking song (*brindisi*), remain in character, the spectators feel little enchantment and much tenseness in view of the dire forebodings. In scene 4 in Act IV participants perceive Lady Macbeth's sadness and tenseness but feel enchantment listening to this famous aria.

These results illustrate the potential role of multiple mechanisms interacting in the elicitation of emotions by different types of music, as suggested by Scherer and Zentner's (2001) route model, in particular specific types of appraisal, contagion, empathy, and entrainment. However, although the similarity between perceived and felt emotions in Acts I and IV might well be linked to empathy and contagion, or even entrainment, the wide divergence between perceived and felt emotions in Act II is most likely due to complex appraisals of the imagined future events and their consequence (see Discussion).

**Study 2 - Richard III (Battistelli)**

The opera, or “dramma per musica”, as the composer prefers to call it in the spirit of Monteverdi’s operas, describes Richard’s rise to power and throne. His ascent to kingdom is made possible by an extraordinarily large number of political assassinations, including those of his young nephews in the Tower of London. The
libretto is directly taken from Shakespeare’s drama, using the original verses, but many parts were cut to center the opera on the character of Richard III and his maniacal ambition, which pushes him to arrange the serial murders.

Musically, the opera proceeds in an unbroken flow of vocal and instrumental fragmented utterances, with forceful repeated figures and a few recurring themes, such as the choral tribute to the newly crowned Richard at the end of Act 1. Wordless choral singing accompanies the recount of the murder of the king’s nephews and the scene in which Richmond (Henry VII), surrounded by corpses, proclaims in a speaking voice the end of the War of the Roses.

Battistelli’s musical language is extremely dense and full of tension. The orchestra, with a reinforced percussion section, deploys dark and incisive tones. The vocal writing rejects every formal style, putting all the potentiality of the voice, from the speaking declamation to the lyric expression, at the service of the drama, to highlight the different emotional aspects of the opera.

In this opera we examined potential differences in emotional response to a key scene and to the opera as a whole. The specific scene chosen (Act II, scene 3) contains the malediction of Richard by his old mother because of the murders he committed.

Procedure. Forty-eight participants (73% female, mean age 26 ± 6 years) were recruited via advertisements at the University of Geneva. During a performance of Richard III, participants were given the βGEMIAC to complete during a specific scene and after the whole opera. For the specific scene, participants were asked to rate the intensity of 12 emotions they might have felt during the scene. For the whole opera, they were asked to rate, after the final curtain, the frequency at which they felt those 12 emotions throughout the opera.
Results. The βGEMIAC profile means for the z-transformed data to compare the ratings for scene (intensity) vs opera as a whole (frequency) are shown in Table S1 in the SOM Section B and displayed in Figure 3.

The Part-Opera differences cannot be analyzed with standard statistical procedures for the significance of differences because the scene ratings were based on intensity and the whole opera rating on the frequency of emotions. In an exploratory vein, we used a profile comparison method (Deshpande, Van der Sluis, and Myers, 2013) to obtain some quantitative information of the differences. The results, shown in Table S1 in the SOM Section B, suggest the strongest profile differences for Awe, Harmony, Tenderness, Boredom. In all of these cases, the difference in the emotional value between the scene and the opera is due to higher scores for the felt emotions in the whole opera compared to those in the specific scene. Very high similarity is found for Tenseness, Interest, Animation and Sadness, possibly due to the frightening nature of the violent malediction scene that has, to some extent, affected the ratings of the whole opera. As could be expected from the representation of a contemporary opera with its innovative musical language, epistemic, knowledge-related emotions (astonishment, curiosity, interest, wonder, surprise, doubt; see Scherer, Coutinho, 2013, p. 125) prevail over other types of emotions, the dominant emotion categories in this opera being Interest, Animation and the absence of Boredom.

Study 3 - Der Rosenkavalier (Strauss)
Hofmannsthal’s libretto and Strauss’ score are dominated by the Marschallin’s reflections and emotional feelings about the flow of time. The different scenes of the opera create a sort of emotional crossover, with a crescendo of positive emotions (tenderness, love, trust, joy) and a parallel decrescendo of negative emotions: The Marschallin’s attitudes and emotions shift from initial melancholy and sadness, and even anger, to the final acceptance in Act III that the passage of time cannot be stopped.

In Act I, the Marschallin acknowledges with fear, sadness, and anger that time relentlessly flows because each instant of the present passes irreversibly into the past. For a long time, her love for the much younger Octavian had been an effort to affirm the persistence of a younger emotional self inside an aging body. At the beginning of the opera, with nostalgia and despair, she finally recognizes that this will no longer be possible and she laments clocks and mirrors, which remind her that time is passing away. At the end of the opera, she becomes conscious that life is a process, in which one continuously takes, holds, and leaves pleasures and emotional experiences. The heart of Der Rosenkavalier is the touching final trio sung by the Marschallin, Sophie, and Octavian, during which the Marschallin leaves the scene to the ecstatic singing of the young lovers, who can now look to a bright future together. But the Marschallin also no longer looks to the past and to the flow of time, but starts to look to her own future.

The question addressed in studying this performance was the difference in emotional response to the three acts of the opera.

Procedure. Eighteen participants (80% female, mean age 27 ± 7 years) were recruited via advertisements at the University of Geneva. Participants were invited to attend a performance of Der Rosenkavalier and were given the βGEMIAC questionnaire to report on their emotional experiences. In this case, no specific scenes
were chosen but participants were asked, after each of the three acts, to rate the
frequency with which they had experienced the 12 emotion classes. The frequency
rating format was chosen because of the length of the acts and the variety of different
scenes which would have made it difficult to remember the respective intensity levels of
affective reactions.

Results. The βGEMIAC profiles for the three acts are displayed in Figure 4 (for
a listing of the main effects revealed by the repeated-measures ANOVA and the mean
values, see Table S2 in SOM Section B).

There was a significant main effect of Part for Animation, $F(1.85, 17) = 6.24, p$
= .006, $\eta^2 = .27$, and a tendency (given the BH criterion) for Interest, $F(1.67, 17) =$
3.87, $p = .039, \eta^2 = .19$, reflecting the higher values attributed to these emotions in Act
II compared to Acts I and III (corroborated by significant quadratic trends in the
ANOVA), a result that may be due to the lively and animated scene around the
confrontation between Octavian and Baron Ochs and his footmen in Act II. The means
also suggest a continuously rising feeling of tenderness of the participants from Act I
and continuing across Act II to Act III (supported statistically by a significant linear
contrast, $F(1, 17) = 6.03, p = .025, \eta^2 = .26$). This may be due to the affectionate duet
between Sophie and Octavian in Act II and the compassionate finale, the Marschallin
gracefully renouncing her lover, encouraging him to obey his heart, upon which the two
young people embrace each other in a soaring duet of love.
Between-Opera Analyses

It is difficult to compare the differential emotional impact across the three operas as a whole, given the differences between scenes and acts within each opera as described above. However, given the major differences between the operas with respect to the period of their creation and their general style, it does seem of interest to attempt at least an approximate comparison in an exploratory fashion. As the final act of an opera is often apotheosis of the work and as spectators' ratings of the last period in the opera may partly reflect the special character of the opera as a whole, we propose to compare the ratings for the last act across the three operas (whole opera for Richard III). It should be noted that, due to the specific questions asked in each of the studies, intensity ratings were used in Macbeth and frequency ratings in Der Rosenkavalier and Richard III. Figure 5 illustrates the differential patterns in the emotional reactions to the three operas. Der Rosenkavalier obtained significantly higher ratings on Harmony and Tenderness, Macbeth received higher ratings on Sadness and Nostalgia, and Richard III higher ratings on Tenseness and Interest. Only Enchantment and Awe are rated in a highly similar fashion for all three operas. These patterns cannot be tested for significance with standard statistical procedures given the difference between frequency and intensity ratings. However, given the similarity of results obtained with frequency and intensity scales in the literature (Krabbe and Forkman, 2012) and the fact that the z-scores in each case indicate the relative predominance of certain emotions in a metrically comparable fashion, we used nonparametric analyses for exploratory comparison. The results for the ratings of the last act in the three operas are shown in Table S3 in SOM Section B. There are important differences for all emotions except Enchantment and Awe, confirming the excellent resolution provided by the GEMIAC scale. In particular, they show that while
established music evaluation dimensions like enchantment and awe do not well
discriminate (as all opera produce these feelings to some extent), the new scales added in
the GEMIAC provide clear differences for the three operas. The standardized profiles
shown in Figure 5 confirm the patterns found in the detailed analyses of the different
operas and the respective acts, providing typical profiles for the overall emotion induction
potential of the different works under study.

The remainder of the article deals with questions 3 and 4, listed at the end of the
introduction. Given the large amount of data to be analyzed in order to respond to these
questions, we decided to perform the analyses with component scores following a
Principal Component Analysis (PCA) of the complete set of the ratings on the 12 fuzzy
emotion categories to facilitate comparisons and to increase the stability of the effects.
The results of the PCA (Eigenvalue > 1; VARIMAX rotation) is shown in Table 3. Five
factors are extracted according to criterion, explaining 68% of the total variance. The
rotated loadings for the five components are readily interpretable and can be labelled as
follows: 1 Aesthetic_Enjoyment (feelings of harmony and clarity being generally
associated with beauty and the absence of sadness and tenseness pointing toward calm
enjoyment), 2 Intellectual_Challenge (animation and excitement being generated by
interest and discovery as well as wonder), 3 Melancholy (feelings of nostalgia and
sadness combined with powerlessness), 4 Awe (feelings of veneration and
transcendence combined with enchantment and wonder), and 5 Tender_Involvement (an
active, outgoing form of tenderness in the absence of bored indifference). These five
component scores will be used as measures of the emotional reactions of the
participants to the opera performances in all analyses reported in the remainder of the article.

---Table 3 here---

**Information Effects.** This section focuses on the third question for this series of studies mentioned in the introduction, the potential role of receiving prior information about the music or the plot/staging on the emotional effects of the opera performances. We manipulated the amount of prior information by organizing introductory information sessions for subgroups of the participants.

**Methods.**

As in the normal introductory talks organized by the Grand Théâtre for its visitors, we organized three different introductory talks in separate rooms at the opera house, one on the plot of the respective opera and the staging by the director given by the chief dramaturge of the Geneva opera and one on the special characteristics of the music and the composer by a musicologist regularly involved in such introductory sessions. The duration of each of these talks was approximately 30 minutes, allowing for questions from the audience. The same experimental design was used for the performances of Macbeth and Richard III - introductions to plot/staging, music/composer, and a control group receiving no information (see details under General Methods; Rosenkavalier had been excluded because of the small number of participants).

In the interest of economy of space and ease of understanding, we used only the data for opera as a whole for Richard III and for Act IV for Macbeth), using the emotion
rating component scores (based on the PCA) for this analysis. As shown in table S4 in SOM Section B) the results of the multivariate ANOVAs for the information session manipulation did not yield a single significant result. The effect sizes were generally very low. The multivariate Fs were also far from significance. These results suggest that receiving information during a brief 30 minute introduction to the opera, either on plot/staging or music, may not be sufficiently powerful to affect the emotional experience of people during the opera.

**Individual Differences.** Question 4 to be examined in this series of studies, as mentioned in the introduction, concerns the prediction by the *route model* that the emotions elicited by the opera might depend on dispositional and situational individual differences due to factors such as personality and mood. As in the preceding analyses, we decided to examine this question by combining the data across the three operas, to be able to generalize results and increase statistical power. Here we also used the component scores for the emotion ratings to reduce the amount of information to be processed.

**Age and gender.** We ran an ANOVA on the component scores for the emotion ratings (Table S5 in the SOM Section B) with age category (< 25 versus > 25) and gender as between subject factors. There was a significant difference in gender for the category Aesthetic-Enjoyment ($F = 14.67$, $p < .001$, partial $\eta^2 = .13$), with males feeling more “calm enjoyment” than females. There were, however, no differences between the two age categories, and no interaction effect of age and gender.
Personality. Before examining the correlations between the emotion ratings and personality, a PCA was run on the personality items to check if they could be grouped as general traits. The analysis extracted 5 components out of the 13 personality items (see table S6 in the SOM Section B), which, as expected given that the TIPI scale is based on the Big Five tradition, match the Big-Five traits: Openness (items open, critical and not reserved), Discipline (items disciplined, health oriented and organized), Extraversion (likeable and extraverted), Conformity (conventional and not rebellious) and Emotional stability (calm, independent and not anxious). We then looked at the correlations between the component scores for the emotion ratings and the component scores for the personality dimensions (Table S7 in the SOM Section B; including the information on a bootstrapping option). The results revealed only one significant correlation between the emotion component Awe and the personality dimension Openness ($r = .33$, $p = .001$, 95% C.I. = .16 - .49), suggesting that people who scored high on the Openness dimension were more fascinated during the performance than other spectators.

Mood. We also ran a PCA on the mood items (see table S8 in the SOM Section B), and extracted three components (largely compatible with the findings of Wilhelm & Schöbi, 2007), which we labeled as follows: Mood-Positive, Mood-Active and Mood-Excited. We found three correlations between the component scores for the mood ratings before the performance and the component scores for the emotion items (see table S9 in the SOM Section B, including information on a bootstrapping option). Mood-Excited was negatively correlated with Aesthetic-Enjoyment ($r = -.25$, $p = .01$, 95% C.I. = -.42 - -.06), suggesting that the more excited a listener was before the performance, the less sensitive he or she was to the aesthetics of the performance. Mood-Active was also
negatively correlated with Melancholy ($r = -.26$, $p = .007$, 95% C.I. = -.42 – -.09).

Although a correlation does not establish causality, it is possible that the more active a person was before the opera, the less melancholic they would feel while listening (melancholy being generally associated with passivity); finally, Mood-Positive was negatively correlated with Awe ($r = -.33$, $p = .001$, 95% C.I. = -.51 – -.13), meaning that people who were in a positive mood before the performance felt less impressed by it. As to the mood ratings after the performance, only one significant correlation emerged -- interestingly exactly the same as the correlation with mood before the performance - Mood-Excited After was also negatively correlated with Aesthetic-Enjoyment to about the same extent ($r = -.26$, $p = .01$, 95% C.I. = -.42 – -.10), suggesting that the excited listeners were somewhat less affected by the performance, possibly being less sensitive to the aesthetics of the performance. These results demonstrate the need to further investigate the special role of aesthetic and epistemic emotions in the domain of music appreciation (Scherer, 2004).

Discussion

Our first question was whether we could identify specific emotion profiles for different operas and for specific acts or scenes of an opera by using an in situ approach and asking a group of participants to record their reactions on a dedicated checklist. The results strongly indicate that this is indeed the case. There were major, and clearly interpretable, differences in participants’ emotional reactions to the different operas, as well as differences between acts and scenes within each opera, particularly in the case of Macbeth, where the different acts elicited different emotions, such as power and
animation (Act I) and sadness and tenseness (Acts II and IV). Significant differences were also found for Richard III (more enchantment and interest for the opera as a whole) and Der Rosenkavalier (higher animation and interest in Act II and a linear trend for increasing tenderness across the three acts of the opera).

Our second question was whether we could differentiate the emotion elicited in the spectator and the emotion perceived as expressed by the music and the performance. This is important in the light of critical evaluations of research on music and emotion which often claim that participants basically report what they feel the music expresses rather than their own feelings. As mentioned in the Introduction the danger of this confusion is all the more serious in opera, given the explicit role of emotion displays in the plot and interpretation of the singers. However, our results for Macbeth echo the earlier findings showing a large degree of independence of these two aspects. Thus, the sleepwalking aria in Scene 4 of Act IV is seen as expressing high tenseness and sadness on the part of Lady Macbeth, which is partly reflected in the participants own feelings, but the aria is also experienced by them as very enchanting. We also find the opposite pattern - joyful music producing tenseness: During the brindisi scene in Act II participants perceive the expression of high power and enchantment but feel themselves high tenseness and sadness, presumably because they are appraising the dire consequences to be expected in the future.

A recent study highlighted empathy as a possible link between perceived and felt emotions in music (Egermann & McAdams, 2013). The findings suggest that participants with higher scores for empathy saw the difference between their felt and perceived scores diminished (see also Miu & Balteş, 2012). Our results for Macbeth are compatible with this pattern: The high level of perceived power/animation (and the low...
level of sadness/nostalgia) perceived in Lady Macbeth's expression of determination to
act in Act I scene 7 seems to be shared by the participants. Empathy and similar
mechanisms may indeed be partly responsible for this similarity between perceived and
felt feeling -- possibly due to the ratings of participants with high dispositional empathy.
As suggested by Scherer & Coutinho (2013, p. 140) the induction mechanisms are
numerous and ubiquitous and can be recruited in various combinations at several levels
of the emotion induction process.

The results for our third question showed that in this case there was no significant
effect of introduction/prior information (i.e., receiving a briefing on music or plot/staging
compared to a control group) on the emotions attributed to the different parts of the
operas. It is possible that the introductions of about 30 minutes were too short to be
effective, although they were delivered by professionals who generally animate such
sessions for the opera public (for about the same length of time). Clearly, this negative
finding would need to be replicated before drawing any further conclusions. One
possibility might be that in music and opera the knowledge domain does not directly
impact on the affect domain. It is interesting, though, that previous studies found no
differences in emotional experience between music experts and non-experts (with a much
greater degree of knowledge and expertise involved; Bigand, Vieillard, Madurell,
Marozeau, & Dacquet, 2005).

We found some interesting effects of individual differences on the emotional
ratings. First, male participants felt more aesthetic enjoyment (which comprises calm,
harmony, and absence of tenseness and sadness) than females. We can only speculate
why this is so. Maybe it is linked to findings showing that females often respond in more
empathetic fashion to music than males, privileging the underlying social relationships (Egermann & McAdams, 2013; Balteş & Miu 2014).

As predicted by the route model outlined in the Introduction, we found significant effects of mood and personality on the emotional reactions. The participants’ mood before the performance affected their emotional experience during the different operas. People who were in an excited mood before the opera felt less aesthetic enjoyment. This is probably explained by the fact that, as shown before, aesthetic enjoyment is part of a “calm” dimension that includes the absence of tenseness and sadness. People who were in a positive mood before felt less awe during the performance which suggests that generally positive mood is not necessarily conducive to being awed by transcendence or the sublime. Similarly, people who were in an active mood before the opera felt less melancholy during the performance.

Personality variables correlated only weakly with participants' affective reactions. However, there was a significant positive correlation between Openness and the feeling of Awe. This corresponds to the finding that individuals who generally score high in Openness on the Big5 personality inventory tend to be particularly sensitive to art and beauty (McCrae, 2007), suggesting that opera spectators with a high degree of Openness to experience might be more sensitive to non-functional qualities of music involving transcendence, the sublime, and awe.

Limitations. This study has a number of limitations, mainly related to the lack of complete experimental control. However, it seems virtually impossible to impose stringent controls in settings such as a performance in a big opera house. To empirically study the elicitation of emotion by operatic performances in situ, one has to accept the
inherent limitations of ecological settings, as there is no other option for obtaining reactions of the public to a live performance. We also used groups of paid participants, mostly young people, rather than the usual population of opera visitors with a higher age level. Although this limits generalizability, it also has a positive effect: the absence of habitual preferences. Our participants were interested in and somewhat knowledgeable about opera but they did not participate in the study because they particularly liked one of the operas. This makes the ratings much more comparable than would be the case with unselected members of a normal audience who may often have chosen to attend because of their preferences or social habits.

There is also an inherent limitation because it is impossible to use the same rating format to get the intensities of different emotions experienced during a brief scene that just ended and to get a summary assessment of the frequency with which different emotions have been experienced throughout a complete opera. This means that it is impossible statistically to compare directly the difference in impact between scenes and whole operas. We have standardized the different measures and analyzed them separately, interpreting both of them in terms of the relative importance or prevalence of certain emotions.

Conclusion

A rapidly growing number of experimental laboratory studies of musically induced emotions, generally using recordings of different types of instrumental or synthetic music, are providing major additions to our knowledge about the underlying mechanisms, especially concerning the role of musical structure or architecture in terms of form, timing, rhythm, or tonal structure. The experimental manipulation of some of these factors allows rigorous testing of specific hypotheses. However, if it is generally
difficult to transpose real-life issues into the laboratory in an ecologically valid fashion, it is all the more difficult in the case of emotional effects of music, for which the performance aspect is a central ingredient. This is especially the case for opera as Gesamtkunstwerk, in which drama, stage setting and lighting, dynamic movement of the protagonists on the stage, affect expression in speech, singing and face/body movements, and powerful embedding in the music played by a large orchestra all contribute to the overall emotional effect. Although large-scale video projections in a university lecture hall can provide an approximation of the overall setting and the performance of singers and musicians (see Coutinho & Scherer, 2016, for a comparison of a church setting for Lieder with a video reproduction in an auditorium), it is highly desirable to study the phenomenon in an authentic performance context, such as an opera house.

A major impediment for studies of this sort has been the nature of the measurement instruments for emotional reactions (e.g., the widely used GEMS scale; Zentner et al., 2008), which generally require participants to describe their affective reactions on a large number of items. The work reported in this article has employed the beta version of the new GEMIAC checklist (Coutinho & Scherer, 2017), specifically developed to be used in the framework of in situ music performances in different public settings. The GEMIAC consists of a checklist with a small number of fuzzy-set items, fitting on a single page but representing the major dimensions of emotions likely to be elicited by different kinds of music performances, including aesthetic emotions. It is particularly well suited to be used during brief pauses in musical works to obtain ratings of emotional reactions that are as close as possible to the musical stimuli that produced them. The checklist can even be employed in a dynamic fashion to capture audience
As indicated in the introduction, in the absence of appropriate research models the current work was conceived as an exploratory venture rather than as a test of specific hypotheses. However, the results reported above can be used to address a certain number of questions and hypotheses posed in the literature. Below, we summarize the contributions of this research to the field.

On the methodological side, the GEMIAC has fully confirmed the hopes that the developers had entertained: Participants completed their ratings reliably and in very short periods of time. They showed themselves highly satisfied with the procedure, especially the fact that there was only a minimal interruption of the enjoyment of the musical event. It is to be hoped that the availability of this new tool will generate an increasing number of empirical studies of emotional reactions to different kinds of music performances in naturalistic settings.

Importantly, our results show that the relatively small number of fuzzy-set emotion descriptions is largely sufficient to differentiate the affective reactions that were targeted in this research. One important issue was to trace the differences in emotional response to different parts of an opera and different operas. The results show very marked differences in the relative importance or prevalence of certain emotions between different acts for Macbeth and between a central scene and the opera as a whole for Richard III demonstrating the fine resolution provided by the instrument and the dynamic affective sensitivity of the spectators. One remarkable finding is the prevalence of epistemic emotions for Battistelli’s opera Richard III with its resolutely
contemporary musical score eliciting strong interest (and, correspondingly, the absence of boredom). Another result that shows the high resolution of the rating scales is the strong increase in feelings of tenderness across the three acts of Der Rosenkavalier, very much in line with musicological analyses of the music (especially in the final scene). In response to question 1) at the end of the introduction, we found that we could indeed identify specific “feeling profiles” for the affective reactions to different operas and for specific acts or scenes of an opera.

As to question 2), the results for Macbeth show that we can differentiate the emotion elicited in the spectator by an operatic performance and the emotion the spectator perceived as being expressed by the performers. This is particularly important with respect to the mechanisms of emotion elicitation suggested in the route or BRECVEM models and the various hypotheses proposed in different contributions to the literature. One particularly popular hypothesis is that felt emotions almost always correspond to expressed emotions, due to entrainment, congruent appraisal, contagion or empathy. While the results for Acts I and IV of Macbeth would be consistent with any of these explanations, Act II really stands out as the felt emotions are essentially diametrically opposite to the emotions expressed by the protagonists in the opera, ruling out entrainment, contagion, or direct empathy based on expression. We have suggested that this discrepancy may be due to the spectators see the impending disaster coming, implying a predictive empathic appraisal process. Further theoretical and empirical work is needed to more clearly define the potential mechanisms and develop appropriate research designs to examine these.

The effect of prior information (concerning the music vs. plot/staging) on the affective responses to the performance, Question 3), seems to be negligible. One might
have assumed that a detailed introduction to the plot would favor empathy or appraisal mechanisms but that does not seem to be the case. This does not mean that the introductory information sessions organized by many opera houses just before a performance are without effect and thus superfluous – they may well have an important effect on cognitive appreciation and thus be an important aspect of the overall enjoyment of a visit to the opera and increase in general culture.

Finally, on question 4, the potential effects of individual differences between spectators, including personality and prior mood, on emotion elicitation, our results suggest that these may be less important than generally believed, although we found some interesting leads that are worth pursuing. However, we only scratched the surface of potentially important individual differences such as musicianship, musical capacity, music preferences and motivations for music use. Recently a new modular tool to broadly assess participant background for music research has become available (MUSEBAQ; Chin, Coutinho, Scherer, & Rickard, in press) which can greatly facilitate future research and provide for better comparability of the samples in different studies.

Overall, the empirical examination of the route model of emotion induction via music of different genres requires a combination of laboratory approaches with in situ studies; allowing the assessment of complex interactions between the different determinants specified by the model. Future research needs to use more experimentally controlled methods to chart some of the most important routes and our results suggest the need to consider that several mechanisms may be active depending on different conditions. Our results confirm, in a real-life musical setting, earlier findings of the important differences between emotions felt while listening to music from emotions simply perceived or inferred from salient features of the music or the performance. This
does not mean that these are entirely independent as they may well influence each other (e.g., via interactions between the different induction routes, see Scherer & Coutinho, 2013). Further work on the nature of these induction mechanisms will need to disentangle these complex processes -- a discovery process that will require both controlled laboratory and in situ research at public music venues.

A major desideratum for future work is to further differentiate the nature of the affective reactions produced by music. Given that the generic term “emotion” often leads to confusion in the sense of implying rather strong reactions, such as “basic” emotions, it is important to fine-tune the dependent measures in this research tradition. Konečni (2008) has rightly pointed out that music can produce a whole panoply of cognitive, affective, and physiological reactions that should be distinguished. In addition, the need to acknowledge the central role of aesthetic feelings and music-specific affective reactions has been highlighted for some time (Konečni, 2005; Scherer, 2004; Scherer & Zentner, 2008). There have been recent efforts to systematically chart the semantic space of aesthetic emotions (Hosoya, Schindler, Beermann, Wagner, Menninghaus, Eid, & Scherer, in press). This work has allowed the development of a new scale to measure aesthetic emotions in response to different works of art, including music – the Aesthetic Emotions Scale (AESTHEMOS). With 21 subscales the instrument covers prototypical aesthetic emotions (e.g., the feeling of beauty, being moved, fascination, and awe), epistemic emotions (e.g., interest and insight), emotions indicative of amusement (humor and joy), negative emotions (e.g., the feeling of ugliness, boredom, and confusion) as well as activating (energy and vitality) and the calming (relaxation) effects of aesthetic experience (Schindler, Hosoya, Menninghaus, Beermann, Wagner, Eid., & Scherer, 2017).
We conclude that, given the scarcity of available models and the exploratory nature of the current research, our findings need to be replicated and extended in future research on different types of music, including vocal music and especially opera. We feel that restricting research on music and emotion to instrumental (pure or absolute) music is stifling progress in this field, especially that all types of music need to be performed by musicians and singers in a certain physical place (or on electronic media). In consequence, the effects of the “pure” music can never be separated from the effects of interpretation and local context. In addition, listeners will always differ in terms of social-demographic variable, personality, mood, prior information and expertise and these will obviously color their emotional experiences. In consequence, studies in this area need to make much more of effort to capture all potential variables of influence and determine their relative effect size and the nature of their interaction. We believe that with the work reported here we have shown that this kind of in situ research is possible and promising, especially as new instruments become regularly available.
References


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.

power of music (pp. 122–145). Oxford: Oxford University Press, United Kingdom.


Figure captions

Figure 1. Mean $z$-scores for each emotion term during Act I scene 7 (top), Act II scene 5 (middle), and Act IV scene 4 (bottom) of Macbeth. The diamonds show the emotions perceived, and the squares show emotions felt by the listeners.

Figure 2. Graphic illustration of the interaction effects between Act and Quality (perceived vs. felt emotions) in Macbeth for four emotion categories (Power, Sadness, Enchantment, and Tenseness).

Figure 3. Mean $z$-score profile for each emotion word in the selected scene (diamonds) and the whole opera (squares) in Richard III.

Figure 4. Mean $z$-scores for each emotion term in the three parts: Act II (diamonds), Act III (squares), and whole opera (triangles) in Der Rosenkavalier.

Figure 5. Emotional profiles for the final rating period of the three operas for the emotion categories: Macbeth (circles), Richard III (diamonds), and Rosenkavalier (triangles).
Table 1

Repeated-measures ANOVA: Main effect of Scene/Act on the ratings for the felt emotion categories in Macbeth

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Scene/Act</th>
<th>Means</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>df</td>
</tr>
<tr>
<td>Enchantment</td>
<td>13.19</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Awe</td>
<td>1.60</td>
<td>2/76</td>
<td>.208</td>
</tr>
<tr>
<td>Tenderness</td>
<td>2.32</td>
<td>2/76</td>
<td>.105</td>
</tr>
<tr>
<td>Nostalgia</td>
<td>27.09</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Calm</td>
<td>1.64</td>
<td>2/76</td>
<td>.201</td>
</tr>
<tr>
<td>Power</td>
<td>38.95</td>
<td>1.3/57.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Animation</td>
<td>34.56</td>
<td>1.4/54.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Tenseness</td>
<td>4.42</td>
<td>2/76</td>
<td>.015</td>
</tr>
<tr>
<td>Sadness</td>
<td>44.63</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Harmony</td>
<td>1.40</td>
<td>2/76</td>
<td>.25</td>
</tr>
<tr>
<td>Interest</td>
<td>27.96</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Boredom</td>
<td>0.24</td>
<td>1.7/66.7</td>
<td>.754</td>
</tr>
</tbody>
</table>

Note. ANOVA = analysis of variance.
Table 2

Repeated-measures ANOVA: (a) Main effects of Scene/Act, (b) Quality, and (c) interaction effect of Scene/Act × Quality on the perceived and felt emotion ratings in Macbeth.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Scene/Act</th>
<th></th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>$df$</td>
<td>$p$</td>
</tr>
<tr>
<td>Enchantment</td>
<td>1.55</td>
<td>2/76</td>
<td>.218</td>
</tr>
<tr>
<td>Awe</td>
<td>0.15</td>
<td>2/76</td>
<td>.863</td>
</tr>
<tr>
<td>Tenderness</td>
<td>1.64</td>
<td>2/76</td>
<td>.201</td>
</tr>
<tr>
<td>Nostalgia</td>
<td>50.92</td>
<td>1.4/54.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Calm</td>
<td>0.58</td>
<td>1.8/66.7</td>
<td>.542</td>
</tr>
<tr>
<td>Power</td>
<td>81.05</td>
<td>1.6/61.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Animation</td>
<td>66.06</td>
<td>1.6/59.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Tenseness</td>
<td>3.60</td>
<td>1.5/56.7</td>
<td>.046</td>
</tr>
<tr>
<td>Sadness</td>
<td>59.49</td>
<td>1.4/59.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Harmony</td>
<td>2.84</td>
<td>2/76</td>
<td>.065</td>
</tr>
<tr>
<td>Interest</td>
<td>29.17</td>
<td>1.6/60.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Boredom</td>
<td>1.06</td>
<td>1.7/62.8</td>
<td>.341</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Quality</th>
<th></th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>$df$</td>
<td>$p$</td>
</tr>
<tr>
<td>Enchantment</td>
<td>0.21</td>
<td>1/38</td>
<td>.648</td>
</tr>
<tr>
<td>Awe</td>
<td>0.48</td>
<td>1/38</td>
<td>.493</td>
</tr>
<tr>
<td>Tenderness</td>
<td>2.12</td>
<td>1/38</td>
<td>.154</td>
</tr>
<tr>
<td>Nostalgia</td>
<td>40.98</td>
<td>1/38</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Calm</td>
<td>1.84</td>
<td>1/38</td>
<td>.183</td>
</tr>
<tr>
<td>Power</td>
<td>77.02</td>
<td>1/38</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Emotion</td>
<td>$F$</td>
<td>$df$</td>
<td>$p$</td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Enchantment</td>
<td>27.18</td>
<td>1.4/23.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Awe</td>
<td>3.61</td>
<td>2/76</td>
<td>.036</td>
</tr>
<tr>
<td>Tenderness</td>
<td>4.64</td>
<td>2/76</td>
<td>.013</td>
</tr>
<tr>
<td>Nostalgia</td>
<td>38.73</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Calm</td>
<td>4.80</td>
<td>1.3/50.5</td>
<td>.024</td>
</tr>
<tr>
<td>Power</td>
<td>20.42</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Animation</td>
<td>14.60</td>
<td>1.8/67.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Tenseness</td>
<td>17.49</td>
<td>2/76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Sadness</td>
<td>57.19</td>
<td>1.4/57.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Harmony</td>
<td>4.81</td>
<td>1.6/60.8</td>
<td>.017</td>
</tr>
<tr>
<td>Interest</td>
<td>4.69</td>
<td>2/76</td>
<td>.012</td>
</tr>
<tr>
<td>Boredom</td>
<td>0.05</td>
<td>1.7/64.7</td>
<td>.927</td>
</tr>
</tbody>
</table>

*Note. ANOVA = analysis of variance.*
Table 3
Rotated factor matrix of a PCA analysis of the complete set of ratings on the 12 fuzzy emotion categories

<table>
<thead>
<tr>
<th></th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Calm</td>
<td>0.813</td>
</tr>
<tr>
<td>Tenseness</td>
<td>-0.793</td>
</tr>
<tr>
<td>Sadness</td>
<td>-0.624</td>
</tr>
<tr>
<td>Harmony</td>
<td>0.526</td>
</tr>
<tr>
<td>Interest</td>
<td>0.766</td>
</tr>
<tr>
<td>Animation</td>
<td>0.725</td>
</tr>
<tr>
<td>Nostalgia</td>
<td>0.786</td>
</tr>
<tr>
<td>Power</td>
<td>-0.693</td>
</tr>
<tr>
<td>Awe</td>
<td>0.846</td>
</tr>
<tr>
<td>Enchantment</td>
<td>0.505</td>
</tr>
<tr>
<td>Tenderness</td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td></td>
</tr>
</tbody>
</table>