The effect of "lower than normal" pitch collections on sadness judgments of unconventionally-tuned melodies

GARY YIM [1]
School of Music, Ohio State University

DAVID HURON
School of Music, Ohio State University

PARAG CHORDIA
Department of Music, Georgia Institute of Technology

ABSTRACT: An experiment was conducted to investigate whether listeners perceived melodies using pitch collections with lowered pitches as sadder. Participants heard melodies in an experiment consisting of two phases: an exposure phase followed by a test phase. In the test phase, participant pairs independently judged the sadness of identical sets of melodies, which were constructed using artificial scales with unconventional tuning. During the prior exposure phase, participant pairs had heard different melodies, constructed using scales that were either raised or lowered versions of the test-phase scales. A between-subjects analysis revealed that test-phase melodies using the lowered version of the exposure-phase scales were judged as sadder. The results are suggestive of an explanation for the sadness of the minor mode, which may be perceived as sad because it is heard as a lowered version of the more prevalent major mode. This explanation is consistent with the observation that the minor mode or similar scales have no sad connotations for listeners not enculturated to a normative major mode. A related study using North Indian Classical Music is also discussed, where listeners judged melodies using *raags* with lowered pitches as more sad (Chordia & Sastry, 2011).

Submitted 2015 November 30.

KEYWORDS: *sadness, low pitch, minor scale, emotion, enculturation, prosody*

INTRODUCTION

THE expression and elicitation of sadness through music has increasingly been an area of interest for music and emotion researchers, exploring such topics as its paradoxical enjoyment (Schubert, 1996; Huron, 2011; Garrido & Schubert, 2011; Vuoskoski et al., 2012; Taruffi & Koelsch, 2014; Weth & Parncutt, 2014; Sachs, Damasio & Habibi, 2015), its role in mood regulation (Saarikallio & Erkkilä, 2007; Van den Tol & Edwards, 2013; 2014) and whether it is truly felt and experienced by listeners (e.g., Vuoskoski & Eerola, 2012). The features or expressive cues associated with sad music have been extensively studied as well: in one review (Juslin & Laukka, 2004) twenty-two cues of sad music are listed, including slow tempo, low sound level, narrow pitch range, dissonance, legato articulation and dull timbre.

Arguably, one of the most reliable cues for sadness in music is the use of the minor mode, at least for Western-enculturated listeners. Numerous studies have shown that in contrast to the major-mode, which tends to be associated with happiness or some other positive affect, the minor mode is strongly associated with sadness and negative affect (e.g., Heinlein, 1928; Hevner, 1935; Crowder 1984; Costa & Fine, 2004; Gomez & Danuser, 2007; Hunter, Schellenberg & Schimmack, 2010). Additionally, the minor mode is correlated with the other commonly identified cues for sadness: minor-mode music tends to be slower (Post & Huron, 2009), quieter (Turner & Huron, 2008), lower in both pitch and pitch variability (Huron, 2008), and tends to be characterized acoustically by less high-frequency energy, which is associated with a darker timbre (Shutz, Huron, Keeton & Loewer, 2008; Huron, Anderson & Shanahan, 2014).

Sad music has many of the same expressive cues as sad speech (Hargreaves, Starkweather, & Blacker, 1965; Williams & Stevens, 1972; Sobin & Alpert, 1999; Juslin & Laukka, 2003; Scherer, Johnstone & Klasmeyer, 2003; Bowling et al., 2010), which is consistent with theories that emotional expression in music may be based on emotional expression in speech (Roederer, 1984; Brown, 2000; Juslin & Laukka, 2003; Mithen, 2009; Huron, 2015). But the minor mode does not seem to have any obvious analogue in speech. This raises a question: if not by a speech relationship, why does the minor mode express or elicit sadness so strongly?

An early explanation for this sad-minor association was offered by Helmholtz (1863/1912), that referred to the harmonic series and essentially attributed the perceived negative affect of minor triads to sensory dissonance (see also Plomp & Levelt, 1964; Crowder, 1984). Other explanations for the sadness of the minor mode depict it as a learned cultural phenomenon (Valentine, 1913; Matthews & Pierce, 1980), which is consistent with the observation that minor-mode music is not sad in all cultures (e.g., Rice, 2004). Research on the development of affect discrimination in infants and children is also consistent with the cultural learning hypothesis (e.g., Kastner & Crowder 1990; Crowder et al., 1991; Nawrot, 2003).

The current investigation agrees with the learning account, and returns to a relationship with expressive speech for an explanation as to how the sad-minor association may be learned. Whereas other recent work considers this relationship with speech via spectral similarities (Bowling et al, 2010) or fundamental intervals (Curtis & Bharucha, 2010), the account provided here considers it via similarities in how lowered pitch may be perceived. (See also Cook, 2007 for another proposed explanation relating pitch to affect via ethology.)

Many speech prosody researchers have shown that lower pitch usually connotes sadness in speech (Skinner, 1935; Fairbanks & Pronovost, 1939; Eldred & Price, 1958; Davitz, 1964; Huttar, 1968; Williams & Stevens, 1972; Bergmann, Goldbeck & Scherer, 1988; Banse & Scherer, 1996; Sobin & Alpert, 1999; Breitenstein, van Lancker & Daum, 2001). The speech research notwithstanding, we can observe that most men do not sound sadder than most women, despite their generally lower-pitched voices. We might suppose, then, that a voice may be heard as sounding sad when its pitch is lower than what is "normally" expected. Ladefoged and Broadbent (1957) have demonstrated that listeners use a process of pitch normalization to infer the natural tessitura or compass of the voice, and that they can use these cues in the perception of intonation. Thus, in addition to absolute pitch height, listeners may also use relative differences (compared to an expectation or norm) in pitch height for affective judgments, where lower than normal pitches may be judged as sadder.

Similarly in music, at least two methods for lowering pitch may be considered. The first, a simple downward (chromatic) transposition, is indeed an adequate cue for negative affect (e.g., Hevner, 1937; Huron, et al., 2006). But such a method of lowering pitch does not relate to the minor mode, since it does not affect intervallic relationships. But there is another way of lowering pitch that does change intervallic relationships, by lowering a subset of pitches within a pitch collection.

In this latter conception, a minor scale may be considered as a "lowered version" of its parallel major scale. For example, the canonical harmonic minor scale may be conceived of as a version of the major scale, derived from it by lowering the third and sixth scale degrees. This meaning of "lowered pitch" is used for this study. Note that the minor mode is considered a lowered version of the major mode (rather than the major mode a raised version of the minor) simply because the major mode is normative in Western cultures: it accounts for 65 to 75 percent of Western music, and furthermore, listeners tend to assume that unfamiliar melodies are major by default (Huron, 2006). Therefore, we theorize that the minor mode, for those enculturated to such a predominantly major-mode culture, is heard as lower than normal, and consequently heard as sad.

In the present study, we test the hypothesis that melodies using a pitch collection that is lower than the norm will be heard as sadder. Although the results may have implications for the understanding the affect of major and minor modes, it is important to avoid actually using major- or minor-mode stimuli in this experiment because existing learned associations may confound the results. Moreover, any pitch collection that even resembles the major or minor modes should be avoided; otherwise, one could not determine whether the results were truly due to the lowering of pitches or due to the confounding effect of lingering associations to the minor mode.

Accordingly, our experiment used artificial scales, constructed uniquely for each participant, with the aim of minimizing the effect of any unknown associations. A norm or expectation was established during an exposure phase with melodies using an initial artificial scale. Then, the experimental hypothesis was tested during a test phase with melodies using another artificial scale in which a subset of tones were either lowered or raised compared to the initial scale.

METHOD

In brief, the experiment consisted of first exposing listeners to unfamiliar melodies based on unfamiliar scales in the exposure phase. In the subsequent test phase, listeners rated the sadness of melodies that used scales that were either relatively higher or relatively lower in pitch than the scales used in the exposure phase. To anticipate the results, the test melodies that were relatively lower than those in the exposure phase tended to be rated as more sad by listeners.

The following sections describe 1) the novel use of a between-subjects design, 2) the synthesis and modifications of the musical stimuli, and 3) the experimental procedure itself.

1. Between-Subjects Design

To test the effects of relatively lower pitch, it was necessary to modify pitches (i.e., raise or lower them) in the context of a scale or melody. However, whenever a pitch or note in a scale or melody is modified, whether by raising or lowering it, its intervallic relationships with all of the other notes in that scale or melody are also changed. This may cause, for example, unintended changes in overall consonance and dissonance, or, in particular, harmonic implications. One way to control for this is to use identical stimuli for both conditions, and manipulate only the immediately preceding listening experience. Thus, a between-subjects design was used for the experiment, where both participants heard the same melodies as each other in the test phase, after they had heard different melodies in the exposure phase. The exposure phase melodies were created differently for each participant, such that the identical test melodies were heard as either relatively lower for one participant in each pair, or relatively higher for the other participant.

All participants rated 60 melodies, with each melody heard twice consecutively. Thirty exposure stimuli (ES), which were different for each participant, were first heard in the exposure phase. Then, 15 test stimuli (TS), which were identical for paired participants, were heard in the test phase. In addition, the test phase was interspersed randomly with 15 other ES melodies, in order to ensure that the ES were heard as normative. Thus, each participant heard 45 ES melodies and 15 TS melodies. Melodies were uniquely generated for each pair of participants, and each participant rated the sadness of all 60 melodies they heard. Thus, each ES melody was heard by a single participant, while each TS melody was heard by a pair of participants. This experimental design is summarized in Table 1 below.

Participant	Test Condition	Exposure Phase	Test Phase
A	Lower than normal	30 high exposure	15 high exposure, 15 test
B	Higher than normal	30 low exposure	15 low exposure, 15 test

Table 1. Summary of melodies heard for each pair of participants.

The critical element of this design is that although the TS were in fact identical for both participants in each pair, one participant heard them as relatively lower (“low TS”), and the other heard them as relatively higher (“high TS”). That is, the low TS melodies were “lower than normal” due to the higher ES, while the high TS melodies were “higher than normal” due to the lower ES. The two sets of ES (higher and lower) resulted from transposing a set of melodies to two slightly different scales: one that had pitches raised from the TS scale, and one that had pitches lowered. Details of scale construction are given in the Stimuli section below.

It was hypothesized that relatively lower pitch (or “lower than normal” pitch) is a cue for sadness. Thus, it is predicted that for each TS melody, rated by a pair of participants, the participant who heard it as relatively lower TS (due to the higher ES in the exposure phase) would rate it as sadder.

2. Stimuli

In this section, the construction of the 60 melodies for each participant is described. For each pair of participants, three sets of scales were generated: one for the test stimuli (TS scale), which was shared by both participants; one for the low exposure stimuli (low ES scale), which was heard by only one of the participants; and one for the high exposure stimuli (high ES scale), which was heard only by the other participant. The three scales were used in conjunction with selected melodies and a synthesized timbre to create a unique set of stimuli for each pair of participants. It will be seen that there are a variety of constraints and desiderata that led to the peculiar method of stimuli generation as described here.

A naïve approach to testing the hypothesis would be to simply use existing Western melodies as stimuli, then modifying (i.e., raising or lowering) a random selection of pitches. However, since the stimuli would compare music in Western scales with music in modified scales, the most salient feature would probably relate to familiarity, rather than relative pitch height. Thus, in order to avoid confounds due to Western musical conventions, stimuli were explicitly designed so as to be unfamiliar to Western-enculturated listeners. At the same time, the stimuli should not be so unfamiliar (or unconventional) that they seem unmusical. This may be important if the participants are expected to engage emotionally with the music. Therefore, while the use of scales and timbre were designed to be unfamiliar, the rest of the listening experience was designed to be as musically conventional as possible.

In sum, the requirements for the stimuli to be created were that: (i) the created scales should sound unfamiliar, (ii) the scales should allow modifications to produce "higher" and "lower" versions, (iii) the timbre should sound unfamiliar, (iv) the timbre should sound musical, (v) the melodies should sound musical, and (vi) the scales should be average pitch height equivalent.

Stimuli were designed with these desiderata in mind. Artificial scales were created, along with an unusual timbre, in order to subvert Western listening habits. Obscure monophonic Germanic folksongs were selected, to have musically conventional melodic structure. The final stimuli were created by combining the folksongs with the artificial scales and timbre, as detailed below.

(I) "THE CREATED SCALES SHOULD SOUND UNFAMILIAR"

Since the participants were all enculturated to Western music, any resemblances to the major/minor scale system might be expected to confound the results. Accordingly, the test stimuli (TS) scales were generated in a way that subverted Western listening habits by avoiding familiar intervals, avoiding equal temperament, and avoiding octave equivalence. The number of tones in the scale was randomly set in a range from five to nine notes. Then, the distance between successive scale steps was set to a random value between 80 and 350 cents (hundredths of semitones). In order to avoid an excessively small or large pitch compass, a further condition was that the span of the complete scale must lie between 800 and 1400 cents (i.e. between 8 and 14 semitones). This method was used to produce a single TS scale for each pair of participants.

(II) "THE SCALES SHOULD ALLOW MODIFICATIONS TO PRODUCE HIGHER AND LOWER VERSIONS"

After a TS scale was generated, two exposure stimuli (ES) scales were derived from it. Recall that in the exposure phase, melodies were based on a scale that was either higher or lower than the TS scales. To create an ES scale, a random number of tones were modified from the TS scale: one tone was modified at a minimum, and one-third of the TS scale tones were modified at a maximum. (It was assumed that the majority of scale tones must be preserved in order to encourage listeners to hear the test scale as a modified version of an exposure scale, rather than as a wholly novel scale.) For the high ES scale, randomly selected tones from the TS scale were raised in pitch. For the low ES scale, those same scale tones were lowered in pitch. The "tonic" was never modified in this way, and the size of the pitch modification (in cents) was always the same for both the high and low ES scales.

In determining the amount of pitch modification, it was important to maintain clear categorical boundaries of pitches. For example, if a modified pitch should be heard as a raised second scale tone, several kinds of confusion should be avoided. That is, it should not be heard as (a) an unmodified second scale tone, (b) a lowered third scale tone, or (c) a new scale tone altogether (increasing the total notes in the scale).

In order to minimize (a), a minimum pitch modification of 50 cents (i.e., a quarter tone) was established. This minimum distance is in the region of the perceptual limen for pitch interval judgments for naïve Western listeners (Burns & Ward, 1978). In order to minimize possible confusions pertaining to (b) and (c), the amount of pitch modification was restricted so that the modified scale tone was always closer to the original scale tone than any neighboring scale tone. Pitch proximity is known to have a powerful perceptual grouping effect (Bregman, 1990), so there was reason to suppose that proximity to the original scale tone would ensure the appropriate perception.

(III) & (IV) "THE TIMBRE SHOULD SOUND UNFAMILIAR" AND "SOUND MUSICAL"

Synthesized stimuli were used in order to create a sound that seemed both "foreign" and "natural." To this end, the sounds were modeled on the Kalimba (African thumb-piano).

(V) "THE MELODIES SHOULD SOUND MUSICAL"

Although the melodies used a foreign-sounding timbre and employed artificially generated scales, the construction of the melodies was intended to sound musically conventional. Thus, all of the stimuli were based on real melodies, preserving their pitch contour, rhythms, and other tonal elements. Specifically, Germanic folksongs in the Essen Folksong Collection (Schaffrath, 1995) were used as melodic templates for the stimuli. This electronic collection contains over six thousand traditional Germanic folksongs assembled from various sources. Each folksong in the Essen collection is encoded in terms of diatonic scale degrees with possible chromatic alterations.

Melodies were randomly selected for each pair of participants, without replacement. Folksongs containing chromatic tones were excluded from selection, as were songs whose lowest note was not the tonic. Of the 6,255 folksongs in the database, there were 1,685 that fit these criteria. For each pair of participants, 15 randomly selected melodies were mapped to the test stimuli (TS) scale, and 45 randomly selected melodies were mapped to both the high and low exposure stimuli (ES) scales. In mapping the melodies, the tonic (which was always the lowest note) in the original folksong was always mapped to the lowest scale tone in the novel scale, which was not modified. This was done to facilitate the perception of a tonal center, which may help listeners in forming a sense of higher or lower than normal.

The folksong melodies were mapped to the novel scales (whether test, high exposure or low exposure) as follows. A folksong with the same number of unique scale tones as the novel scale was randomly selected, using the criteria above. For example, if a test scale contained six tones, then, the corresponding randomly selected folksong also had to contain six unique scale tones. The lowest scale tone in the folksong was then mapped to the lowest tone in the novel scale, the second lowest to the second lowest, and so on. This permitted a consistent one-to-one mapping between the scale tones in the folksong and the novel scale.

In this way, "homologous" melodies were created in which the "tonic" pitches were the lowest pitch in both the source folksong melodies and in the mapped artificial-scale versions. This approach preserved whatever correlations may exist between tonic pitches and rhythmic or phrase-related patterns in the original folksong. For example, if tonic pitches tend to end phrases, then the designated "tonic" for the novel scales would also tend to end phrases.

(VI) "THE SCALES SHOULD BE AVERAGE PITCH HEIGHT EQUIVALENT"

One final consideration relates to the absolute pitch height of a passage. Huron, Kinney and Precoda (2006) found that exact transposition of a melody influenced perceived affect. Namely, they found that transposing a melody down an octave caused it to be judged as more negatively valenced. In the current experimental design, note that the modification of pitches in a scale will very slightly increase or decrease the overall average pitch height. That is, manipulating relative pitch will also result in a slight change of absolute pitch, which may potentially influence affective judgments. In order to adjust for this possible confound, the ES scales were transposed slightly upward or downward in order to have the same average pitch height as the TS scale from which they were derived. Pitch compensation was applied to the scale as a whole, rather than to the individual melodies. In practice, these compensating transpositions were quite small (typically less than 30 cents). The compensating transpositions did not seem easily detectable to us, even though we were conscious of the manipulations.

The eight sets of generated scales used in the experiment are given in the Appendix, while Figure 1 below summarizes the construction of the TS scales and the derived ES scales.

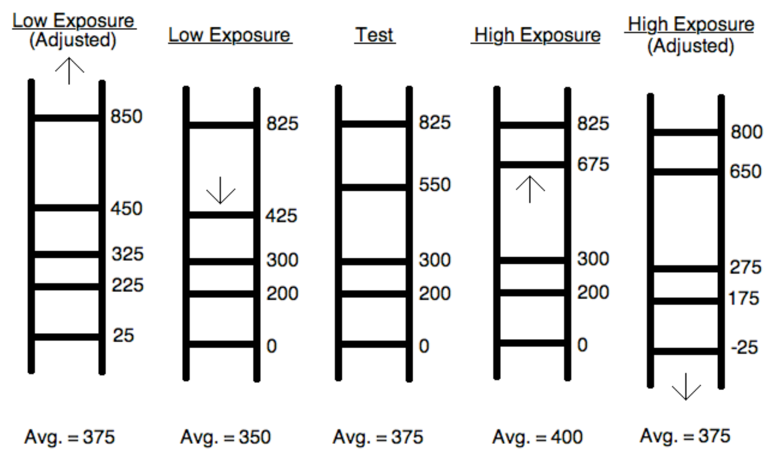

Fig. 1. Schematic for scale construction. Each scale is represented by a ladder, with each scale tone represented by a rung. The central ladder, the TS (test stimuli) scale, is created first, with random values (in cents) for each rung, or pitchwe. To create the ES (exposure stimuli) scales, one of the rungs from the TS scale is randomly selected and then raised or lowered by a random amount, to create the high ES scale or the low ES scale. Then, both ES scales are slightly adjusted to maintain the same overall average pitch height. Only the TS scale and the adjusted ES scales are used for the melodies in the experiment.

3. Procedure

Sixteen participants were recruited from the Ohio State University School of Music subject pool. This was one of several experiments that could be selected by participants in order for them to receive partial course credit. In the recruitment materials, participants were informed that the task would involve judging musical sadness in non-Western melodies. Participants were undergraduate music majors, 7 males and 9 females. All participants reported that they did not possess absolute pitch. Written instructions were provided, and the text was read aloud by the experimenter:

> "In this experiment, you will hear 60 short melodies. Your task is to rate how sad each melody seems to you. If you think the melody is very sad, you should rate it towards the right side of the scale. If it is not sad at all, you should rate it towards the left. After the melody is played twice, there will be an indication on the screen prompting you to respond."

We only considered degrees of sadness (that is, more or less sad), and not happiness directly. While some may consider happiness and sadness as opposites on a single continuum, others may argue that they are qualitatively separate experiences. Accordingly, the rating scale had 7 discrete points, and only the endpoints were labeled, with "Very sad" on the right and "Not sad at all" on the left.

The exposure phase (30 exposure melodies) and test phase (15 test melodies and 15 exposure melodies interspersed) were joined as one continuous session (approximately 30 minutes in total), such that there was no indication to participants that there were two sets of stimuli. Participants completed the experiment individually to avoid any indication of the paired, between-subjects design. Responses were collected after two presentations of each melody (approximately 30 seconds each), and participants responded by using a mouse to click on the rating scale displayed on a computer screen. Stimuli were presented using a computer connected to loudspeakers inside an Industrial Acoustics Corporation sound attenuation room. The statistical analyses were performed in R 3.2.2.

RESULTS

Since different participants may not have used the rating scale in a comparable way, standardized ratings were obtained. Fifteen test stimuli (TS) melodies were rated by each of the 16 participants, for a total of 240 TS sadness ratings (120 ratings from "high TS" participants, and 120 ratings from their "low TS" counterparts). The TS ratings were Z-standardized for each participant, using the means and standard deviations for their ratings of the 30 ES melodies in the exposure phase. In order to test the experimental hypothesis, the ratings of the participant who heard the TS as lower than normal ("low TS") were compared with those of the participant who heard the same set of TS as higher than normal ("high TS"). Recall that the low TS and high TS are actually identical stimuli, but are relatively lower or higher than the ES in the exposure phase. According to the experimental hypothesis, the low TS melodies would be rated as more sad, while the high TS melodies would be rated as less sad.

Indeed, the mean standardized sadness rating for the low TS melodies was higher than the mean standardized sadness rating for the high TS melodies. Since each TS melody was heard twice (by paired participants, once as "high TS" and once as "low TS"), it is appropriate to carry out a matched pairs *t*-test (one-tailed). The difference in mean ratings was significant, with low TS melodies having a standardized sadness rating that was 0.392 greater (95% CI [0.117, 0.666], t = 2.6, df = 234.16, p-value = 0.005), consistent with the prediction that participants who heard a higher exposure scale rated the test melodies as more sad than those who heard a lower exposure scale.

Upon inspection of the data, it was found that the assumption of normality was not met (Shapiro-Wilk's W = 0.97, p = 0.02 for the low ES participants; W = 0.95, p < 0.001 for the high ES participants.) Thus, a non-parametric alternative was also conducted, the Wilcoxon signed rank test with continuity correction (one-tailed). Again, the results were significant (W = 4519, p = 0.010), in agreement with the paired *t*-test. However, both the *t*-test and the rank sum test require an assumption of data independence, but each participant provided 15 ratings (and the sample size is not large).

To account for the dependence of the data, a linear mixed effects model was constructed, using random-intercepts for each participant pair, with test type ("high" or "low" TS) as the predictor. Participant pair was used for the random variable instead of individual participants, since each individual participant only heard one of the test type, while participant pairs heard both test types. Moreover, each pair heard a unique set of constructed melodies based on similar scales. Thus the "pair" variable accounts for individual differences in melodies heard between participant pairs, but not for differences in pitch height. Despite the fact that this test does not leverage the pairing of the individual melodies, a significantly higher sadness rating for the low TS melodies was again found (p = 0.0073, 95% CI [0.104, 0.679], adjusted R-squared = 0.082).

DISCUSSION

An experiment was conducted on the perception of sadness in melodies. The results of the experiment were generally consistent with the hypothesis that melodies using a pitch collection containing lowered notes are perceived as sadder. The study was designed to avoid two particular confounding factors. First, artificial scales were used instead of Western major or minor scales, in order to minimize the effects of enculturation. Additionally, a between-subjects design was used in order to minimize confounds related to pitch manipulations. Participants heard two kinds of melodies: exposure scale (ES) melodies and test scale (TS) melodies, and rated each melody for perceived sadness. ES melodies were heard in the first part of the experiment, in order to establish an expectation or norm. TS melodies were heard subsequently, and were either relatively higher or relatively lower than the ES melodies. That is, the TS melodies were actually identical for participant pairs, but the ES melodies were manipulated such that the TS melodies were heard as either relatively higher (high TS) or relatively lower (low TS) for either participant in a pair. As predicted, the low TS melodies were rated as sadder (after they were Z-standardized using each participants ES ratings). Thus, it may be concluded that melodies heard as lower than the norm were perceived as sadder.

This result offers a potential explanation for the sadness of the minor mode for some listeners: if minor-mode melodies are heard as lower than the normative major mode, this lowering of pitch may be heard as sad. This account suggests that a resemblance to the minor mode is not sad in and of itself. Rather, having pitches lowered from a normative scale is the critical factor. A consequence of this "lower than normal" theory is that affective connotations are expected to change with changing scale norms, and that the sad affect of minor-like scales may vary across cultures with different normative scales.

Considering Non-Western Cultures: Why Minor is Not Always Sad

The Dorian mode sounds somewhat sad to modern ears (e.g., Temperley & Tan, 2013). However, medieval commentators describe the Dorian mode as relatively positive or happy. Note that during the Middle Ages, the Dorian mode was the most common mode. The view suggested here is that without it being lower than another normative mode, there is no reason to expect the Dorian mode to evoke sad connotations. Only in the sixteenth-century, with the transition to the major-minor system, do we find evidence of changing affective descriptions, as in a famous characterization by Giosseffo Zarlino: "The first mode [Dorian] has a certain effect midway between sad and cheerful ... By nature it is somewhat mixed." (Zarlino, 1558)

As another example, ethnomusicologists have long noted that scales similar to the Western minor scale exist in several regions of the world, including North Africa, the Middle East and the Balkans (e.g. Rice, 2004). The observation that these scales evoke no sad connotations for enculturated listeners is consistent with the "lower than normal" account, when we consider that the major scale is not normative in such cultures. Again, without a comparison to a normative major mode, there would be no reason to expect these minor-like scales to be perceived as sadder, since they would not be perceived as lower.

The present experiment did not directly use the major scale as the "normative" scale, since any results from such an experiment could not be distinguished from enculturation generally, rather than a "lower than normal" effect specifically. For this reason, our experimental design used artificial materials that were arguably poor in ecological validity. Consequently, one of us (Chordia & Sastry, 2011) initiated a second experiment to see whether the results of this experiment would generalize when melodies based on raags from North Indian Classical music (NICM) were used. As this study was never published, we summarize it briefly here.

A web-based study was conducted in which participants ($N = 544$, 88% male) were asked to judge the happiness and sadness of short musical excerpts (15-20 seconds in length). These stimuli were based on one of three raag types, out of hundreds of possible raags used in NICM. In contrast to the modern Western notion of a scale, a raag might be regarded as a melodic prototype in which certain pitch patterns or figures provide a template for improvisation. Nevertheless, a raag defines a collection of pitches akin to a Western scale. The three raags used in the study differed in the number of lowered scale degrees.

Rather than pairing participants (as in the current study), participants were randomly assigned to either the low or high exposure group. Those in the low-exposure group heard 10 melodies composed with the "low" raag (Raag Todi), which contains three lowered scale degrees; those in the high-exposure group heard 10 melodies in the "high" raag (Raag Khamaj), which contains one. Both groups then heard five melodies in the "test" raag (Raag Kafi), which contains 2 lowered scale degrees. This design (as in the current study) allowed for identical test melodies to be heard as "higher" for the former group and "lower" for the latter group. Consistent with the "lower than normal" hypothesis proposed presently, the group that heard the test melodies as lower heard it as more sad, with a standardized z-score difference of 0.36 ($p < 0.001$). Unsurprisingly, they also heard those melodies as less happy, with a standardized z-score difference of -0.40 ($p < 0.001$). (Tests were corrected for multiple comparisons using the Tukey-Cramer statistic.) It is also notable that this significant result was obtained using an exposure phase of only ten melodies, heard once, totaling three minutes of musical exposure. (Recall that the present study used thirty exposure melodies, heard twice, totaling 15 minutes.)

The online participants also reported their familiarity with NICM: "Not at all familiar" (35%), "A little familiar" (36%), "Somewhat familiar" (20%), and "Very familiar" (8%). The difference in sadness ratings and happiness ratings was greatest for the group that was "very familiar" with NICM, and lowest for the group that was "not at all familiar" with NICM. That is, the effect was greatest for the group that was most enculturated to a non-Western musical language. This observation suggests that the results of the present experiment may be culturally generalizable, since hearing lower-than-normal melodies as sadder does not seem dependent on Western enculturation (if we assume that those most enculturated to NICM are least enculturated to Western music), and that further cross-cultural studies may be fruitful.

What is Normal?

Central to this this "lower than normal" account for perceived sadness is identifying what is meant by "normal." We propose that the normative scale, or the most prevalent scale in a given culture (as in the major scale, for Western music), is used as a frame of reference for assessing lowered pitches, and that the use of such lowered pitches (as in the minor scale) may be a cue for sadness. A recent paper by Temperley and Tan (2013) raised a critical issue of interpretation regarding an earlier presentation of the current study (Huron, Yim, & Chordia, 2010), which we believe may be resolved by applying this definition.

As an alternative explanation for the perception of sadness/happiness in melodies, Temperley and Tan (2013) presented research that was consistent with the "line-of-fifths" model. In a two-alternative forced choice paradigm, they presented melodies in transpositions of two different church modes. Participants were asked to judge which melody sounded happier. They found that as the number of flats was increased, starting with the Mixolydian mode and ending with the Phrygian mode, happiness ratings decreased. At face value, this result is consistent with our results and conclusions, and both accounts (i.e., their line-of-fifths model and our "lower than normal" model) make the same predictions for the happy or sad connotations of the diatonic modes.

However, Temperley and Tan argue that the two models would make different predictions for the pentatonic modes. In Figure 2 below, the five pentatonic modes are listed. In their line-of-fifths account, moving from left (Ionian) to right (Phrygian) is equivalent to "adding flats," and decreasing perceived happiness (or increasing perceived sadness) would be predicted, in line with musical intuition. Temperley and Tan argue, however, that our "lower than normal" account would predict the exact opposite trend, contrary to musical intuition. For example, the third pentatonic mode (Dorian) may be considered a relatively higher version of the second pentatonic mode (Mixolydian), by raising the fifth scale degree, from A-natural to A-sharp (or B-flat). Thus, they argued that moving from left to right would be equivalent to "raising pitches", and consequently, and increasing perceived happiness (or decreasing perceived sadness) would be perceived.

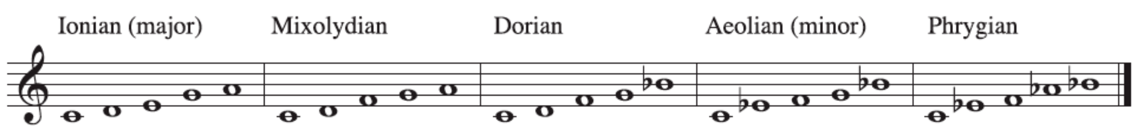

Fig. 2. Five pentatonic modes. Modes left-to-right represent decreasing perceived happiness according to the line-of-fifths model (reproduced from Temperley & Tan, 2013).

Note that this argument assumes a major pentatonic scale as the starting point—with successively raised notes yielding the modes from Ionian to Phrygian. The rebuttal offered here is that since pentatonic modes are not normative, they should not be expected to be the basis for determining which notes are considered raised or lowered. For Western-enculturated listeners, one may expect that the pentatonic modes be heard relative to the diatonic major scale instead. For example, the second pentatonic mode (Mixolydian) has no lowered notes relative to the diatonic major scale, while the third pentatonic mode (Dorian) has one (B-flat). From this perspective, moving left to right through the pentatonic modes, Ionian to Phrygian, would be equivalent to "lowering pitches", and would be expected to sound progressively sadder – which is the same intuitive prediction as the line-of-fifths account.

CONCLUSION

In summary, listeners heard melodies in an experiment where test melodies were based on pitch collections either higher or lower in pitch compared to initial exposure melodies. All melodies used unconventional scales, in order to avoid learned associations with the major-minor modes. Additionally, participants were paired such that the identical melodies were judged in the test phase, but were heard as either higher or lower depending on the melodies heard in exposure phase. The results were consistent with the hypothesis that melodies using lowered pitches (compared to what is "normal") are perceived as sadder. Additionally, the results are suggestive of an explanation for the association of the sadness of the minor mode: it's heard as sad when listeners hear it as lower than the normative major mode.

END MATTER

NOTES

[1] Correspondence can be addressed to: Gary Yim, School of Music, 1866 N College Rd., Columbus, OH 43210, yim.28@osu.edu.

REFERENCES

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Bergmann, G., Goldbeck, T., & Scherer, K. (1988). Emotionale Eindruckswirkung von prosodischen Sprechmerkmalen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *35*(2), 167-200.

Bowling, D. L., Gill, K., Choi, J. D., Prinz, J., & Purves, D. (2010). Major and minor music compared to excited and subdued speech. *Journal of the Acoustical Society of America, 127*(1), 491-503.

Bregman, A. S. (1990). *Auditory scene analysis: the perceptual organization of sound.* Cambridge: MIT Press.

Breitenstein, C., Van Lancker, D., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Journal of Cognition and Emotion, 15*(1), 57-79.

Brown, S. (2000). The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of Music,* pp. 271-300. Cambridge: MIT Press

Burns, E. M., & Ward, W. D. (1978). Categorical perception - phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America, 63*(2), 456-468.

Chordia, P., & Sastry, A. (2011). The effect of pitch exposure on sadness and happiness judgments: Further evidence for "lower-than-normal" is sadder, and "higher-than-normal" is happier. Paper presented at the Society for Music Perception Conference, Rochester, New York (August 13, 2011).

Cook, N. D. (2007). The sound symbolism of major and minor harmonies. *Music Perception, 24*(3), 315-319.

Costa, M., Fine, P., & Bitti, P. E. R. (2004). Interval distributions, mode, and tonal strength of melodies as predictors of perceived emotion. *Music Perception, 22*(1), 1-14.

Crowder, R. G. (1984). Perception of the major/minor distinction: I. Historical and theoretical foundations. *Psychomusicology, 4*(1-2), 3-12.

Crowder, R. G., Reznick, J. S., & Rosenkrantz, S. L. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society, 29*(3), 187-188.

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion, 10*(3), 335-348.

Davitz, J. R. (1964). Auditory correlates of vocal expressions of emotional meaning. In J. R. Davitz (Ed.), *Communication of Emotional Meaning,* pp. 101-112. New York: McGraw-Hill.

Eldred, S. H., & Price, D. B. (1958). A linguistic evaluation of feeling states in psychotherapy. *Psychiatry, 21*(2), 115-121.

Fairbanks, G., & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotion. *Speech Monographs, 6*(1), 87-104.

Garrido, S., & Schubert, E. (2011). Negative emotion in music: what is the attraction? A qualitative study. *Empirical Musicology Review, 6*(4), 214-230.

Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion, 7*(2), 377-387.

Hargreaves, W. A., Starkweather, J. A., & Blacker, K. H. (1965). Voice quality in depression. *Journal of Abnormal Psychology, 70*(3), 218-220.

Heinlein, C. P. (1928). The affective characters of the major and minor modes in music. *Comparative Psychology, 8*(2), 101-142.

Helmholtz, H. L. (1912). *On the Sensations of Tone as a Physiological Basis for the Theory of Music (4th ed.).* (A. J. Ellis, Trans.) London: Longmans, Green and Co. (Original work published 1877)

Hevner, K. (1935). The affective character of the major and minor modes in music. *The American Journal of Psychology, 47*(1), 103-118.

Hevner, K. (1937). The affective value of pitch and tempo in music. *The American Journal of Psychology, 49*(4), 621-630.

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

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge: MIT Press.

Huron, D. (2008). A comparison of average pitch height and interval size in major- and minor-key themes: Evidence consistent with affect-related pitch prosody. *Empirical Musicology Review, 3*(2), 59-63.

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

Huron, D. (2015). Affect induction through musical sounds: an ethological perspective. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 370*(20140998). http://dx.doi.org/10.1098/rstb.2014.0098

Huron, D., Anderson, N., & Shanahan, D. (2014). "You can't play a sad song on the banjo:" Acoustic factors in the judgment of instrument capacity to convey sadness. *Empirical Musicology Review, 9*(1), 29-41.

Huron, D., Kinney, D., & Precoda, K. (2006). Influence of pitch height on the perception of submissiveness and threat in musical passages. *Empirical Musicology Review, 1*(3), 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 … (Eds.), *Proceedings of the 11th International Conference on Music Perception and Cognition,* Seattle, Washington, 23-27 August.

Huttar, G. (1968). Relations between prosodic variables and emotions in normal American English utterances. *Journal of Speech and Hearing Research, 11*(3), 481-487.

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

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research, 33*(3):217-238.

Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception, 8*(2), 189-201.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America, 29*(1), 98-104.

Matthews, M. V., & Pierce, J. R. (1980) Harmony and nonharmonic partials. *Journal of the Acoustical Society of America, 68*(5), 1252-1257.

Mithen, S. (2009). The music instinct. *Annals of the New York Academy of Sciences, 1169,* 3-12.

Nawrot, E. S. (2003). The perception of emotional expression in music: Evidence from infants, children and adults. *Psychology of Music, 31*(1), 75-92.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38*(4), 548-560.

Post, O., & Huron, D. (2009). Western Classical music in the minor mode is slower (Except in the Romantic period). *Empirical Musicology Review, 4*(1), 2-10.

R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rice, T. (2004). *Music in Bulgaria: Experiencing music, expressing culture.* Oxford: Oxford University Press.

Roederer, J. G. (1984). The search for a survival value of music. *Music Perception, 1*(3), 350-356.

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music, 35*(1), 88-109.

Sachs, M.E., Damasio, A., & Habibi, A. (2015). The pleasures of sad music: A systematic review. *Frontiers in Human Neuroscience, 9*(404)*,* 1-12. http://dx.doi.org/10.3389/fnhum.2015.00404

Schaffrath, H. (1995). The Essen folksong collection. In D. Huron (Ed.), *Database containing 6,255 folksong transcriptions in the Kern format and a 34-page research guide [computer database].* Menlo Park, CA: CCARH.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, K. R. Scherer, & H. Goldsmith (Eds.), *Handbook of Affective Sciences,* pp. 433-456. Oxford: Oxford University Press.

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

Shutz, M., Huron, D., Keeton, K., & Loewer, G. (2008). The happy xylophone: acoustic affordances restrict an emotional palate. *Empirical Musicology Review, 3*(3), 126-135.

Skinner, E. R. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness; and a determination of the pitch and force of the subjective concepts of ordinary, soft, and loud tones. *Speech Monographs, 2*(1), 81-137.

Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness and joy. *Journal of Psycholinguistic Research, 28*(4), 347-365.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: an online survey. *PLoS ONE, 9*(10), 1-17. http://dx.doi.org/10.1371/journal.pone.0110490

Temperley, D., & Tan, D. (2013). Emotional connotations of diatonic modes. *Music Perception, 30*(3), 237-257.

Turner, B., & Huron, D. (2008). A comparison of dynamics in major- and minor-key works. *Empirical Musicology Review, 3*(2), 64-68.

Valentine, C. W. (1913). The aesthetic appreciation of musical intervals among school children and adults. *British Journal of Psychology, 6*(2), 190-216.

Van den Tol, A. J. M., & Edwards, J. (2011). Exploring a rationale for choosing to listen to sad music when feeling sad. *Psychology of Music, 41*(4), 440-465.

Van den Tol, A. J. M., & Edwards, J. (2015). Listening to sad music in adverse situations: How music selection strategies relate to self-regulatory goals, listening effects, and mood enhancement. *Psychology of Music, 43*(4), 473-494.

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

Vuoskoski, J. K., & Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception, 29*(3), 311-317.

Weth, K., & Parncutt, R. (2014, 18-20 September). Functions of sad music in daily life: An explorative study on the influence of exposure, coping strategies and personality. In Jakubowski, K., Farrugia, N., Floridou, G.A., & Gagen, J. (Eds.), *Proceedings of the 7th International Conference of Students of Systematic Musicology (SysMus 14)*, London, UK, 18-20 September. http://ojs.gold.ac.uk/index.php/sysmus14/article/view/261

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America, 52*(4), 1238-1250.

Zarlino, G. (1983). *Le Institutioni hamoniche, Book 4.* (V. Cohen, Trans.) Yale University Press. (Original work published 1558)

APPENDIX

Scale Group	Scale Variant	Scale Degree								
		#1	#2	#3	#4	#5	#6	#7	#8	#9
1	Test	0	181	320	421	681	903	1083	1225	1325
	Low	15	196	285	436	696	828	1098	1240	1340
	High	-155	166	355	406	666	978	1068	1210	1310
2	Test	0	146	255	522	747	932	1079	1186	--
	Low	20	112	275	430	767	954	1099	1206	--
	High	-20	118-	235	614	727	914	1059	1166	--
3	Test	0	267	492	680	825	933	1199	--	--
	Low	10	277	502	618	835	943	1209	--	--
	High	-10	257	482	742	815	923	1189	--	--
4	Test	0	171	299	591	840	1047	1216	--	--
	Low	12	183	311	603	852	1059	1144	--	--
	High	-12	159	287	579	828	1035	1288	--	--

5	Test	0	278	514	712	868	986	--	--	--
	Low	26	304	441	738	835	1012	--	--	--
	High	-26	252	587	686	901	960	--	--	--
6	Test	0	103	364	588	769	913	--	--	--
	Low	18	121	270	606	787	931	--	--	--
	High	-18	85	458	570	751	895	--	--	--
7	Test	0	132	422	670	881	--	--	--	--
	Low	21	153	443	691	797	--	--	--	--
	High	-21	111	401	649	965	--	--	--	--
8	Test	0	127	418	667	878	--	--	--	--
	Low	21	148	439	667	794	--	--	--	--
	High	-21	106	397	646	962	--	--	--	--

Fig. 3. Relative cents values for experimental scales described in methodology. Eight groups of generated scales are represented, with each group having a test scale, a relatively “low” exposure scale and a relatively “high” exposure scale. Each group of scales had five to nine notes per scale. The relative cents values, with 0 as middle C, is given for each of the notes.