



It's not what you play, it's how you play it: Timbre affects perception of emotion in music

Julia C. Hailstone, Rohani Omar, and Susie M. D. Henley

Dementia Research Centre, Institute of Neurology, University College London, London, UK

Chris Frost

Dementia Research Centre, Institute of Neurology, University College London, London, UK, and Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK

Michael G. Kenward

Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK

Jason D. Warren

Dementia Research Centre, Institute of Neurology, University College London, London, UK

Salient sensory experiences often have a strong emotional tone, but the neuropsychological relations between perceptual characteristics of sensory objects and the affective information they convey remain poorly defined. Here we addressed the relationship between sound identity and emotional information using music. In two experiments, we investigated whether perception of emotions is influenced by altering the musical instrument on which the music is played, independently of other musical features. In the first experiment, 40 novel melodies each representing one of four emotions (happiness, sadness, fear, or anger) were each recorded on four different instruments (an electronic synthesizer, a piano, a violin, and a trumpet), controlling for melody, tempo, and loudness between instruments. Healthy participants (23 young adults aged 18–30 years, 24 older adults aged 58–75 years) were asked to select which emotion they thought each musical stimulus represented in a four-alternative forced-choice task. Using a generalized linear mixed model we found a significant interaction between instrument and emotion judgement with a similar pattern in young and older adults ($p < .0001$ for each age group). The effect was not attributable to musical expertise. In the second experiment using the same melodies and experimental design, the interaction between timbre and perceived emotion was replicated ($p < .05$) in another group of young adults for novel synthetic timbres designed to incorporate timbral cues to particular emotions. Our findings show that timbre

Correspondence should be addressed to J. D. Warren, Dementia Research Centre, Institute of Neurology, University College London, Queen Square, London, WC1N 3BG, UK. E-mail: jwarren@drc.ion.ucl.ac.uk

We thank the volunteers for their participation. We are grateful to Jonas Brolin of the London College of Music for recording the trumpet stimuli and to Professor Tim Griffiths of the University of Newcastle upon Tyne for assistance in writing the Matlab[®] script to synthesize artificial timbres. This work was supported by the Wellcome Trust, the UK Medical Research Council, and the Alzheimer Research Trust. R.O. is supported by a Royal College of Physicians/Dunhill Medical Trust Joint Research Fellowship. J.D.W. is supported by a Wellcome Trust Intermediate Clinical Fellowship. The Dementia Research Centre is an Alzheimer's Research Trust Co-ordinating Centre. This work was undertaken at University College London who received a proportion of funding from the Department of Health's National Institute for Health Research (NIHR) Biomedical Research Centres funding scheme.

(instrument identity) independently affects the perception of emotions in music after controlling for other acoustic, cognitive, and performance factors.

Keywords: Timbre; Emotion; Music; Auditory object.

Our most salient sensory experiences often have a strong emotional tone. However, the neuropsychological relations between perceptual characteristics of sensory objects and the affective information they convey are poorly defined. Music is relevant to this issue because it contains a variety of distinctive auditory objects (musical instruments and melodies) and because the emotional content of music is for most listeners its paramount attribute (Blood & Zatorre, 2001; Dalla Bella, Peretz, Rousseau, & Gosselin, 2001; Sloboda & O'Neill, 2001). Music is also a useful stimulus because it has relatively well-defined structural perceptual features that can be manipulated simply and independently. Normal listeners are able to achieve high levels of agreement about the emotional content of particular musical pieces, demonstrating that alteration in the structural properties of music can reliably alter its emotional valence (Gabrielsson & Juslin, 1996; Juslin, 1997; Robazza, Macaluso, & Durso, 1994; Terwogt & Van Grinsven, 1988). Lesion and functional imaging studies in human participants have shown that the processing of emotion in music involves mesolimbic and other brain regions associated with the processing of strong emotion in a variety of contexts (Blood & Zatorre, 2001; Gosselin, Peretz, Johnsen, & Adolphs, 2007; Menon & Levitin, 2005; Peretz, Gagnon, & Bouchard, 1998). This suggests that music, despite its abstract nature, recruits neural machinery for processing emotion similar to animate sensory objects (Dalla Bella et al., 2001; Juslin & Laukka, 2003). Music is therefore potentially a useful model for studying the effects of manipulating perceptual features on the emotional content of a sensory stimulus.

Perhaps the most basic and widely studied structural feature of a musical piece is the sequence of pitches comprising the melody. Normal

listeners can reliably distinguish a "happy" from a "sad" tune (Dalla Bella et al., 2001; Robazza et al., 1994; Terwogt & Van Grinsven, 1988), and manipulations of pitches in the melodic line can alter expression of emotion between happiness and sadness. Many other structural characteristics of a musical piece besides the melodic line, including sound level, vibrato, tempo, rhythm, mode, and harmony (consonance–dissonance), have been shown to influence psychophysiological correlates of emotion (Sloboda, 1991) and judgments of happiness, sadness, and other basic emotions in music (Gabrielsson & Juslin, 1996; Juslin, 1997; Peretz et al., 1998). The choice of musical instrument or combination of instruments is also highly relevant to the expression of emotion (Balkwill & Thompson, 1999; Behrens & Green, 1993; Gabrielsson, 2001; Gabrielsson & Juslin, 1996; Juslin, 1997). Composers select particular instruments (either consciously or intuitively) to convey specific emotional qualities: The plaintive cor anglais solo in the *New World Symphony* and the comical bassoon part in *Peter and the Wolf* are striking examples, and there are countless other instances in all genres of music (Copland, 1957; Gabrielsson & Juslin, 1996; Sloboda & Juslin, 2001) and across musical cultures (Jackendoff & Lerdahl, 2006; Schneider, 2001).

The identity of a musical instrument is conveyed by its timbre: This is defined operationally as the acoustic property that distinguishes two sounds of identical pitch, duration, and intensity and is integral to the identification of all kinds of auditory objects, including musical instruments, voices, and environmental sounds (Bregman, Levitan, & Liao, 1990; Griffiths & Warren, 2004; McAdams & Cunible, 1992). Timbre is a complex property with several distinct dimensions (Caclin, McAdams, Smith, & Winsberg, 2005; McAdams, 1993), including spectral envelope,

temporal envelope (attack), and spectral flux (the dynamic interaction of spectral and temporal properties). Timbre might in principle occupy a critical interface linking the distinctive structural features of an auditory object, such as a musical instrument, with the emotional content of that object. Human psychophysical work has shown that instrumental timbre contributes to emotion judgements in music (Balkwill & Thompson, 1999), while electrophysiological evidence suggests that alterations in musical emotional expression can be detected at the level of single notes (Goydke, Altenmuller, Moller, & Munte, 2004). Aspects of timbre such as attack and frequency spectrum contribute to the perception of particular emotions in music (Gabrielsson & Juslin, 1996; Juslin, 1997). This evidence suggests that timbral features can at least in part determine the emotional content of music independently of melodic and other cues or if these other cues are absent. Evidence from human lesion studies further supports an interaction between timbre and emotion: Relatively selective impairment of timbre perception occurs following acute damage in the region of the non-dominant posterior superior temporal lobe, and affected individuals commonly describe a loss of pleasure and reduced emotional responsiveness to music (Kohlmetz, Muller, Nager, Munte, & Altenmuller, 2003; Stewart, von Kriegstein, Warren, & Griffiths, 2006). As timbre is a general property of different categories of sounds, including music, voices, and environmental noises, it may allow us to identify generic mechanisms that enable object features to influence emotion recognition, both in sounds and in sensory objects in other modalities. However, in order to isolate any specific effects of timbral variation on perceived emotion in music, it is necessary to control structural features such as melody, tempo, dynamic range, pitch range, and semantic effects based on learned associations with familiar music.

Here we investigated whether the identification of emotion in melodies by normal participants is influenced by the timbre (instrument) conveying the melody independently of other musical features. In designing this study, we were particularly

concerned to assess the role of timbre after controlling for factors such as structural cues to emotion perception, familiarity, musical background, and exposure, and the role of human performance variables, which have tended to confound the interpretation of any purely timbral effects in previous work. In order both to control musical structural characteristics and to minimize learned associations with particular musical pieces, we created a set of novel melodies. Two experiments were conducted. In Experiment 1, melodies were conveyed by instruments familiar to most Western listeners, and two age groups (young adults and older adults) were studied. Inclusion of older adults was designed to assess the stability of any timbral effect over different musical exposures and over the course of the lifetime. Older adults have demonstrated worse performance on other cognitive musical tasks (Dowling, Bartlett, Halpern, & Andrews, 2008) and greater negative emotional reactions to music in everyday listening (Sloboda & O'Neill, 2001). It is likely a priori that different age cohorts will have had different musical exposures throughout the course of a lifetime and may therefore have learned distinct patterns of association between particular musical instruments and the emotions they are used to convey. Furthermore, an age-related decline has been reported for identification of facial expressions (Calder et al., 2003; Fischer et al., 2005) and vocal emotions (Ruffman, Henry, Livingstone, & Phillips, 2008), and there may be an age-related interaction with particular emotions and modalities (Ruffman et al., 2008), raising the possibility that ageing per se affects the perception of emotion in music. In Experiment 2, melodies were conveyed exclusively by synthetic timbres: These novel "instruments" were designed to incorporate timbral cues to particular emotions and to minimize the potentially confounding effects of learned emotional associations with real instruments and to remove human performance cues. The study was approved by the local research ethics committee, and all participants gave informed consent in accord with the principles of the Declaration of Helsinki.

EXPERIMENT 1

Novel melodies were composed with the intention of conveying a defined emotional tone, since it is difficult to create a truly emotionally “neutral” melody, and in everyday listening music generally has an affective tone. Basic emotions (rather than more abstract feeling states) were chosen to align with the fundamental work on emotion perception from facial expressions by Ekman (1992) showing that these emotions are universally recognized by normal human participants, and because they can be represented in music. The melodies were performed on instruments that are currently in widespread use and were likely to have been heard before by most listeners without musical training. We also sought to determine whether any effect of timbre on perceived emotion was robust to individual differences in musical training and the frequency with which individuals currently listen to music. The effect of musical listening experience on any timbre–emotion effect was further examined using two age cohorts (young and older adults). We reasoned that these age groups should have had the opportunity to form a range of specific emotional associations with musical instruments used in popular songs and other contemporary music.

Method

Participants

Healthy volunteers with no history of hearing symptoms or neurological or psychiatric illness were recruited. Participants comprised two age ranges: young adults, aged 18–30 years ($n = 23$, 12 females; mean age 24.7 years, $SD = 3.0$) and older adults, aged 58–75 years ($n = 24$, 14 females; mean age 66.7 years, $SD = 5.2$). None of the participants was a professional musician. All participants completed a questionnaire in which they were asked to detail any formal musical training (instrumental or vocal) and to estimate the number of hours per week they currently listened actively to music; responses on this questionnaire were used to index participants’

level of musical training and current exposure to music. There was substantial variation in musical background and listening experience among participants. However, the majority of participants (14/23 younger participants, 20/24 older participants) indicated that they listened regularly to Western classical music (a full summary of participant characteristics is available on request).

Stimulus composition

Novel melodies were composed to represent one of four emotions: happiness, sadness, anger, or fear. A total of 80 single-line melodies (20 representing each of the four target emotions) containing 8 bars (two 4-bar phrases) of music were composed by one of the authors (J.C.H.) using Western musical theory. Melodies were composed within a two-octave range in the treble clef, contained a regular pulse, and used a range of major and minor keys and meters. From this pool, 10 melodies that most successfully represented each of the four target emotions based on independent ratings by the authors (J.C.H., R.O., S.M.D.H., J.D.W.) were selected for inclusion in the final stimulus set (none of the authors participated in the subsequent experiments).

The characteristics of the selected melodies are summarized in Table 1. Notations of four exemplars are shown in Figure 1, and all melodies are available on request. Mode, key, tempo, rhythm, meter, and dynamic range varied between stimuli from different emotional categories, as these structural acoustic factors are likely to play an important role in representing particular emotions in music (Dalla Bella et al., 2001; Gabrielsson & Juslin, 1996; Juslin, 1997; Peretz et al., 1998). Despite variation in mean tempo between emotional categories, there was extensive overlap in tempi of individual stimuli from the categories of happiness, anger, and fear, and overlap between the tempi of sadness and fear stimuli. A total of 15 of the total set of 40 melodies contained the same pattern of pitches as at least one other stimulus conveying a different emotion; most of these shared pitch structures (12/15) occurred between “happy” and “sad” melodies, and the target emotion was conveyed by altering mode,

Table 1. Structural characteristics of melodies

| Structural property | | Intended emotion | | | |
|----------------------|------------------------|------------------|-------------|-------------|----------------|
| | | Happiness | Sadness | Anger | Fear |
| Tonality | Mode | Major | Minor | Minor | Minor |
| | No. chromatic notes | None | None | Few | Many |
| Tempo (beats/minute) | Mean (<i>SD</i>) | 189 (31) | 86 (11) | 223 (32) | 151 (53) |
| | Min:max | 156 : 257 | 68 : 100 | 182 : 274 | 64 : 245 |
| Metre (% of stimuli) | 4/4 | 50 | 20 | 60 | 60 |
| | 3/4 | 30 | 70 | 20 | 20 |
| | 2/4 | 10 | — | 10 | — |
| | Compound | 10 | 10 | 10 | 20 |
| Dynamics | Range | Small range | Small range | Small range | Wide variation |
| | At beginning of phrase | loud | soft | loud | variable |
| | At end of phrase | loud | soft | loud | variable |
| | No. accented notes | none | none | few | many |

tempo, and rhythm. These “repeat” melodies were included because they were successful representatives of particular emotions (based on pilot ratings); the study was not designed to assess any effect from repetition of pitch structures across emotion conditions.

Stimulus recording

All stimuli were recorded on four instruments with distinctive timbres: piano (keyboard/percussion timbre), violin (string timbre), trumpet (brass timbre), and electronic synthesizer (“electronic”

timbre). Stimuli were performed on the “natural” instruments by two experienced conservatory-trained classical musicians, and the stimuli for the piano and violin were recorded by the same performer (J.H.). No constraints were imposed on performance techniques. The electronic synthesizer was generated using Sibelius v4® (2005) composition software (<http://www.sibelius.com>). Melodies written in the treble clef could be performed using the same pitch range for the piano, violin, trumpet, and synthesizer. The tempo and loudness of the stimuli were controlled to ensure there were no systematic differences between instruments with regard to these properties. The average tempo for each emotion was controlled by providing the performer with the metronome tempo based on the initial piano recordings made by the composer; durations of the stimuli in the final set ranged from 9–31 s (mean 17 s). All stimuli were stored as digital wave files on a laptop computer, and the mean (root mean square) intensity level was fixed for all stimuli after recording.

happiness



sadness



anger



fear



Figure 1. Notations for representative melodies exemplifying each target emotion.

Experimental procedure

Stimuli were delivered to participants via headphones at a comfortable loudness level, in a four-alternative forced-choice paradigm under Matlab v7.0® (2004) (<http://www.mathworks.com>). Participants were instructed to select which one of the four

emotion labels (“Happiness”, “Sadness”, “Anger”, or “Fear”) was represented by each melody. Four practice examples on the piano were presented initially to ensure that the participant understood the task. No feedback about performance was given during the experiment. Each participant was presented with 160 stimuli in four blocks. Each block comprised the 40 musical stimuli played on a single instrument. For each participant a different ordering of melodies was used in each block, with the ordering selected using a restricted randomization amongst all possible permutations that ensured that the same emotion could not occur more than twice consecutively. This design ensured that successive presentations of the same melody were temporally spaced and prevented the priming that might occur were the melodies repeated in the same order in the different blocks. The study was planned to include 24 younger and 24 older participants with the order of instrument presentation (or period) following that in a four-period cross-over design (Jones & Kenward, 2003). This design required that each of the 24 ($4 \times 3 \times 2$) possible instrument orderings was followed by a single participant from each age group.

Statistical analysis

Generalized linear mixed regression models (Verbeke & Molenberghs, 2000), fitted separately to the young and older participant groups, were used to investigate the influence of instrument and target emotion on the probability of a “correct” (intended by the experimenters) response. A logistic link was used as is standard for binary outcomes (intended versus other response), and participant identity was included as a random effect to allow for associations between responses from the same individuals.

In the basic model the predictors were instrument, target emotion, period (1, 2, 3, 4), and the interaction between target emotion and instrument. A global Wald test of interaction was carried out, and odds of an intended response relative to that for a “happy” melody on the piano were estimated. Happiness on the piano was arbitrarily selected as the baseline category on the premise that both the piano timbre and the concept of

happiness represented in music should be highly familiar to all participants. Linear combinations representing the mean response over instruments were also calculated for each emotion. A joint Wald test of these linear combinations was employed to investigate global differences between target emotions. An analogous technique was employed to investigate global differences between instruments. Additional predictor variables were added to the basic model to investigate the effects of (a) self-reported years of musical training and (b) listening hours per week on the probability of a correct response, and the interactions of both these factors with target emotion.

The basic model was extended to permit comparisons between age groups in an analysis of all participants. Additional between-subject predictor variables were age group and its interaction with instrument and emotion. As an additional check on the robustness of the findings in musically “naïve” participants, the basic model was fitted to the data from the subset of participants having less than a year’s formal musical training ($n = 18$).

Results

For young participants, the mean number of “correct” responses (score /10) for each target emotion are displayed in Figure 2, and the odds ratios for a “correct” (intended) response for each combination of instrument and emotion (adjusted for order of instrument presentation), relative to the odds of a “correct” (intended) response for “happy” melodies played on the piano, are shown in Table 2. Odds ratios that are larger than 1 are indicative of an increased probability of a correct response relative to that for a “happy” melody on the piano. Those that are smaller than 1 are indicative of a reduced probability. The 95% confidence intervals that do not include 1 are statistically significant at the 5% level, but individual results must be interpreted cautiously given the number of pairwise comparisons that can be made. Strictly, participant responses here can be assessed only relative to the target emotion intended by the experimenters, since the criteria for a “correct” emotion judgement were not absolute but were

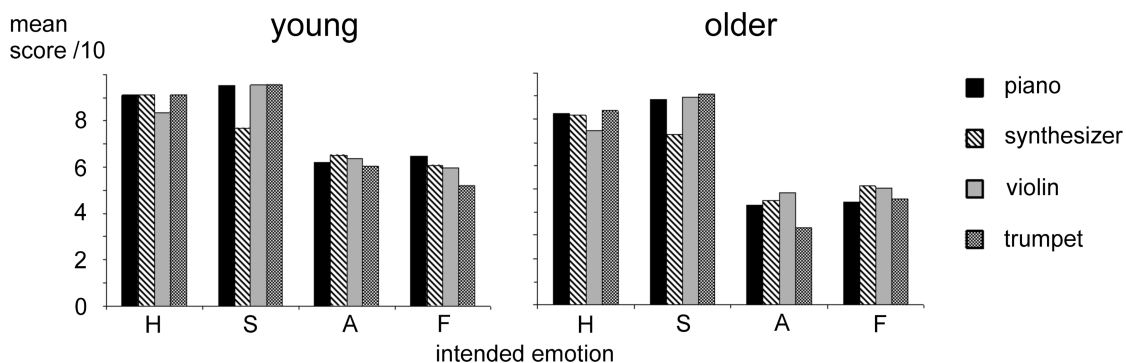


Figure 2. Mean scores (/10) for each intended emotion for each “real” instrument in young (left) and older (right) adult participants. H, happiness; S, sadness; A, anger; F, fear.

defined by the experimenters. Although a global test of differences between the odds of an intended (correct) response by instrument (averaging over emotions) was statistically significant ($p = .006$), the probability of an intended response was similar for each instrument (range .74–.78). However, the odds of an intended response differed substantially ($p < .0001$) between emotions. The mean probability of an intended response was similar for “happy” and “sad” melodies and markedly reduced for “angry” and “fearful” melodies, though identification of both anger and fear were still well above what would be expected by chance. For each melody the target emotion was selected most frequently. There was a statistically

significant interaction between instrument and emotion judgement ($p < .0001$): For “happy” melodies the odds of an intended response were lower for the violin than for the other instruments, while for “sad” melodies, the odds of an intended response were lower for the synthesizer than for the other instruments.

Older participants showed a very similar pattern of results. The mean number of “correct” responses (score /10) for each target emotion are displayed in Figure 2, and odds ratios for an intended response for each combination of instrument and emotion (adjusted for order of instrument presentation) relative to the odds of an intended response for “happy” melodies played

Table 2. Odds of a “correct” response relative to “happy” melodies played on piano for young and older adults in Experiment 1

| Participants | Instrument | Intended emotion | | | | | | | |
|--------------|-------------|------------------|------------|---------|------------|-------|------------|------|------------|
| | | Happiness | | Sadness | | Anger | | Fear | |
| | | OR | CI | OR | CI | OR | CI | OR | CI |
| Young | Piano | 1 | | 1.92 | 0.89, 4.12 | 0.15 | 0.08, 0.25 | 0.16 | 0.10, 0.28 |
| | Synthesizer | 1.01 | 0.52, 1.95 | 0.31 | 0.18, 0.54 | 0.17 | 0.10, 0.29 | 0.14 | 0.08, 0.24 |
| | Violin | 0.48 | 0.27, 0.86 | 2.14 | 0.97, 4.72 | 0.16 | 0.09, 0.27 | 0.13 | 0.08, 0.22 |
| | Trumpet | 1.04 | 0.54, 2.01 | 2.21 | 1.00, 4.87 | 0.14 | 0.08, 0.23 | 0.10 | 0.06, 0.16 |
| Older | Piano | 1 | | 1.63 | 0.96, 2.74 | 0.15 | 0.10, 0.23 | 0.16 | 0.10, 0.24 |
| | Synthesizer | 0.93 | 0.58, 1.50 | 0.57 | 0.36, 0.88 | 0.16 | 0.10, 0.24 | 0.21 | 0.13, 0.32 |
| | Violin | 0.62 | 0.40, 0.97 | 1.83 | 1.07, 3.13 | 0.18 | 0.12, 0.28 | 0.20 | 0.13, 0.31 |
| | Trumpet | 1.09 | 0.67, 1.77 | 2.12 | 1.22, 3.71 | 0.10 | 0.06, 0.15 | 0.17 | 0.11, 0.25 |

Note: “Correct” = intended response; OR = odds ratio of an intended response; CI = 95% confidence intervals.

on the piano are shown in Table 2. Averaged over all emotion categories, the probability of an intended response was similar for each instrument (range .63–.66), and a global test of differences in odds was not statistically significant ($p = .14$). Across instruments, the probability of an intended response was similar for “happy” and “sad” melodies and was markedly reduced for both anger and fear ($p < .0001$, global test for differences in odds by emotion). As with the young participants, identification of both anger and fear was above that which would be expected by chance. There was a significant interaction between instrument and emotion judgement ($p < .0001$): For “happy” melodies, the odds of an intended response were lower for the violin than for the other instruments ($p < .05$); for “sad” melodies, the odds of an intended response were lower for the synthesizer than for the other instruments ($p < .05$); and for “angry” melodies, the odds of an intended response were lower for the trumpet than for the other instruments ($p < .05$).

Comparing the two age groups, the young participants were significantly more likely ($p < .0005$) to obtain the intended responses across emotion categories and instruments than were the older participants, and there was a formally statistically significant ($p = .03$) interaction between age and emotion. This age–emotion interaction primarily reflected the relatively worse performance of older adults for recognition of anger (and to some extent happiness) in melodies (see Figure 2). However, in view of the borderline nature of the statistical significance of this interaction, this result should be interpreted cautiously.

There was considerable individual variation in performance both within and between age cohorts. We postulated that individual performance might be influenced by years of musical training and listening hours per week (time each week currently spent playing an instrument was not modelled, as most participants in both age cohorts no longer played regularly). After extending the model to incorporate years of musical training and current listening hours per week, the main effect of age was reduced but remained statistically significant ($p = .006$); however, the

interaction between age and emotion no longer reached statistical significance. In this model there was a main effect of musical training and an interaction of this factor with emotion judgements ($p < .01$). The odds of a correct response increased with years of musical training, with this response being most marked for “happy” melodies and next most marked for “angry” ones. There was no such effect for weekly listening hours.

A model restricted to those participants with one year or less of musical training across age cohorts ($n = 18$) showed similar results to those fitted to the young and older cohorts separately. Similar trends were apparent by age and emotion, and the interaction between age and emotion judgement was highly statistically significant ($p < .0001$).

Discussion

These findings show that timbre (instrument identity) affects perception of emotion in music by normal participants. The interaction between perceived emotion and instrument identity was not dependent on the level of “correct” (intended) performance. Since the melodies were novel, the effect cannot be attributed to previously learned associations between particular instruments and melodies. A similar pattern of findings was observed in both young and older adults and after controlling for factors relevant to musical experience (years of musical training and current estimated listening exposure). The effect also was observed in a subset of participants with less than one year of formal musical training on an instrument. These observations affirm that the effect was not determined by prior musical expertise. While our findings support and extend previous work suggesting that timbre can influence judgements of emotion in music (Balkwill & Thompson, 1999; Behrens & Green, 1993; Gabrielsson & Juslin, 1996; Juslin, 1997), the present study did not replicate certain specific instrument emotion interactions observed in other studies. For example, one earlier study (Behrens & Green, 1993) found that young

normal participants are more accurate at recognizing sadness and fear in improvisations performed on the violin than those performed on the trumpet. These apparent discrepancies are likely to be at least partly attributable to methodological differences between the studies; in the study of Behrens and Green (1993), musical structural characteristics (melody, pitch range, tempo, dynamics, and amplitude level) and overall performance accuracy across the emotions was not controlled as performers were allowed to freely improvise the melodic content. Moreover, different instruments were used, and responses were based on a Likert scale rather than a forced-choice categorization task. The specific effects of particular instruments in modulating emotion perception are not easy to interpret due to the multi-dimensional timbral “space” that they occupy and the difficulty of predicting a priori the affective properties of particular timbral components embedded in real instrument tones. We found that the violin reduced the odds of identification of “happy” melodies, and the electronic synthesizer reduced the odds of identification of “sad” melodies. Differences between the synthesizer and the other instruments might in principle be due to expressive cues added by the human performers on the natural instruments, as such factors have been shown previously to contribute to emotion judgements (Gabrielsson & Juslin, 1996; Juslin, 2000, 2001). This issue is addressed explicitly in Experiment 2.

While there was general agreement among participants on the emotion conveyed by particular melodies, older adults were overall less likely than young adults to identify the target emotion label for a particular melody. The overall difference between the age groups was not eliminated by adjusting for years of musical training or musical listening hours per week, and is unlikely to be attributable to the style of music in which the melodies were composed as a greater proportion of older adults reported regularly listening to classical music. This suggests that there may be a systematic reduction in ability to recognize emotions in music over the course of the lifespan, which is broadly consistent with findings concerning the

effects of ageing on other modalities of emotion recognition (Calder et al., 2003; Fischer et al., 2005). However, the interaction between age and emotion identification in music (such that older adults were relatively worse at recognizing anger and to some extent happiness) was attributable entirely to the effect of duration of formal musical training. This implies that musical training may facilitate identification of specific features, perhaps including mode, key, articulation, or accent (Juslin, 2000), that are important for the identification of particular emotions in music, and it is consistent with evidence that accuracy at identifying emotions in music can improve with expertise (Balkwill & Thompson, 1999; Juslin, 2000; Sloboda, 1991).

EXPERIMENT 2

Having established that the perception of emotion in melodies is influenced by timbre for familiar instruments, we set out to replicate these findings in a new group of participants using unfamiliar synthetic timbres. Our objective in this experiment was two-fold. Firstly, we wished to create novel timbres that incorporated timbral cues predicted to convey particular musical emotions based on previous work (Gabrielsson & Juslin, 1996; Juslin, 1997; Sloboda, 1991; Sloboda & O'Neill, 2001b). We sought to make these novel timbres complex and perceptually distinct while avoiding close similarities with real musical instruments, thereby minimizing the effects of learned emotional associations with particular instruments. Secondly, synthetic stimuli remove effects (such as fine manipulations of dynamics or tempo) that may be added by human performers and which may modulate emotional impact.

Method

Participants

A total of 18 healthy young adults (13 females) aged 19–28 years (mean age 24.3 years, $SD = 2.6$), none of whom had participated in Experiment 1, were recruited. None of the

participants was a professional musician. Participant characteristics (available on request) were comparable to those of the young participant group in Experiment 1.

Stimuli

Four novel timbres were created digitally using Matlab©. Timbres were defined by specifying harmonic content, spectral envelope shape, and the temporal envelope parameters of “attack” and temporal modulation rate and depth (vibrato), using the procedure described by Warren, Jennings, and Griffiths (2005). Characteristics of the timbres are summarized in Table 3. Timbral characteristics have been associated with particular musical emotions (Gabrielsson & Juslin, 1996; Sloboda & Juslin, 2001), and we sought to manipulate the key properties of spectral composition (spectral envelope), attack, and vibrato (temporal envelope) when designing our novel timbres. Four synthetic instrument sounds were created with the intention of creating perceptually distinctive timbres by systematically varying these timbral parameters. There is a potentially vast space of spectrotemporal parameters that can be used to define a timbre, and real musical instruments share constituent timbral parameters despite being perceptually distinctive (Caclin et al., 2005; McAdams, 1993; McAdams & Cunible, 1992): The perceptual effect of a novel synthetic timbre is therefore difficult to predict with precision. Here we used combinations of timbral features that have been shown to differentiate between expressed emotions in music: The novel timbres 1 and 3 (Table 3) incorporated properties including high

energy at lower harmonics (conveying a “dull” quality), slower attack, and slow vibrato, which have been associated with sadness, while Timbres 2 and 4 incorporated properties including higher energy at higher harmonics, “notched” spectra (conveying a bright or “brash” quality), and more rapid attack, which have been associated with anger or happiness, and rapid vibrato, which has been associated with anger and fear.

The four synthetic timbres were used to write digital wavefiles corresponding to each of the melodies used in Experiment 1. The resulting stimuli were identical in pitch (f_0) values, rhythm, tempo, and average intensity level to the stimuli used in Experiment 1 (examples available on request).

Experimental procedure

Using analogous methods to those in Experiment 1, each participant was presented with 160 stimuli in four “synthetic timbre” blocks. The order in which melodies were presented within each block (the period) followed the same principles as those in Experiment 1. However as 18 participants participated in Experiment 2, only 18 of the 24 possible orderings of synthetic timbres were used. These were chosen so as to ensure as close a balance as possible on instruments and periods.

Statistical analysis

The data was analysed using the basic model described under Experiment 1.

In addition, we wished to establish that participant responses to the synthetic timbres were not driven by similarity between the synthetic timbres and real instruments. This was assessed

Table 3. *Characteristics of novel timbres for Experiment 2*

| <i>Timbral property</i> | <i>Timbre</i> | | | |
|-------------------------|------------------------------------|--------------------------|--------------------------|---|
| | <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> |
| Spectral content | Strong middle and low frequencies. | Strong high frequencies. | Strong low frequencies. | Few harmonics, “notched” spectral envelope. |
| Temporal envelope | Slow attack and decay. | Fast attack, slow decay. | Slow attack, fast decay. | Fast attack and decay. |
| Vibrato rate and depth | None. | Fast and mid amplitude. | Slow and low amplitude. | Fast and high amplitude. |

following the main experiment based on judgements made by 11 musically experienced participants (with five or more years of musical training). These participants were presented with each of the four novel timbres playing a “happy” melody and were asked to rate on a scale of 0 – 4 how strongly each synthetic timbre resembled a piano, violin, trumpet (target instruments from Experiment 1), clarinet (a widely used wind instrument), or any other real instrument nominated by the participant.

Results

The mean number of “correct” responses (score / 10) for each target emotion are displayed in Figure 3. The odds ratios for an intended response for each combination of synthesized timbre and emotion (adjusted for order of instrument presentation), relative to the odds of an intended response for “happy” melodies played on Timbre 1 (as explained in the Method section) are shown in Table 4. Results were similar to Experiment 1: Averaged over all emotion categories, the odds of a correct response were similar for each timbre, and a global test of differences was not statistically significant ($p = .7$). However, the odds of an intended response differed markedly and were statistically significantly ($p < .0001$) between emotions. There was a statistically significant interaction ($p < .05$) between timbre and emotion judgement: Anger was

better identified using Timbres 2 and 4 than using Timbres 1 and 3, while, conversely, sadness was better identified using Timbres 1 and 3 than using Timbres 2 and 4.

Considered as a group, the novel timbres were not rated by 11 musically experienced participants as strongly resembling real instruments. For the experimenter-nominated instruments (piano, violin, trumpet, clarinet), none of the timbres had a mean similarity rating greater than 1.9 (range 0–4). For three of the timbres (1, 2, and 4), only 1 or 2 participants rated that timbre as similar to another real instrument. A total of 5 participants rated Timbre 3 as somewhat similar to a French horn (mean rating for those participants was 2.6).

Discussion

The findings of Experiment 2 support and extend those of Experiment 1: The previously observed interaction between timbre and perceived emotion in music continues to operate when the timbres are novel and in the absence of additional performer-dependent expressive cues. It is unlikely that the interaction here was driven by learned emotional associations with particular timbres, as the novel synthetic timbres considered as a group did not strongly resemble real instruments. The pattern of the interaction is broadly in keeping with empirical predictions concerning the timbral properties relevant to particular musical emotions from previous work (Gabrielsson & Juslin, 1996; Juslin, 1997; Sloboda & O’Neill, 2001). Previous studies have shown considerable variability in the relative importance attributed to particular timbral features in conveying particular emotions in music, and it would therefore be premature to draw specific conclusions about the role of individual timbral parameters. Despite this caveat, it is likely that timbral properties convey affective valences that are relevant to both real and synthetic instruments (Juslin, 1997, 2001; Juslin & Laukka, 2001, 2003) and to at least some basic emotions. It is likely that the relative importance of this timbral effect varies for different emotions: Fear, for example, may be difficult to convey using purely

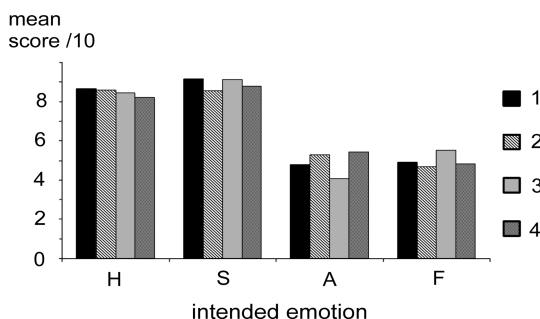


Figure 3. Mean scores (/10) for each intended emotion for synthesized timbres 1–4, in young adult participants. *H*, happiness; *S*, sadness; *A*, anger; *F*, fear.

Table 4. Odds of a "correct" response relative to "happy" melodies on *Timbre 1* for young adults in *Experiment 2*

| <i>Novel timbre</i> | <i>Intended emotion</i> | | | | | | | |
|---------------------|-------------------------|--------------|----------------|--------------|--------------|--------------|-------------|--------------|
| | <i>Happiness</i> | | <i>Sadness</i> | | <i>Anger</i> | | <i>Fear</i> | |
| | <i>OR</i> | <i>CI</i> | <i>OR</i> | <i>CI</i> | <i>OR</i> | <i>CI</i> | <i>OR</i> | <i>CI</i> |
| 1 | 1 | | 1.70 | (0.86, 3.36) | 0.14 | (0.08, 0.23) | 0.14 | (0.08, 0.24) |
| 2 | 0.96 | (0.52, 1.75) | 0.91 | (0.50, 1.67) | 0.17 | (0.10, 0.28) | 0.13 | (0.08, 0.22) |
| 3 | 0.83 | (0.46, 1.51) | 1.58 | (0.80, 3.10) | 0.10 | (0.06, 0.17) | 0.18 | (0.11, 0.31) |
| 4 | 0.71 | (0.4, 1.26) | 1.10 | (0.59, 2.05) | 0.18 | (0.11, 0.30) | 0.14 | (0.08, 0.24) |

Note: "Correct" = intended response; OR = odds ratio of an intended response; CI = 95% confidence intervals.

timbral cues without recourse to tempo and dynamic variations (Gabrielsson & Juslin, 1996; Reniscover, Salovey, & Repp, 2004; Robazza et al., 1994; Sloboda & O'Neill, 2001; Terwogt & Van Grinsven, 1988).

GENERAL DISCUSSION

Taken together, the two experiments described here show that the perception of emotion conveyed by a melody is affected by the identity or timbre of the musical instrument on which it is played. We observed an interaction between performance on an emotion categorization task and instrumental timbre in Experiment 1 using familiar timbres and in Experiment 2 using synthesized timbres. This interaction was observed after controlling for potential confounding musical structural factors (the specific sequence of pitches comprising the melody, loudness, tempo, harmonic cues), cognitive factors (familiarity of the melody and instrument, order of presentation of instruments), and performance factors (human instrumentalists versus synthetic instruments). This evidence suggests that the interaction between perceived emotion and timbre does not depend simply on lower level acoustic or structural factors nor on the individual's past experience of music, but is a robust effect driven directly by changes in timbral characteristics such as attack and spectral content. This is consistent with previous evidence on the effects of timbral changes on emotion judgements in cross-cultural musical

contexts (Balkwill & Thompson, 1999; Balkwill, Thompson, & Matsunaga, 2004).

The interaction between perceived emotion and timbre was consistently observed in both experiments in the context of a main effect of emotion recognition for different emotions across instruments. Sadness and happiness were well recognized, whereas anger and fear were less well recognized. It is likely that aspects of timbre interact with other prominent cues for emotion (e.g., melody and harmony). Under most circumstances music is polyphonic (containing multiple musical parts): It is likely that more complex timbral effects operate at the level of instrument combinations and may facilitate emotion recognition (Heaton, Hermelin, & Pring, 1999). These effects may be particularly relevant to the perception of anger and fear (Juslin, 2001). In addition, anger and fear are both associated with minor harmonies and irregular rhythms, and it is likely that such shared cues contributed to reciprocal confusions between these emotions here and in previous studies (Reniscover et al., 2004; Robazza et al., 1994; Terwogt & Van Grinsven, 1988). Furthermore, in everyday music listening fear is the least frequently represented of the four target emotions (Gabrielsson, 2001), and it is perhaps even less frequently conveyed in works for unaccompanied instruments (Reniscover et al., 2004). Despite these caveats in regard to particular emotions in music, the use of synthetic timbres in Experiment 2 demonstrates that timbral properties can be specifically manipulated in line with empirical and theoretical predictions about their

emotional effects (Gabrielsson & Juslin, 1996; Juslin, 1997; Sloboda & O'Neill, 2001) in order to bias listeners' judgements about the emotional content of a particular piece of music. In light of the very large space of timbral parameters, this is clearly an area ripe for further systematic study.

Based on neuropsychological evidence of doubly dissociated deficits for identification of perceptual features and emotion in music (Gosselin et al., 2007; Peretz et al., 1998), it has been proposed that affective and perceptual characteristics of music are processed at least partly in parallel. The present findings show that there is an important interaction between these perceptual and affective dimensions in normal listeners. Furthermore, this interaction occurs at the level of an object-level property (timbre), rather than more elementary acoustic properties such as tempo or intensity, which have generally been emphasized in previous work (Balkwill & Thompson, 1999; Balkwill et al., 2004; Dalla Bella et al., 2001; Gabrielsson & Juslin, 1996; Juslin, 1997). This suggests that the interaction between perceptual and affective information may be hierarchical: The encoding of properties such as tempo, intensity, mode, and dissonance may serve to distinguish positive and negative emotional states in music (Dalla Bella et al., 2001; Juslin, 1997; Peretz et al., 1998) while more fine-grained affective discriminations may depend on higher level (object-level) properties such as timbre (Juslin, 2001). This part serial, part parallel organization of early perceptual analysis followed by interacting object identification and affective processing mechanisms would be consistent with current models of auditory object processing (Griffiths & Warren, 2002, 2004) and with emerging evidence from functional brain imaging work in normal participants (Blood & Zatorre, 2001; Koelsch, Fritz, Von Cramon, Muller, & Friederici, 2006; Menon & Levitin, 2005; Olson, Plotzker, & Ezzyat, 2007).

The influence of instrument timbre (a key auditory object property) on emotional valence may have wider relevance to the processing of other types of complex sounds in different contexts. It has been argued that the brain mechanisms that process timbre in music evolved for the

representation and evaluation of vocal sounds (Juslin & Laukka, 2003). There are several candidate theoretical mechanisms by which timbral properties might acquire musical affective connotations. Musical timbres might, for example, share acoustic characteristics with vocal emotional expressions (Juslin & Laukka, 2001, 2003). Timbral features in music could represent expressive contours (via temporal characteristics such as attack) or qualities (via spectral properties such as low or high frequency content): For example, a "dull" spectral quality is associated with sadness in music, whereas the "brash" quality conferred by prominent high frequencies is associated with anger (Juslin, 2001). These may generalize to the expression of emotion via other structural cues in music, the expression of vocal emotion, or expression of emotion in other modalities such as gesture (Sloboda & Juslin, 2001). Interpretation of musical cues may be facilitated by or contribute to "emotional intelligence" in nonmusical contexts (Reniscover et al., 2004). Abnormal coupling between timbre and affective processing mechanisms might contribute to altered emotional reactions to music and environmental sounds observed in conditions such as Williams syndrome (Levitin, Cole, Lincoln, & Bellugi, 2005), temporal lobe epilepsy (Boeve & Geda, 2001; Rohrer, Smith, & Warren, 2006), and frontotemporal lobar degenerations (Boeve & Geda, 2001).

Original manuscript received 26 June 2007
Accepted revision received 18 December 2008
First published online 17 April 2009

REFERENCES

- Balkwill, L. L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*, 43–64.
- Balkwill, L. L., Thompson, W. F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research, 46*, 337–349.
- Behrens, G. A., & Green, S. B. (1993). The ability to identify emotional content of solo improvisations

- performed vocally and on three different instruments. *Psychology of Music*, 21, 20–33.
- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, USA*, 98, 11818–11823.
- Boeve, B. F., & Geda, Y. E. (2001). Polka music and semantic dementia. *Neurology*, 57, 1485.
- Bregman, A. S., Levitan, R., & Liao, C. (1990). Fusion of auditory components—effects of the frequency of amplitude-modulation. *Perception & Psychophysics*, 47, 68–73.
- Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, 118, 471–482.
- Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., et al. (2003). Facial expression recognition across the adult life span. *Neuropsychologia*, 41, 195–202.
- Copland, A. (1957). *What to listen for in music* (Rev. ed.). New York: Signet Classics.
- Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, 80, B1–B10.
- Dowling, W. J., Bartlett, J. C., Halpern, A. R., & Andrews, W. M. (2008). Melody recognition at fast and slow tempos: Effects of age, experience, and familiarity. *Perception & Psychophysics*, 70, 496–502.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6, 169–200.
- Fischer, H., Sandblom, J., Gavazzeni, J., Fransson, P., Wright, C. I., & Backman, L. (2005). Age-differential patterns of brain activation during perception of angry faces. *Neuroscience Letters*, 386, 99–104.
- Gabrielsson, A. (2001). Emotions in strong experiences with music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 431–449). Oxford, UK: Oxford University Press.
- Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, 24, 68–91.
- Gosselin, N., Peretz, I., Johnsen, E., & Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia*, 45, 236–244.
- Goydke, K. N., Altenmuller, E., Moller, J., & Munte, T. F. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cognitive Brain Research*, 21, 351–359.
- Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, 25, 348–353.
- Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews. Neuroscience*, 5, 887–892.
- Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine*, 29, 1405–1410.
- Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100, 33–72.
- Jones, B., & Kenward, M. G. (2003). *Design and analysis of cross-over trials* (2nd ed.). London: Chapman & Hall/CRC Press.
- Juslin, P. N. (1997). Perceived emotional expression in synthesised performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae*, 2, 225–256.
- Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1797–1813.
- Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 309–337). Oxford, UK: Oxford University Press.
- Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, 1, 381–412.
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129, 770–814.
- Koelsch, S., Fritz, T., Von Cramon, D. Y., Muller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, 27, 239–250.
- Kohlmetz, C., Muller, S. V., Nager, W., Munte, T. F., & Altenmuller, E. (2003). Selective loss of timbre perception for keyboard and percussion instruments following a right temporal lesion. *Neurocase*, 9, 86–93.
- Levitin, D. J., Cole, K., Lincoln, A., & Bellugi, U. (2005). Aversion, awareness, and attraction: Investigating claims of hyperacusis in the Williams

- syndrome phenotype. *Journal of Child Psychology and Psychiatry*, 46, 514–523.
- McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146–198). Oxford, UK: Oxford University Press.
- McAdams, S., & Cunible, J. C. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 336, 383–389.
- Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *NeuroImage*, 28, 175–184.
- Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: A review of findings on social and emotional processing. *Brain*, 129, 2945–2956.
- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68, 111–141.
- Reniscow, J., Salovey, P., & Repp, B. (2004). Is recognition of emotion in music performance an aspect of emotional intelligence? *Music Perception*, 22, 145–158.
- Robazza, C., Macaluso, C., & Durso, V. (1994). Emotional reactions to music by gender, age, and expertise. *Perceptual and Motor Skills*, 79, 939–944.
- Rohrer, J. D., Smith, S. J., & Warren, J. D. (2006). Craving for music after treatment for partial epilepsy. *Epilepsia*, 47, 939–940.
- Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, 32, 863–881.
- Schneider, A. (2001). Sound, pitch, and scale: From “tone measurements” to sonological analysis in ethnomusicology. *Ethnomusicology*, 45, 489–519.
- Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, 19, 110–120.
- Sloboda, J. A., & Juslin, P. N. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 71–104). Oxford, UK: Oxford University Press.
- Sloboda, J. A., & O’Neill, S. A. (2001). Emotions in everyday listening to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 415–430). Oxford, UK: Oxford University Press.
- Stewart, L., von Kriegstein, K., Warren, J. D., & Griffiths, T. D. (2006). Music and the brain: Disorders of musical listening. *Brain*, 129, 2533–2553.
- Terwogt, M. M., & Van Grinsven, F. (1988). Recognition of emotions in music by children and adults. *Perceptual and Motor Skills*, 67, 697–698.
- Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer.
- Warren, J. D., Jennings, A. R., & Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *NeuroImage*, 24, 1052–1057.