

Beyond Happiness and Sadness: Affective Associations of Lyrics with Modality and Dynamics

LAURA TIEMANN

Ohio State University, School of Music

DAVID HURON[1]

Ohio State University, School of Music

ABSTRACT: A study is reported investigating the relationship between modality (major/minor) and dynamics (piano/forte) on four affects – as evident in the content of musical lyrics. Forty solo vocal works were sampled: 10 in the major mode with a loud (forte) dynamic level, 10 in the major mode with a quiet (piano) dynamic level, 10 in the minor mode with a loud dynamic level, and 10 in the minor mode with a quiet dynamic level. Sampled compositions were all tonal works from the Western vocal repertoire. Without hearing the music, 60 native-speakers of English, German, and French judged the language-appropriate lyrics according to four affects: *sadness*, *happiness*, *passion*, and *tenderness*. Results were consistent with predicted associations between minor-piano music and *sadness*, major-forte music and *happiness*, and minor-forte music and *passion*. A fourth predicted association between major-piano music and *tenderness* was skewed in the predicted direction, but was not statistically significant.

Submitted 2011 April 12; accepted 2011 July 17.

KEYWORDS: *emotion, lyrics, vocal music, happy, sad, tender, passionate*

SINCE at least the sixteenth century, there has been an assumption in Western music linking the minor mode with sadness (e.g., Zarlino, 1558). In addition to informal observations, formal empirical evidence in support of this view was provided by Heinlein (1928) and Hevner (1935). Hevner, for example, arranged short musical passages transposed into both the major and minor modes and asked listeners to judge the affective content. As well as “sadness,” she found that listeners selected a number of other affective descriptors, including pathetic, melancholic, plaintive, yearning, mysterious, weird, dark, dreamy, hushed, serious, and spiritual.

Valentine (1913-1914) suggested that the purported sadness of the minor mode might arise from learned associations. That is, those sounds or sound patterns heard in sad contexts (such as memorial services) or accompanied by sad lyrics, would, after sufficient exposure, tend to evoke sad connotations on their own as a conditioned response. Consistent with this interpretation, many ethnomusicologists have noted that the minor scale does not have sad connotations in a number of non-Western cultures. It is widely assumed that the affective content associated with the minor mode arises due to enculturation. Research in developmental psychology also converges with this interpretation. For example, studies by Crowder and his colleagues imply that, while the association of the minor mode with sadness is evident in both adults and children 3 years of age, there is little evidence that the major/minor distinction is salient for 6-month old children (Crowder, 1985; Kastner & Crowder, 1990; Crowder, Reznick & Rosenkrantz, 1991).

More recent experimental research suggests that there may be some common principles that contribute to the perception of sadness, such as principles associated with sad speech prosody (e.g., Juslin & Laukka, 2003; Patel, 2008). In prosodic research, a sad voice is associated with a relatively lower pitch. Huron, Yim and Chordia (2010), for example, have shown that artificial scales are judged as sounding sadder when random tones are lowered compared with a normative exposure. Such research suggests that while there may be nothing inherently sad about the minor mode, it would be difficult to establish the reverse associations (i.e., with the minor mode linked to happiness and the major mode linked with sadness) due to the lowering of pitches in the minor mode compared with the major mode.

In contrast to the evidence suggesting that lower-than-normal pitch is associated with sadness, Huron, Kinney and Precoda (2008) found that low transpositions of melodies caused them to be heard as more aggressive, more dominant, and heavier. In short, in some circumstances lower pitch seems to be associated with sadness, whereas in other circumstances lower pitch seems to be associated with aggression. As an informal observation, loudness seems to be implicated as a mediating factor. That is, low and loud may be more likely to be heard as aggressive (rather than sad), while low and quiet may be more likely to be heard as sad (rather than aggressive).

This conjecture provides a possible interpretation for an anomaly reported in Post and Huron (2009). In that study, a test was carried out of the hypothesis that music in the minor mode is associated with slower tempo. Indeed, minor-mode music was found to be slower for samples of Baroque and Classical music. However, music from the early Romantic period was found to exhibit a reverse correlation. That is, in general, music from the Romantic period exhibited a faster tempo for works in the minor mode compared with the major mode.

Post and Huron speculated that this reverse observation may be related to the affective language commonly referred to as *Sturm und Drang*. This artistic movement originated in the late Classical period. Especially with regard to compositions from the mid to late eighteenth century, music scholars have described loud minor-mode musical passages as conveying “passion” rather than “sadness” (e.g., Brook, 1970; Eggebrecht, 1955). This affective language continued beyond the Classical period, and became symptomatic of the Romantic period.

Ladinig and Huron (2010) formally tested this interpretation by comparing the dynamic levels in keyboard works from the eighteenth and nineteenth centuries. In general, notated dynamic levels tend to become quieter over the period in question. Using the mean dynamic level for each period as a reference, music in the minor mode during the nineteenth century exhibited higher dynamic levels as contrasted with corresponding music in the major mode. These observations are consistent with the idea that Romantic music is more likely than music of the preceding Classical period to employ the minor mode to represent or convey *passion* rather than *sadness*.

These studies suggest that researchers pay attention to how modality interacts with other musical parameters such as tempo and dynamics. When an experiment forces listeners to choose between “happy” and “sad” responses, an unduly restricted affective palette may be imposed – especially regarding the use of the minor mode.

In carrying out research of this sort, there are both advantages and disadvantages to examining sound recordings versus notated scores. It is difficult to infer the dynamic level from sound recordings due to the standard practice of compressing the dynamic range. For example, one cannot determine whether a recorded passage is *forte* or *piano* merely by observing a volume units meter. Conversely, tempo is difficult to infer from notated scores. Firstly, tempo is not always specified in notated music. Furthermore, even when a work specifies a metronome marking (such as quarter-note equals 60 beats per minute) a listener may perceive an eighth-note tactus corresponding to 120 beats per minute. That is, even when the tempo is clearly specified in the score, it can be challenging for the researcher to infer that the music would be perceived as fast or slow. Consequently, it is difficult to study both the relationships of tempo and dynamics to modality when employing solely notated or solely recorded samples.

For the purpose of this study, we therefore chose not to take tempo into consideration. Instead, we focus on the relationship between mode and dynamic level. Specifically, we will examine four conditions: major/*forte*, major/*piano*, minor/*forte*, and minor/*piano*.

In studies of affect, psychologists have tended to concentrate on prototypical or so-called “basic” emotions such as anger, fear, sadness, disgust, and joy. Many of the early studies examining music-related emotions sought acoustic correlates with these presumed fundamental emotion categories. However, Scherer has suggested that such basic emotions may be less relevant for the arts (Scherer, 2004). For example, a follow-up study by Zenter, Grandjean and Scherer (2008), chronicled a large number of affective descriptions in response to musical passages, and distilled these descriptions to nine emotions that are more pertinent to music; these nine emotions differ from the common lists of presumed basic emotions. In particular, a commonly reported music-related affect not found in lists of basic emotions is what Zenter *et al.* have referred to as *tender-longing*.

In any research on music and emotion, a central problem is how one characterizes the presumed affective content of a passage or stimulus. Several approaches are possible. A common method asks listeners to judge the affective content of the sounded music and then relates these judgments to objective properties of the musical organization. Notice that this approach may inadvertently introduce various

demand characteristics. For example, hearing that the music is loud, a listener may think that “sad music is never loud” and therefore refrain from judging the passage as sad. That is, listener hypotheses about how music induces affect may influence their judgments of the conveyed or expressed affect. A useful way of circumventing such demand characteristics is to have participants judge the affective content of the lyrics alone. This approach will be used in this study. That is, we will have participants judge the emotional content of song lyrics without hearing the songs themselves. This method makes the assumption that both the music and the lyrics are intended to express or convey the same emotional content.

In general, research on emotion has tended to employ one of two different approaches. In an open-ended approach, the researcher invites participants to describe the affective content without pre-defined response categories. This has the advantage of reducing experimenter bias, and allows for descriptive terms that may more accurately capture the participant’s experience. However, such an open-ended approach raises difficulties in analyzing and summarizing the descriptive content. In a contrasting approach, the researcher provides a set of pre-defined categories or terms and asks participants to rate the stimulus according to the given labels. This has the advantage of facilitating analysis and allows more direct hypothesis testing. However, this approach can flounder when participants are forced to make judgments according to categories that they find ill-suited to the stimulus.

In the current study, we propose to test specific hypotheses and so will ask participants to make judgments using a limited set of categories. In order to minimize the problem of forcing participants to make judgments using categories they deem ill-suited to the stimuli, we will ask the participants to specifically judge the applicability of the pre-defined categories for each stimulus.

Hypotheses

In light of the above discussion, one can imagine an interaction between loudness and modality that might account for four affects that appear to be musically pertinent: *happiness*, *sadness*, *tenderness*, and *passion*. If music in the minor mode also exhibits the quiet dynamic level characteristic of sad speech, one might predict the music to convey or express sadness. By contrast, when linked with a loud dynamic level, music in the minor mode might be predicted to convey or express the passionate affects of *Sturm und Drang*. In the case of the major mode, a high dynamic level might be predicted to convey or express happiness, whereas when the major mode is linked with a low dynamic level, one might predict the music to convey or express tenderness. Accordingly, we propose to test the following four hypotheses:

- H1.** Minor and piano is associated with sadness,
- H2.** Minor and forte is associated with passion,
- H3.** Major and piano is associated with tenderness,
- H4.** Major and forte is associated with happiness.

METHOD

In brief, our method involved assessing the affective content of lyrics for songs selected according to the four combinations of mode and dynamics. Specifically, we administered a survey where participants were asked to judge the emotional character of printed lyrics for songs representing the four conditions of minor/piano, minor/forte, major/piano, and major/forte.

Musical Sample

Ideally, we want the results of this study to be representative for all vocal music created using the major/minor system. In order to facilitate the testing of our hypotheses, we therefore need to select music that is unambiguous with regard to mode, and unambiguous with regard to dynamics. Accordingly, we set out to sample vocal works that are clearly major/forte, minor/forte, major/piano, and minor/piano. In the end, we were able to identify 40 works (10 for each condition); unfortunately, the sample was limited by the difficulty of finding suitable works in the minor mode.

The character of many musical works changes over the length of the work. A work that begins with one dynamic level may end with a contrasting dynamic, and works sometimes change mode. Rather than attempting to accommodate such changes, we made the simplifying operationalization that the

beginning of a work is representative of the work as a whole. Hence, our sampling method focused solely on the initial dynamic and the initial mode of the individual works.

In determining the dynamics of a musical work, a simple operationalization is to identify the notated dynamic marking. This assumes that performers interpret the works accordingly. Our sampled scores included vocal works from different periods, different styles, and different languages. Specifically, our sample drew from the Baroque, Classical and Romantic periods, as well as Western popular music, and included lyrics in English, French, and German. A work was deemed to be *forte* if the initial dynamic marking was either *mf*, *f*, *ff*, or *fff*. Conversely, a work was deemed to be *piano* if the initial dynamic marking was either *mp*, *p*, *pp*, or *ppp*.

The mode of a musical work was determined by examining the score. Commonly, the prevailing key is established at the beginning of the work. The experimenters determined the key of the work by examining the key signature and the initial harmonic framework. For each work examined, the key was judged to be either (i) obviously major, (ii) obviously minor, or (iii) not obviously major or minor. We sampled only those works deemed to be in an unambiguously major or minor mode.

Assessment of Lyrics

In assessing the affective content of the lyrics, we aimed to avoid potential experimenter bias by recruiting independent judges who characterized the emotional content of printed lyrics without hearing any of the songs or seeing any notated scores. For each lyric, they were asked to rate the expressed *happiness*, *sadness*, *tenderness*, and *passion* on separate 7-point scales. In addition, they rated the *pertinence* for these same four affects on separate 5-point scales. For each song, judges were further asked whether or not the lyrics were familiar. In order to limit the amount of work involved, each judge rated the lyrics for just eight songs. Since the lyrics included English, French, and German materials, we recruited native speakers of these three languages. In assembling our sample of music, the available art-music scores tended to be dominated by materials in the German language, whereas the pop-music scores tended to be dominated by materials in the English language. In order to avoid excessively long questionnaires, we consequently created two different questionnaires in English and in German, with a single questionnaire pertaining to the French material. Hence, five independent questionnaires were distributed to five groups of respondents.

English native speakers were recruited from the Ohio State University School of Music undergraduate student subject pool. French speakers were enlisted from the Université de Montréal, and German speakers were recruited through informal contacts of the first author. In total, 44 English, 11 French, and 21 German respondents participated in the content assessment of the lyrics. Participants received one language-specific questionnaire each as an e-mail attachment. The English version of the instructions is given below:

INSTRUCTIONS

Thank you for agreeing to participate in our research on the emotional quality of lyrics.

Below you will see the lyrics for 8 songs. For each of the lyrics we would like you to indicate how

(a) happy, (b) sad, (c) tender, and (d) passionate their content appears to you. Please also make a judgment on how pertinent those four characteristics seem to be for describing the respective lyrical content.

We would like you to make your judgments on the basis of the ideas, story and topic conveyed by the lyrics, not on whether the lyrics make you feel a particular way. By way of example, you may judge a particular lyric to have a very happy topic, even though the lyrics don't make you feel happy at all. In this case you would identify the most joyful lyrics without regard to how they make you feel.

If the lyrics for any given song are familiar to you, please insert a "yes" on the line saying "I am familiar with this song."

For your orientation, each rating scale contains all numbers between 1 and 7, or 1 and 5, respectively. Please delete all the numbers but the one that reflects your judgment best. Please leave no question unanswered.

As noted, respondents were asked to indicate whether they were familiar with the songs. If the participant indicated that they were familiar with one or more of the lyrics, then the entire questionnaire was excluded

from the data analysis. After excluding 16 participants, a total of 60 questionnaires remained for data analysis.

RESULTS

The results for all five questionnaires are summarized in Figures 1-4. The figures show the participants' ratings according to the four musical conditions (minor/piano, minor/forte, major/piano, major/forte) for *happiness* (Figure 1), *sadness* (Figure 2), *tenderness* (Figure 3), and *passion* (Figure 4). Consistent with the four hypotheses, on average, *happiness* judgments were highest for the major/forte condition, *sadness* judgments were highest for the minor/piano condition, *tenderness* judgments were highest for the major/piano condition, and *passionate* judgments were highest for the minor/forte condition. However, not all of these associations were statistically significant.

In order to test for statistical significance, we carried out a repeated measures analysis of variance. In the case of *happiness*, the Huynh-Feldt correction for the degrees of freedom revealed an overall effect of the 2x2 conditions of minor/major/piano/forte, $F(2.80/165.26=15.88, p<.05)$. The Bonferroni post-hoc test showed that all conditions but minor/forte and major/piano, and minor/forte and major/forte received statistically different ratings for *happiness*. The major/forte conditions received the highest average happiness rating with a mean of 3.53.

Ratings for the expressed emotion *happiness*

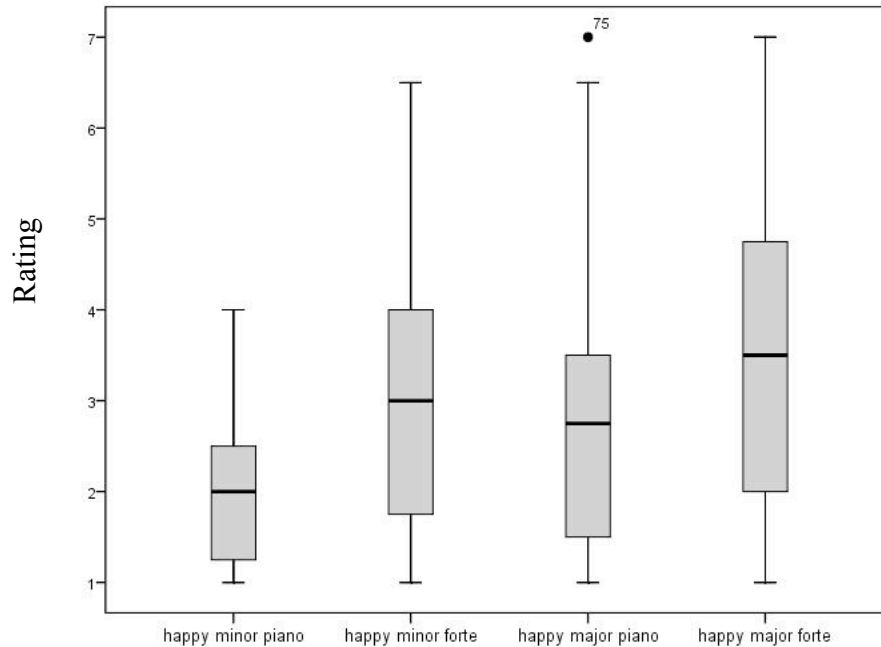


Fig. 1. Box plot illustrating ratings of *happiness* according to the four musical conditions. The range, 25th, 50th (median), 75th percentile, and an outlier are shown. The vertical axis represents increasing *happiness* judgments on a 7-point scale.

For *sadness*, an overall effect of the 2x2 conditions of minor/major/piano/forte, $F(3/177=29.73, p<.05)$ was found. The Bonferroni post-hoc test revealed a significant difference between all conditions, with the exception of minor/piano and major/piano, and minor/forte and major/piano. The minor/piano conditions received the highest average sadness rating with a mean of 5.00.

For *tenderness*, an overall effect of the 2x2 conditions of minor/major/piano/forte, $F(3/177=3.70, p<.05)$ was given. The Bonferroni post-hoc test showed a significant difference between the tenderness ratings for the major/piano and major/forte conditions; all other differences were not significant. The major/piano conditions received the highest average tenderness rating with a mean of 4.07.

Ratings for the expressed emotion *sadness*

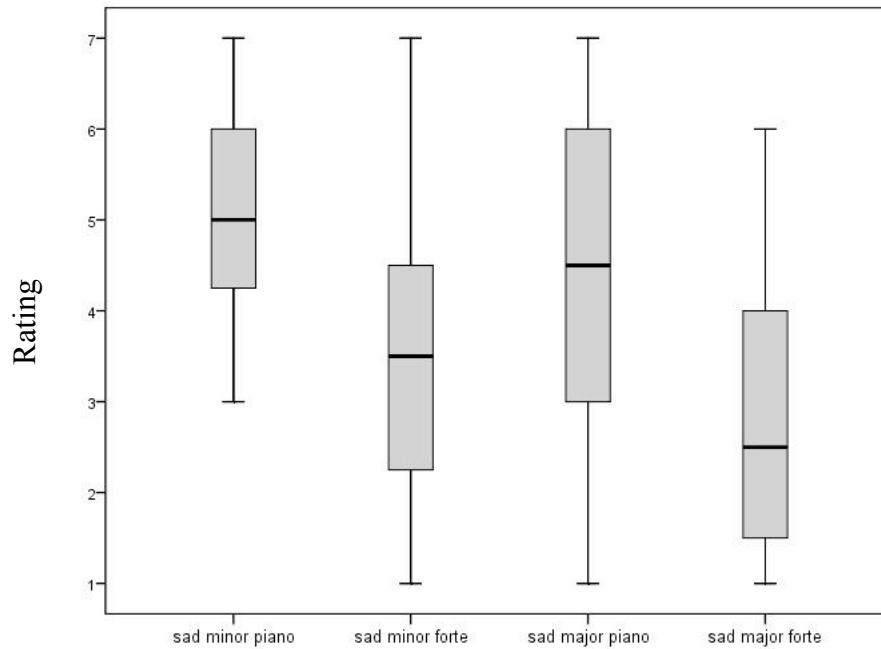


Fig. 2. Box plot illustrating ratings for *sadness* according to the four musical conditions. The vertical axis represents increasing *sadness* ratings on a 7-point scale.

Ratings for the expressed emotion *tenderness*

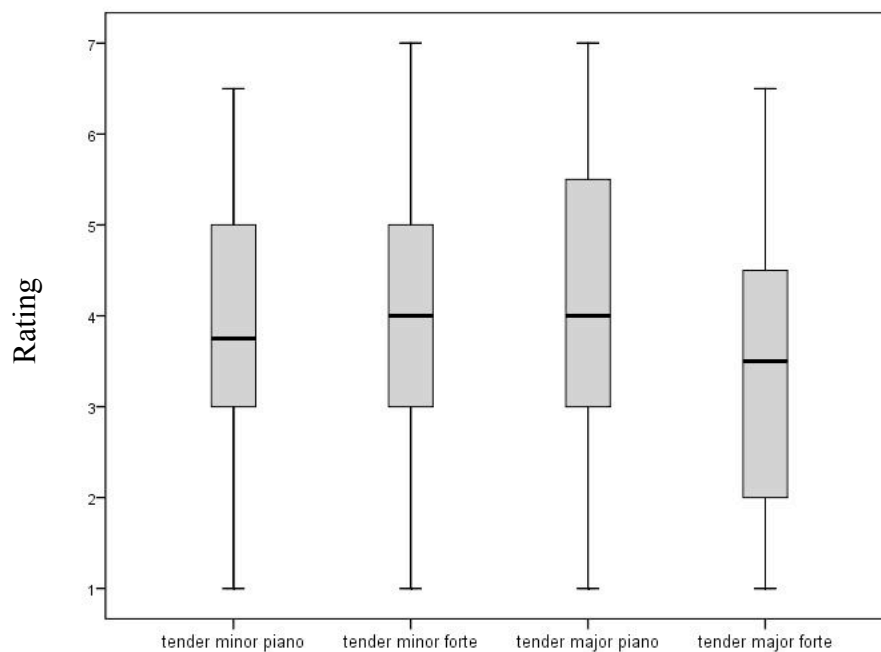


Fig. 3. Box plot illustrating ratings for *tenderness* according to the four musical conditions. The vertical axis represents increasing *tenderness* judgments according to a 7-point scale.

For *passion*, the Huynh-Feldt correction revealed no overall effect of the 2×2 conditions of minor/major/piano/forte, $F(2.75/162.08=15.88, p=.14)$. The Bonferroni post-hoc test likewise showed no significant differences between the conditions. The minor/forte conditions received the highest average *passion* rating with a mean of 5.08.

Ratings for the expressed emotion *passion*

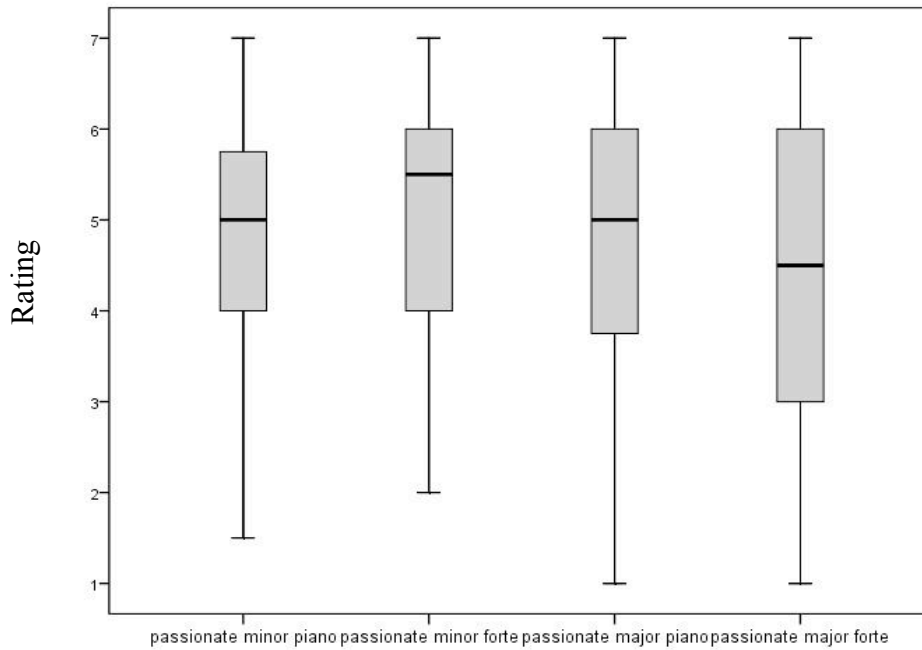


Fig. 4. Box plot illustrating ratings for *passion* according to the four musical conditions. The vertical axis represents increasing *passion* judgments on a 7-point scale.

DISCUSSION

The results are consistent with our first hypothesis that sad lyrics are most often accompanied by minor/piano music. However, although the minor/piano lyrics received the highest sadness ratings, these ratings were not significantly different from the sadness ratings for the major/piano lyrics. The findings are also consistent with the hypothesis that happy lyrical content is predominantly accompanied by major/forte music. Still, although the major/forte lyrics received the highest happiness ratings, these ratings were not significantly different from the happiness ratings for the minor/forte lyrics. Thus, for both *sadness* and *happiness* the chosen dynamic level seemed to underline the emotional content more strongly than the mode, even though the dynamic level was sampled only at the beginning of each work. For the remaining two emotions, *tenderness* and *passion*, the results were skewed in the predicted directions with major/piano receiving the highest *tenderness* ratings and minor/forte receiving the highest *passion* ratings. However, for these two emotions, only the *tenderness* ratings for major/piano and major/forte were significantly different. Considering that our results were all skewed in the predicted directions, the findings suggest that a follow-up study employing a larger sample of both songs and participants would be warranted; however, our considerable efforts to expand the sample were thwarted by the relative rarity of music in the minor mode. Furthermore, an inclusion of other musical parameters, such as tempo, might provide a more complete picture of the relationship between the expressed emotional content of lyrics and the type of musical setting employed to enhance the implied affective content.[2]

NOTES

[1] Address correspondence to David Huron, School of Music, 1866 College Road, Ohio State University, Columbus, Ohio, 43210, U.S.A.

[2] Our thanks to Kristin Precoda for providing French translations of the questionnaire instructions and to Isabelle Peretz for assisting in recruiting French-speaking respondents.

REFERENCES

- Brook, B. (1970). Sturm und Drang and the Romantic Period in Music. *Studies in Romanticism*, Vol. 9, No. 4, pp. 269-284.
- Crowder, R. (1985). Perception of the major/minor distinction: II. Experimental investigations. *Psychomusicology*, Vol. 5, pp. 3-24.
- Crowder, R., Reznick, J. S., & Rosenkrantz, S. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society*, Vol. 29, No. 3, pp. 187-188.
- Eggebrecht, H. H. (1955). Das Ausdrucks-Prinzip im musikalischen Sturm und Drang, *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte*, Vol. 29, pp. 323-349.
- Heinlein, C. P. (1928). The affective characteristics of the major and minor modes in music. *Journal of Comparative Psychology*, Vol. 8, No. 2, pp. 101-142.
- Hevner, K. (1935). The affective character of the major and minor modes in music. *American Journal of Psychology*, Vol. 47, pp. 103-118.
- Huron, D., Kinney, D., & Precoda, K. (2006). Influence of pitch height on the perception of submissiveness and threat in musical passages. *Empirical Musicology Review*, Vol. 1, No. 3, pp. 170-177.
- Huron, D., Yim, G., & Chordia, P. (2010). The effect of pitch exposure on sadness judgments: An association between sadness and lower than normal pitch. In S.M. Demorest, S.J. Morrison, P.S. Campbell (editors), *Proceedings of the 11th International Conference on Music Perception and Cognition*. Seattle, Washington: Causal Productions, pp. 63-66.
- Juslin, P.N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, Vol. 129, No. 5, pp. 770-814.
- Kastner, M., & Crowder, R. (1990). Perception of the major/minor distinction IV: Emotional connotations in young children. *Music Perception*. Vol. 8, No. 2, pp. 189-201.
- Ladinig, O., & Huron, D. (2010). Dynamic levels in Classical and Romantic keyboard music: Effect of musical mode. *Empirical Musicology Review*, Vol. 5, No. 2, pp. 51-56.
- Patel, A.D. (2008). *Music, Language, and the Brain*. Oxford: Oxford University Press.
- Post, O. & Huron, D. (2009). Western classical music in the minor mode is slower (except in the Romantic period). *Empirical Musicology Review*, Vol. 4, No. 1, pp. 1-9.
- Scherer, K. R. (2004). Which emotions can be induced by music? *Journal of New Music Research*, Vol. 33, No. 3, pp. 239-251.
- Valentine, C.W. (1913/1914). The aesthetic appreciation of musical intervals among school children and adults. *British Journal of Psychiatry*, Vol. 6, pp. 190-216.
- Zarlino, G. (1558). *Le istituzioni harmoniche*.
- Zentner, M., Grandjean, D., & Scherer, K.R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, Vol. 8, No. 4, pp. 494-521.