A Cross-Cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues

LAURA-LEE BALKWILL & WILLIAM FORDE THOMPSON

Atkinson College, York University

Studies of the link between music and emotion have primarily focused on listeners' sensitivity to emotion in the music of their own culture. This sensitivity may reflect listeners' enculturation to the conventions of their culture's tonal system. However, it may also reflect responses to psychophysical dimensions of sound that are independent of musical experience. A model of listeners' perception of emotion in music is proposed in which emotion in music is communicated through a combination of universal and cultural cues. Listeners may rely on either of these cues, or both, to arrive at an understanding of musically expressed emotion. The current study addressed the hypotheses derived from this model using a cross-cultural approach. The following questions were investigated: Can people identify the intended emotion in music from an unfamiliar tonal system? If they can, is their sensitivity to intended emotions associated with perceived changes in psychophysical dimensions of music? Thirty Western listeners rated the degree of joy, sadness, anger, and peace in 12 Hindustani raga excerpts (field recordings obtained in North India). In accordance with the raga-rasa system, each excerpt was intended to convey one of the four moods or "rasas" that corresponded to the four emotions rated by listeners. Listeners also provided ratings of four psychophysical variables: tempo, rhythmic complexity, melodic complexity, and pitch range. Listeners were sensitive to the intended emotion in ragas when that emotion was joy, sadness, or anger. Judgments of emotion were significantly related to judgments of psychophysical dimensions, and, in some cases, to instrument timbre. The findings suggest that listeners are sensitive to musically expressed emotion in an unfamiliar tonal system, and that this sensitivity is facilitated by psychophysical cues.

IT is generally agreed that humans tend to make associations between emotion and music, but many questions remain as to the nature of these associations. For example, although it is well established that people make similar judgments when they are asked to describe musically expressed

Address correspondence to Laura-Lee Balkwill or William Forde Thompson, Department of Psychology, Atkinson College, York University, Toronto, Ontario M3J 1P3, Canada. (e-mail: balkwill@yorku.ca or billt@yorku.ca)

emotion (Behrens & Green, 1993; Dolgin & Adelson, 1990; Gabrielsson & Juslin, 1996; Gerardi & Gerken, 1995; Haack, 1980; Kratus, 1993; Robazza, Macluso, & D'Urso, 1994; Terwogt & van Grinsven, 1991; Thompson & Robitaille, 1992), it is not known to what extent these judgments are based on knowledge of cultural conventions or on basic perceptual cues that transcend cultural boundaries.

Some theorists favor an enculturation explanation—that musical meaning is determined exclusively by cultural convention (Blacking, 1973; Feld & Keil, 1994; Walker, 1996). These theorists may acknowledge certain universals in music, such as octave equivalence (Dowling & Harwood, 1986), but they assume that the interpretation of emotional meaning in music is determined solely by enculturation to a specific tonal system. Feld and Keil (1994, p. 85) provide an example of this perspective: "All musical sound structures are socially structured in two senses: they exist through social construction, and they acquire meaning through social interpretation."

Empirical evaluation of this issue has been complicated by the unicultural focus of previous research.[1] With few exceptions (Deva & Virmani, 1975; Gregory & Varney, 1996; Hoshino, 1996), the existing research on emotional meaning in music has focused on music from the Western tonal system and on the judgments of Western listeners. This body of research does not, therefore, provide a basis for evaluating the enculturation hypothesis of the perception of emotion in music.

Although no one would deny the influence of culture on individuals' judgments of emotion in music, there may also be universal influences underlying musically expressed emotion. In addition to the conventions of a specific tonal system, listeners may be influenced by variations in psychophysical dimensions of music, such as tempo, timbre, and stimulus complexity.

For our purposes, a psychophysical dimension will be defined as any property of sound that can be perceived independent of musical experience, knowledge, or enculturation. For example, speed of pulse, or tempo, is a psychophysical dimension, whereas a perfect cadence is not. Complexity is also a psychophysical dimension, because judgments of complexity do not rely on familiarity with the stimulus. Early studies of experimental aesthetics revealed that judgments of complexity are strongly influenced by such factors as the number of elements perceived, and the degree of information redundancy (for a review, see Berlyne, 1971). Such variables can be manipulated in unfamiliar or otherwise meaningless stimulus patterns (e.g., random dot patterns), resulting in reliable changes in judged complexity.

A number of studies have reported strong relationships between emotional judgments and psychophysical dimensions such as tempo, timbre,

1. The term "uni-cultural" refers to studies in which the listeners and the music stimuli are of the same culture.

and loudness (Behrens & Green, 1993; Gabrielsson & Juslin, 1996; Gerardi & Gerken, 1995). Other evidence suggests that increases in perceived complexity may be associated with arousal responses (Berlyne, 1971, p. 69) and shifts in aesthetic evaluation (Day, 1967, 1968; Vitz, 1966b). These dimensions of music are not culture-specific, and hence, they may function as universal cues for the emotional evaluation of all auditory stimuli. Although the basis for such evaluations is unknown, they may have originated as adaptive responses to acoustic input—for example, to threatening and nonthreatening sounds.

A Model of the Perception of Emotion in Music

Studies of the perception of unfamiliar music have illustrated that listeners who are unfamiliar with a musical idiom are nonetheless able to apprehend structural aspects of that music by attending to basic statistical properties. Without knowledge of the stylistic conventions of the music, listeners attend to the statistical distribution of pitches as one way of interpreting North Indian music (Castellano, Bharucha & Krumhansl, 1984), Balinese music (Kessler, Hansen, & Shepard, 1984), and artificial tonalities (Oram & Cuddy, 1995). These findings illustrate that listeners are highly adaptable in their ability to interpret unfamiliar musical styles.

It is reasonable to assume that the perception of emotion in music is similarly adaptable. Consciously or intuitively, composers and performers may draw upon basic perceptual cues (in the form of psychophysical dimensions of music) as well as culturally determined conventions, to express emotion in music. Listeners, in turn, may attend to either of these sources of emotional meaning. Because emotion is conveyed through both culture-specific and perceptual cues, there is often redundancy in how a piece reinforces a specific emotion. The more cues that are present in the music—both culture-specific and perceptual—the stronger the expression of emotion. Thus, listeners' understanding of the musically expressed emotion should be affected both by their familiarity with the conventions of the tonal system and also by their sensitivity to basic perceptual cues.

In Figure 1, we present a model of how listeners attend to the cultural and universal cues embedded in a piece of music to arrive at an understanding of musically expressed emotion. The expression of emotion in the music of a given culture should be most salient to listeners of the same culture, due to the shared understanding of the conventional representation of emotion within that tonal system.

When familiar cultural cues are absent, the listener must pay more attention to basic perceptual cues, such as tempo and complexity. These cues allow listeners to gain a general understanding of the intended emotion.
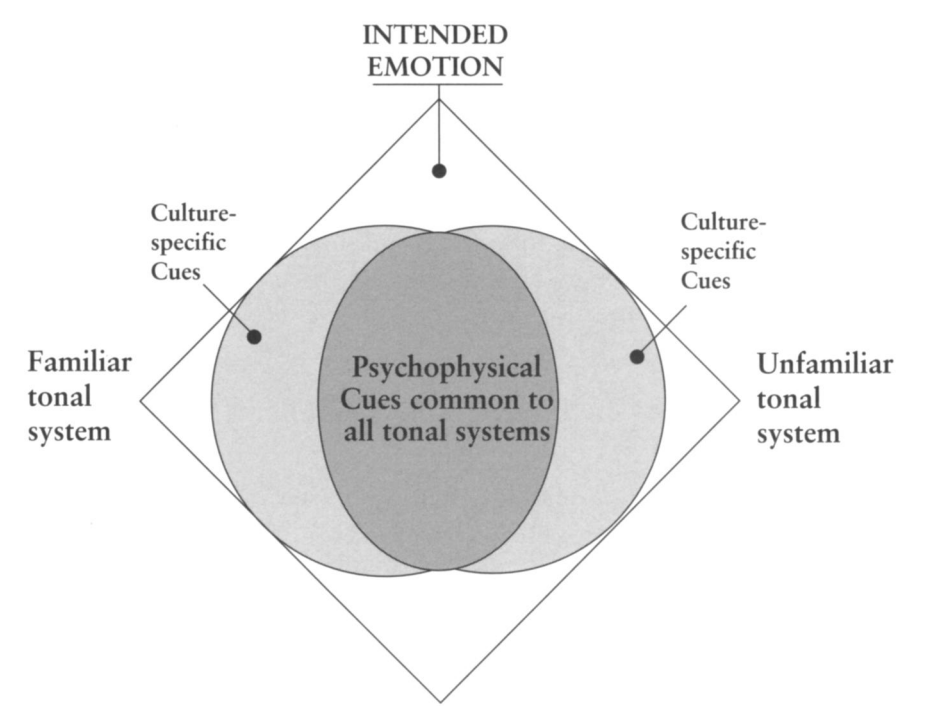
Fig. 1. The figure illustrates several points: (1)Each tonal system has its own distinct cultural cues pertaining to musically expressed emotion that may not overlap with the culture-specific emotion cues of another tonal system. (2) Psychophysical cues pertaining to emotion are present in all tonal systems and thus provide an overlap of information facilitating cross-cultural recognition of musically expressed emotion.

It is possible to argue that the association of psychophysical dimensions of music with the perception of emotions is also the result of enculturation and may differ between cultures. If so, then listeners should be unable to identify an intended emotion in music from an unfamiliar tonal system. If these associations are not culturally determined, however, then listeners should be able to transcend cultural boundaries and recognize intended emotions in music from an unfamiliar tonal system. The present study was designed to investigate this possibility.

Our first hypothesis was that listeners enculturated to one tonal system can accurately perceive the intended emotion in music from an unfamiliar tonal system. Our second hypothesis was that listeners use psychophysical dimensions of music as cues to identify musically expressed emotions. To evaluate these hypotheses, listeners enculturated to Western tonal music were asked to judge the emotional content of music performed in an unfamiliar tonal system from another culture. Listeners also provided judgments of several psychophysical dimensions in the same musical stimuli.

Pilot Study

As an initial assessment of the model, a pilot study was conducted. Fourteen students at York University, Toronto, were exposed to six short music samples (less than 2 min each) in the tonal systems of Kyrgystan, Hindustan, and the Navaho Indians.[2] Two samples, one representing joy and one representing sadness, were taken from each tonal system. Half of the students were asked to rate their perceptions of tempo and complexity for each sample on scales numbered one to nine (low to high). The other half were asked to rate their perceptions of joy and sadness on similar scales for each sample.

The results indicated that listeners assigned high ratings of joy (mean = 6.75) and low ratings of sadness (mean = 3.14) to the three samples intended to convey joy. Conversely, they assigned high ratings of sadness (mean = 5.62) and low ratings of joy (mean = 2.33) to the three samples intended to convey sadness. These results support the view that listeners can perceive musically expressed emotion in music from an unfamiliar tonal system.

A correlational analysis revealed the relationship between listeners' ratings of joy and tempo, and sadness and tempo. The correlation between joy and tempo, $r(10) = .87$, $p = .023$, and between sadness and tempo, $r(10) = -.86$, $p = .027$, supported the view that listeners' ratings of emotion in the music were influenced by the psychophysical dimension of tempo. The direction of these relationships is identical to those reported by previous studies involving only Western tonal music (see Haack, 1980, for a review). The correlations between ratings of joy and complexity and ratings of sadness and complexity were not significant.

The results of this pilot study suggested that listeners of one culture can perceive the intended emotion in music from unfamiliar tonal systems. Furthermore, there was a very strong relationship between judgments of joy and sadness and judgments of tempo.

A larger experiment was conducted to corroborate and expand upon these preliminary findings with a wider range of emotion judgments and psychophysical judgments. We focused on Western listeners' perceptions of emotions and psychophysical dimensions in samples of Hindustani ragas.

The Hindustani classical tonal system is particularly suited to the study by virtue of the *raga-rasa* aspect of Hindustani music theory. Hindustani classical theory specifically outlines traditional relationships between the ragas and their corresponding moods, or rasas. An interesting aspect of

2. Music stimuli: Track 2 and 10 from *Songs of the Navajo*. (1993). JVC World Sounds, Albuquerque, New Mexico; Track 2 and 14 from *Songs of Kyrgyzstan*. (1992). World Music Library; Bhupali alap, and Jogya alap, taken from field recordings (India, Dec. 1996).

classical Hindustani music is the performer's freedom to reinforce the expressed rasa of each piece. The emotional theme of a raga is most strongly expressed during the *alap*,[3] or improvisational, section of the performance (for a more detailed description of Hindustani music theory, see Kaufmann, 1968).

By using field recordings (obtained during a research trip to India in December 1996), we were able to obtain alap samples for which the individual performer's intention to evoke a specific rasa could be verified. The two performers who participated in this study were asked to play the alap portions of ragas that they would normally choose if they wished to convey four specific emotions: joy/*hasya*, sadness/*karuna*, anger/*raudra*, and peacefulness/*shanta*.

These four target emotions have a well-documented history in the psychological literature and they are the nearest equivalents of four of the nine rasas identified in Hindustani music theory (A. Ranade, S. Rao, & S. Devasthali, personal communication, December 1996 to January 1997). Joy and sadness have been the most reliably identified and distinguishable musically expressed emotions in Western music by Western listeners (Haack, 1980). Anger has also been identified in Western music, albeit with less consistency (Behrens & Green, 1993; Terwogt & van Grinsven, 1991; Thompson & Robitaille, 1992). Anger is not often conveyed in Hindustani music: The corresponding rasa, raudra, is very rarely used and is often considered to be incongruent with the North Indian view of music as inherently peaceful (A. Ranade, S. Rao, & S. Devasthali, personal communication, December 1996 to January 1997). Peace, on the other hand, is considered the basic rasa of all ragas. Peace was also included as a suitable opposite to anger and because several studies have indicated that peacefulness (a.k.a. restfulness) is reliably perceived in music (Terwogt & van Grinsven, 1991; Thompson & Robitaille, 1992).

A number of psychophysical dimensions of music have been linked with judgments of emotion in music, including tempo, melodic contour, harmonic complexity, melodic complexity, rhythmic complexity, articulation, dynamics, consonance/dissonance, pitch register, and timbre (Crozier, 1974; Gabrielsson & Juslin, 1996; Gerardi & Gerken, 1995; Hevner, 1935, 1936; Holbrook & Anand, 1990; Kastner & Crowder, 1990; Kratus, 1993; Nielzén & Cesarec, 1982; Rigg, 1940; Scherer & Oshinsky, 1977; Vitz, 1966a; Wedin, 1972). Five of these dimensions were examined: tempo, melodic complexity, rhythmic complexity, pitch range, and timbre.

___

3. The alap is the opening section of any traditional Hindustani raga and is traditionally used to set the mood, or rasa, of the rag before the set composition (the gat) section. Although performers must only use the notes within the chosen raga scale, the duration, tempo, and rhythm of this section can be varied at their discretion.

Among these psychophysical dimensions, tempo has been most consistently associated with emotional content. Hevner (1935, 1936) reported a strong association between ratings of sadness and joy in music and slow and fast tempi, respectively. These findings have been replicated in a number of studies (for a review, see Haack, 1980) and more recently by Gabrielsson and Juslin (1996). In accordance with these findings for Western tonal music, we predicted that higher ratings of joy would be associated with higher ratings of tempo, whereas higher ratings of sadness would be associated with lower ratings of tempo.

Melodic and rhythmic complexity may also be useful in the differentiation of emotions in music. Our predictions were based on both theoretical and empirical considerations. First, according to Berlyne's (1971) optimal complexity model, listeners should prefer stimuli perceived to have a middle level of complexity. Because the ragas were unfamiliar to listeners, only relatively simple melodies and rhythms should fall within this optimal level. Second, researchers have found that simpler melodies, that is, those containing fewer variations of melodic contour and more repetition, are associated with positive and peaceful emotions, while more complex melodies, that is, those featuring more variation and less repetition, are associated with anger and sadness (Crozier, 1974; Imberty, 1974; Vitz, 1966a). Finally, a number of studies have reported an association between irregular rhythms and negative moods and between regular, repetitive, or flowing rhythms and positive moods (Hevner, 1935; Rigg, 1964). Based on these considerations, we predicted that positive emotions would be associated with relatively simple melodies and rhythms, whereas negative emotions would be associated with relatively complex melodies and rhythms.

For pitch range, we predicted that ragas perceived as having a wider pitch range would be associated with negative emotions. A number of studies have provided strong evidence that listeners expect pitch proximity in melodies (e.g., Cuddy & Lunney, 1995; Krumhansl, 1995a, 1995b; Narmour, 1990, 1991; Schellenberg 1996, 1997; Thompson, Cuddy & Plaus, 1997; Thompson & Stainton, 1998). Thus, ragas perceived to involve a wider range in pitch may be less predictable, yielding a higher level of arousal. In addition, ragas involving a narrow pitch range may be processed as one auditory stream (Bregman, 1990) and, hence, may be easier to process. Ragas that are both more predictable and processed as a single auditory stream should be assigned relatively high ratings for positive emotions (joy and peace) and relatively low ratings for negative emotions (anger and sadness).

Timbre has also been found to influence listeners' ratings of emotion in music. Behrens and Green (1993) presented listeners with improvisations intended to evoke specific emotions, performed on violin, timpani, and

voice. They reported that their listeners were more sensitive to sadness and fear expressed in violin and vocal improvisations and to anger expressed in timpani improvisations. Gabrielsson and Juslin (1996) include timbre in some of their psychophysical profiles: music perceived as happy was characterized as having a "bright" timbre, music perceived as angry was characterized as having a "harsh" timbre, and music perceived as having no expression was characterized having a "cold" timbre.

Because the alap samples recorded for this study were performed in two distinctive timbral classes (strings and flute), this variable was not rated by listeners but included as a coded variable. Interviews conducted during a field research trip to India with experts on Hindustani music indicated that the rasa of raudra/anger is most effectively expressed on traditional stringed instruments, particularly sitar, and it is almost never expressed on the bansuri flute (A. Ranade, S. Rao, and S. Devasthali, personal communication, December 1996 to January 1997). Therefore, we predicted that the flute timbre would be associated with higher ratings of peace and the string timbre would be associated with higher ratings of anger.

Method

LISTENERS

Fifteen men and 15 women from the York University community, ranging in age from 23 to 46 years, volunteered to participate in the study. All 30 were raised in a Western culture and exposed to music of the Western tonal system. Two of them reported a minor amount of exposure to Hindustani music as interested listeners but had no knowledge of the conventions of the Hindustani tonal system. The remaining 28 listeners indicated no familiarity with Hindustani music.

Four experts on Hindustani music also participated in the study. These listeners had either taught or studied Hindustani music for an average of 33 years each. The experts were asked to identify each raga in addition to specifying the dominant rasa. All of them asserted that the performances they heard were competent renditions of the ragas and that these ragas were associated with the intended rasas.

MATERIALS

Cassette recordings of the alap portions of 12 raga performances were used in this experiment. The alaps in this study were field recordings (obtained during a 1996–1997 field research trip in India) of performances by two professional Hindustani musicians: a bansuri flute player and a stringed instrument player (sitar, dilruba, and surbahar). Both performers had studied classical Hindustani music for over 20 years and had extensive experience playing before audiences and making professional recordings.

During the recording sessions (conducted separately), each performer was asked to choose several ragas they would normally choose to evoke each of four target emotions: joy, sadness, anger, and peace. Then they were asked to play only a short alap from each raga they had selected. The performers supplied the names of the ragas they had chosen, which were later verified for spelling and associated moods (Danielou, 1991; Kaufmann, 1968).

The performances were recorded on a portable AIWA cassette recorder. Three improvisations for each target emotion were recorded (for a total of 12) and dubbed onto separate cassettes for randomization of presentation to listeners. The length of these recordings ranged from 52.8 to 247.3 s (mean = 165.28 s). Table 1 lists the ragas selected by each of the two performers. It may be noted that the performers independently chose two of the same or similar ragas to convey the same emotions: *Bhupali* for joy and *Yaman/Yaman Kalyan* for peace.

Ragas were presented to listeners on a CS 703D AKAI compact cassette deck through Sennheiser HD 480 headphones. After each presentation, listeners' ratings were entered directly onto a Power Macintosh computer.

A brief demographic questionnaire was completed by each listener. Some of the items included were age, sex, cultural background, and familiarity with Western and Hindustani music.

PROCEDURE

Listeners were tested individually. After filling out a demographic questionnaire, they were told they would be hearing 12 ragas ranging in length from one to four minutes and that for each raga they should indicate which of the four target emotions was most dominant. The ragas were presented in a different order to each of the 30 listeners.

After identifying the dominant emotion in each piece, listeners were asked to rate the degree to which they believed that emotion was conveyed on a scale from one (emotion not conveyed) to nine (emotion strongly conveyed). Next, they were asked to rate each piece on similar scales for the other three target emotions, as well as for tempo, rhythmic complexity, melodic complexity, and pitch range. Listeners were given definitions and verbal examples of each of these psychophysical dimensions as follows:

For judgments of tempo, they were asked to think about the overall pace of the excerpt; was it very fast (a rating of 9) or very slow (a rating of 1) or somewhere in between? For judgments of melodic complexity, listeners were asked to think about how much was going on melodically in the excerpt; was there a lot of repetition as in "Mary had a Little Lamb"? Would it be easy to sing it back to themselves (very simple—a rating of 1)? Did the melody

TABLE 1
Musical Stimuli

Performer	Raga	Rasa	Instrument
1	Bhupali	Hasya	Bansuri flute
2	Khamaj	(Joy)	Sitar (stringed)
2	Bhupali		Surbahar (stringed)
2	Jogya	Karuna	Dilruba (stringed)
1	Bhairavi	(Sadness)	Bansuri flute
1	Bhopali-Todi		Bansuri flute
1	Hindol	Raudra	Bansuri flute
2	Adana	(Anger)	Sitar (stringed)
1	Sohini		Bansuri flute
1	Yaman Kalyan	Shanta	Bansuri flute
2	Yaman	(Peace)	Surbahar (stringed)
2	Bhilaskhani		Dilruba (stringed)

change a great deal and go in unexpected directions; would it be difficult to sing it back to themselves (very complex—a rating of 9)?

For judgments of rhythmic complexity, listeners were asked to imagine a drummer accompanying the excerpt. Would the drummer be doing something very repetitious and easy to follow (very simple—a rating of 1), or would the drummer be doing something very changeable and difficult to follow (very complex—a rating of 9)? For pitch range, listeners were asked to think about how low and how high the melody ranged. Did it stay within the same few notes as in "Mary had a Little Lamb" (a rating of 1), or did it go very high and very low like an Italian aria (a rating of 9)? They were encouraged to ask questions and seek clarification of any terms of which they were unsure.

Results and Discussion

Table 2 illustrates the correlations between mean ratings of the four target emotions in the 12 raga alaps. There were two significant correlations—between joy and sadness and between anger and peace. These negative relationships were expected: joy and sadness were perceived as opposing qualities, as were anger and peace.

Table 3 displays mean ratings of each emotion for ragas intended to convey joy, sadness, anger, and peace. The highest ratings for joy and sadness were, as predicted, assigned to joyful and sad ragas, respectively. However, the results were less clear for ratings of anger and peace. Repeated measures analyses of variance were conducted for each emotion scale, to evaluate whether the mean ratings for each emotion differed significantly among the 12 ragas. There was a significant main effect of raga for joy, sadness, anger, and peace ($p < .001$ for each). These findings indicate that listeners reliably differentiated between the ragas on the basis of emotional content, even though they had no familiarity with North Indian music.

TABLE 2
Correlations of Mean Ratings of Emotions Perceived in Hindustani Ragas

	Joy	Sadness	Anger	Peace
Joy	1.000	−0.96	−0.32	−0.37
	$p = .$	$p = .000$	$p = .303$	$p = .233$
Sadness	−0.96	1.000	0.27	0.38
	$p = .000$	$p = .$	$p = .396$	$p = .218$
Anger	−0.32	0.27	1.000	−0.64
	$p = .303$	$p = .396$	$p = .$	$p = .024$
Peace	−0.37	0.38	−0.64	1.000
	$p = .233$	$p = .218$	$p = .024$	$p = .$

NOTE—The values represent correlations between mean ratings of the 12 ragas.

<div align="center">

TABLE 3

Mean Ratings of Emotions Perceived in Hindustani Ragas

</div>

Emotion	Type of Melody in Terms of Intended Emotion			
Rated	Joy	Sadness	Anger	Peace
Joy	5.61	3.12	2.30	4.40
SE	(0.41)	(0.34)	(0.38)	(0.35)
Sadness	2.99	5.27	2.75	5.06
SE	(0.38)	(0.39)	(0.36)	(0.40)
Anger	5.12	3.54	3.37	3.69
SE	(0.47)	(0.41)	(0.46)	(0.39)
Peace	2.59	5.58	3.50	4.61
SE	(0.35)	(0.39)	(0.39)	(0.48)

NOTE—The values represent means across ragas and listeners. Values in parentheses are mean standard errors.

Contrast analyses were conducted to assess the prediction that mean ratings of a given emotion would be significantly higher for the ragas that were intended to convey that emotion than for ragas not intended to convey that emotion. The mean ratings for the three ragas corresponding to the rated emotion were compared with the nine ragas that were intended to convey a different emotion. These analyses indicated that listeners were sensitive to the intended emotion in ragas when that emotion was joy, sadness, or anger, but not when that emotion was peace.

Table 4 illustrates these results. First, the mean rating of joy for the three ragas intended to convey joy was higher than the mean rating for the nine ragas not intended to convey joy. Second, the mean rating of sadness for the three ragas intended to convey sadness was significantly higher than the mean rating for the nine ragas not intended to convey sadness. Third, the mean rating of anger for the three ragas intended to convey anger was significantly higher than the mean rating for the nine ragas not intended to convey anger. Finally, the mean rating of peace for the three ragas intended to convey peace was not significantly higher than the mean rating for the nine ragas not intended to convey peace. For three of the four emotions considered, these results support the hypothesis that people enculturated to one tonal system can accurately perceive the intended emotion in music of an unfamiliar tonal system.

An analysis of sex revealed that women were significantly more sensitive to the emotion of joy in the ragas, $F(1,28) = 6.54, p < .05$. No other differences between sexes were observed.

The next set of analyses addressed our second research question: Were listeners' ratings of emotion in the ragas influenced by their perception of

TABLE 4

Mean Ratings of Emotion for Ragas Intended to Convey Specific Emotions vs Ragas Intended to Convey Other Emotions

Emotion Rated	Mean Rating of Ragas		Mean Contrast Results	
	Intended (n = 3)	Other (n = 9)	F	P
Joy	5.61 (0.41)	3.57 (0.40)	60.43	.0001
Sadness	5.27 (0.39)	4.80 (0.40)	23.45	.0001
Anger	3.37 (0.46)	2.85 (0.38)	4.41	.0366
Peace	4.61 (0.48)	4.45 (0.42)	.70	.4032

NOTE—Values in parentheses are mean standard errors.

psychophysical cues? To evaluate this question, we entered mean ratings for each of the four rated psychophysical cues into stepwise multiple regressions for each emotion. One additional (unrated) predictor, timbre, was also included in the analysis. Timbre was coded as a dichotomous variable: ragas performed on a stringed instrument were coded as 0, and ragas performed on the flute were coded as 1. Tables 5–8 show, for each of the four emotions, the correlation and regression results.

Table 5 reveals that ratings of joy were strongly associated with ratings of tempo, rhythmic complexity, and melodic complexity, but not with pitch range or timbre. Stepwise multiple regression revealed that tempo and melodic complexity provided unique predictive power. As predicted, ratings

TABLE 5

Mean Joy Ratings: Correlations and Stepwise Multiple Regression

	Correlations				
	Tempo	Rhythmic Complexity	Melodic Complexity	Pitch Range	Timbre
Joy	.85	–.76	–.66	.44	.13
p	.000	.002	.010	.075	.685

Regression
$R^2 = .91$
Adjusted $R^2 = .89$
$F = 47.39$; Signif $F = .0001$

Variables in the Equation

Variable	B	SE B	Beta	T	Sig T
Tempo	1.20	.18	1.28	6.85	.0001
Melodic complexity	–.65	.30	–.40	–2.17	.0578

TABLE 6
Mean Sadness Ratings: Correlations and Stepwise Multiple Regression

	Correlations				
	Tempo	Rhythmic Complexity	Melodic Complexity	Pitch Range	Timbre
Sadness	−.92	.75	.64	−.53	−.003
p	.000	.003	.012	.039	.990

Regression
$R^2 = .91$
Adjusted $R^2 = .89$
$F = 44.19$; Signif $F = .0001$

Variables in the Equation

Variable	B	SE B	Beta	T	Sig T
Tempo	−.98	.14	−1.33	−6.93	.0001
Melodic complexity	.61	.24	.49	2.53	.0321

of joy were positively related to ratings of tempo. These results indicate that mean joy ratings can be predicted by a weighted sum of mean tempo ratings and mean melodic complexity ratings. Together, these two variables accounted for 91% of the variance in mean joy ratings.

Ratings of sadness were significantly correlated with ratings of tempo, rhythmic complexity, melodic complexity, and pitch range, but not timbre.

TABLE 7
Mean Anger Ratings: Correlations and Stepwise Multiple Regression

	Correlations				
	Tempo	Rhythmic Complexity	Melodic Complexity	Pich Range	Timbre
Anger	.007	.27	.32	.12	.56
p	.492	.203	.155	.356	.03

Regression
$R^2 = .31$
Adjusted $R^2 = .24$
$F = 4.51$; Signif $F = .0596$

Variables in the Equation

Variable	B	SE B	Beta	T	Sig T
Timbre	1.07	.50	.56	2.12	.0596

56   Laura-Lee Balkwill & William Forde Thompson

TABLE 8
Mean Peace Ratings: Correlations and Stepwise Multiple Regression

		Correlations			
	Tempo	Rhythmic Complexity	Melodic Complexity	Pitch Range	Timbre
Peace	−.62	−.77	−.69	−.40	−.55
p	.015	.002	.007	.099	.031

Regression
R^2 = .74
Adjusted R^2 = .68
F = 12.48; Signif F = .0025

Variables in the Equation

Variable	B	SE B	Beta	T	Sig T
Rhythm	−.52	.13	−.68	−3.88	.0037
Timbre	−.63	.29	−.38	−2.25	.0607

As expected, ratings of sadness had a strong negative relationship with ratings of tempo. Mean tempo ratings and mean melodic complexity ratings contributed unique predictive power in the regression equation. These two predictors were highly significant and together accounted for 91% of the variance in mean ratings of sadness.

None of the listener-rated dimensions were significantly correlated with ratings of anger. Timbre, however, was significantly correlated with these ratings. There was a positive relationship between ratings of anger and the ragas performed on stringed instruments. In the stepwise multiple regression, timbre was the only significant predictor of mean anger ratings, accounting for 31% of the variance.

Mean ratings of peace were significantly correlated with mean ratings of tempo, rhythmic complexity, melodic complexity and timbre, but not pitch range. As predicted, high ratings of peace were associated with low ratings of tempo. In the stepwise regression, mean ratings of rhythmic complexity and timbre contributed significant predictive power in the equation. These two variables accounted for 73% of the variance in mean ratings of peacefulness.

Although initial results indicated that Western listeners were not reliably sensitive to the intended emotion of peacefulness (see Table 4), it appears that simpler rhythms and flute timbre were used by listeners as psychophysical cues in assigning ratings of peacefulness in the ragas.

EXPERT DATA

Table 9 shows a comparison between mean ratings, made by the 4 expert and 30 naive listeners, of ragas intended to convey each emotion and the ragas not intended to convey that emotion. It should be noted that there is a similar pattern of mean ratings for expert and naive listeners: for all four emotions, both groups assigned higher ratings to ragas intended to convey that emotion than to ragas not intended to convey that emotion. Experts assigned relatively low ratings of anger overall. This observation is consistent with the North Indian view, previously noted, that anger is rarely employed in ragas because of its incongruency with the peaceful nature of Hindustani music.

The correlation between mean ratings assigned by expert and naive listener ratings was significant for joy, $r(10) = .61, p = .036$; and sadness $r(10) = .83. p = .001$; but not for anger, $r(10) = .36, p = .253$; and peace, $r(10) = .47, p = .12$. A closer look at the differences between mean ratings for each emotion indicates that experts made somewhat clearer distinctions than naive listeners. Given the experts' knowledge of the tonal conventions of Hindustani music and the raga-rasa system, this result is consistent with our model.

General Discussion

The results of this study support our model of the perception of emotion in music. Western listeners were sensitive to intended emotions in Hindustani ragas, even though they were unfamiliar with the tonal system and the raga-rasa system of conveying moods within that tonal system. Furthermore, listeners' ratings of the psychophysical dimensions in the ragas were predictive of their judgments of emotions in the ragas.

TABLE 9
Mean Ratings of Emotion in Ragas by Naive and Expert Listeners

Emotion Rated	Naive Group		Expert Group	
	Intended (3)	Other (9)	Intended (3)	Other (9)
Joy	5.61 (0.41)	3.57 (0.40)	6.25 (0.76)	3.25 (0.37)
Sad	5.27 (0.40)	4.80 (0.40)	6.17 (0.68)	4.28 (0.71)
Anger	3.37 (0.46)	2.85 (0.38)	2.5 (0.25)	1.00 (0.00)
Peace	4.61 (0.48)	4.45 (0.42)	6.25 (0.38)	5.06 (0.37)

NOTE—Ratings were made on scales numbered one to nine. Values in brackets are mean standard errors.

    In the absence of relevant culture-specific knowledge regarding the conventional associations of emotion and music in the Hindustani tonal system, listeners relied upon the only resource at their disposal: psychophysical cues. This strategy succeeded for the perception of joy, sadness, and anger but not for the perception of peacefulness. According to our model, this indicates that the psychophysical cues for joy, sadness, and anger were salient enough to enable listeners to overcome their unfamiliarity with culture-specific cues and to make an accurate assessment of the intended emotion. Particularly in the case of ragas intended to represent joy and sadness, ratings made by naive listeners were strongly correlated with ratings made by expert listeners. That naive listeners demonstrated such a high level of agreement with expert listeners, who were deeply familiar with the culture-specific cues embedded in the music samples, is remarkable.

    The current results are compatible with the results of studies in which listeners judged emotion in music from their own culture. Gabrielsson and Juslin (1996) found fast tempi were associated with listeners' perceptions of joyfulness and anger, and slow tempi were associated with listeners' perceptions of sadness and peacefulness. For ragas intended to evoke joy and sadness, the same pattern was observed. In this investigation, 73% of the variance in mean ratings of joy and 84% of the variance in mean ratings of sadness could be accounted for by mean ratings of tempo. These results illustrate that listeners were able to assess tempo in the Hindustani ragas and that a strong relationship exists between the perception of tempo and the interpretation of emotion in music.

    Listeners also perceived rhythmic and melodic complexity in the ragas, and their ratings of these elements were related to their ratings of joy, sadness, and peacefulness. Perceptions of joy were associated with perceptions of simpler melodies, and perceptions of sadness were associated with perceptions of more complex melodies. As with tempo, these findings parallel the results of studies involving Western listeners and Western music. Hevner (1935, 1936) examined listeners' perceptions of rhythm and harmonic complexity, as they related to their judgments of musically expressed emotion. She found that listeners associated simple harmony with happiness and serenity and associated complex harmony with excitement or sadness. "Firm" rhythm was associated with "dignity and vigor" and "flowing" rhythm was associated with "happy feelings." Rigg (1940) investigated structural elements of music and their effect upon listeners' judgments of emotion in music. He found that listeners linked perceptions of joy with faster tempi, simple harmony, iambic or anapestic rhythms, and perceptions of "lamentation" with slow tempi, trochaic rhythms, dissonance, and legato phrasing. Imberty (1974) evaluated the music of Debussy according to an index of formal complexity and another index of dynamism. In his study, high complexity was related to responses of "a melancholic or sorrow kind," and low complexity was related to responses of "joy and well-

being" (see Nielzén & Cesarec, 1982, for a review). Our findings illustrate that these relationships between emotion judgments and psychophysical dimensions of music is preserved across cultural boundaries.

Tempo was not a significant predictor of mean ratings of peacefulness or anger. It is possible that, because both peaceful and sad ragas employed slow tempi, listeners relied on other cues to distinguish between them. In particular, ratings of rhythmic complexity were significantly related to ratings of peacefulness, with simple rhythms associated with greater peacefulness. However, listeners were unsuccessful overall in their attempts to distinguish peaceful ragas from other ragas. Thus, it appears that the psychophysical cue of rhythmic complexity was insufficient to distinguish peaceful ragas from other ragas.

Timbre was a significant predictor of mean ratings of peacefulness and anger. The direction of this influence was expected: the flute timbre was strongly associated with peacefulness, and the string timbre was strongly associated with anger. However, as with rhythmic complexity, the psychophysical cue of timbre was insufficient to distinguish peaceful ragas from other ragas.

Pitch range failed to emerge as a strong predictor of listeners' ratings of any of the four target emotions. Previous evidence has indicated that pitch range plays a role in expectancy (Narmour, 1991), and expectancy, in turn, has been linked to emotion (Meyer, 1956). Quite possibly, however, the "emotions" associated with expectancy mechanisms may be generalized arousal, rather than specifiable emotions. This interpretation would explain why pitch range did not appear to influence Western listeners' identification of specific musically expressed emotion in Hindustani ragas.

OTHER MODELS OF PERCEPTION OF EMOTION IN MUSIC

The proposed model assumes that, in addition to culture-specific cues, there are psychophysical cues (common to all auditory signals) that influence listeners' perceptions of emotion in music. The connection between music and emotion has been explored in a number of other theories that we will now discuss and compare with our own theory.

One of the more influential theories of emotion in music is that of Suzanne Langer. In her philosophical treatise of music and emotion (1957), she develops the idea that "the forms of human feeling are much more congruent with musical forms than with the forms of language. . . ." This idea stems from a rejection of previous theories that sought to explain the emotive character of music solely as a result of music's mimicry of human vocal expressions of emotion.

Langer does not focus on the issue of whether the perception of emotion in music requires an understanding of specific musical idioms. Neither does she explore whether more elementary properties in the music (e.g.,

psychophysical ones such as changes in loudness or tempo) might map directly onto emotional content quite apart from the culturally specific codes and traditions of the music.

Langer rejects the idea that music can represent specific emotions, arguing that music imitates only the form of emotion, not the content. For example, music may be very emotional, but equally happy and sad. This view suggests that the representation of emotion in music is inherently vague. However, it has difficulty accounting for findings that listeners readily attach emotional labels to pieces of music (Francès, 1958, Expt. 14; Hevner, 1935, 1936).

Mandler (1984) proposed a theory of emotion in music based on Meyer's earlier work on musical expectancies. Mandler's account draws a connection between the emotional qualities of music and the biological adaptive value of emotional reactions. Specifically, the interruption of ongoing schemata, that is, a violation of our expectancies, brings about biological arousal, which acts as an alerting signal. This alerting signal then stimulates an interpretation of the stimuli, which, when combined with the biological arousal, is felt as emotion.

Similar to Langer's theory, Mandler's theory holds that music does not signify specific emotions. Fulfillments and violations of musical expectancies lead to biological arousal, which may lead to an emotional interpretation. However, biological arousal is essentially contentless, so emotional qualities cannot be inferred directly from the properties of arousal. The content of the emotional reaction is an interpretation that occurs post hoc.

In contrast to these views are theories of emotion in music that hold that music can represent specific emotions (Cooke, 1959; Kivy, 1980). Cooke set out to analyze elements of musical expression and to establish a lexicon of these elements in terms of the specific emotions they were able to convey. He attempted to test this theory empirically and took as his material Western tonal music composed since 1400 A.D. He identified musical patterns of sound that were commonly used to express emotion in this body of work. He then analyzed these melodic motifs into their component intervals and asserted that each "basic term" had a particular emotive character (i.e., scale notes 5-1-2-3 in a minor mode = a bold acknowledgment of the existence of tragedy). Efforts to substantiate this theory empirically have been unsuccessful. Gabriel (1978) asked listeners to rate the emotive quality of Cooke's "basic terms" and found no significant pattern of agreement with Cooke's descriptions of their emotional content.

Kivy's theory (1980) also centers on the view that music can convey specific emotions to the listener. Like Cooke (1959), Kivy maintained that certain melodic patterns are conventionally associated with specifiable expressive predicates, such as "is sad," "is angry," or "is joyful". However, he acknowledged that some links between music and emotion may be "natural."

Kivy proposed a distinction between two types of cues to emotion in music: "contour" and "convention." In his discussion of contour, he identified elements of music that he characterized as having natural connections to "emotive life" (emotion as expressed by vocal utterances, posture, gesture, and bodily motion), such as melodic contour, tempo, and rhythm. These elements were also featured in his discussion of convention. Falling melodic lines signifying sadness, leaping melodic motion signifying joy, and slow tempo with measured rhythm signifying dignity were all conventions that Kivy identified as being typical of Western classical music.

According to Kivy, all conventional associations between music and emotion ultimately derive from natural connections between music and "emotive life." In turn, he stated that these natural connections are culture-specific, in that the music of a culture can resemble the emotive life of only that culture. Thus, listeners from one culture should be unable to appreciate the natural connections between music and the emotive life of an unfamiliar culture. Kivy further argued that listeners of one culture would be unable to perceive music from an unfamiliar tonal system as music per se, but rather as a sequence of nonmusical sounds.

To illustrate his point, he postulated a cross-cultural exploration of his theory, using the example of North Indian raga music. "The Western ear, nourished on horizontal or (at least) vertical polyphony (i.e., harmony) will naturally tend to read Indian music harmonically or polyphonically and, of course, fail" (Kivy, 1980, p. 90). According to Kivy's predictions, Western listeners should hear ragas not as music, but rather as an unintelligible collection of alien sounds. "To the uninitiated ear, every raga presents about the same mood: a kind of exotic stupor" (Kivy, 1980, p. 89).

Kivy offered no empirical evidence to support his predictions about Western listeners' inability to perceive emotion in the music of North India. The results of our investigation contradict Kivy's assumptions. Western listeners were able to perceive emotional expressiveness in the music of an unfamiliar tonal system, and further evidence suggested that their perceptions were based on psychophysical elements such as tempo, timbre, and complexity.

Our model provides a useful framework for exploring the connection between psychophysical dimensions and the representation of specific emotions in music. It may be that listeners' sensitivity to emotion in music is a highly evolved aspect of human sensitivity to psychophysical dimensions of our auditory environment. Only by stripping away the layers of culture-specific cues and interpretations can we further explore these possibilities.[4]

References

Behrens, G. A., & Green, S. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music, 21,* 20–33.

Berlyne, D. E. (1971). *Aesthetics and psychobiology.* New York: Appleton Century Crofts.

Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge, MA: The MIT Press.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113,* 394–412.

Cooke, D. (1959). *The language of music.* London: Oxford University Press.

Crozier, J. B. (1974). Verbal and exploratory responses to sound sequences varying in uncertainty level. In: Berlyne, D. E. (Ed.), *Studies in the new experimental aesthetic: Steps toward an objective psychology of aesthetics appreciation.* Washington, DC: Hemisphere Publishing Corp.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception & Psychophysics, 57,* 451–462.

Danielou, A. (1991).*The ragas of northern Indian music.* New Delhi: Munshiram Manoharlal Pub.

Day, H. I. (1967). Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. *Perception & Psychophysics, 2,* 281–286.

Day, H. I. (1968). The importance of symmetry and complexity in the evaluation of complexity, interest and pleasingness. *Psychonomic Science, 10,* 339–340.

Deva, B. C., & Virmani, K. G. (1975). A study in the psychological response to ragas. *Research Report II of Sangeet Natak Akademi.* New Delhi: Indian Musicological Society.

Dolgin, K. G., & Adelson, E. H. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music, 18,* 87–98.

Dowling, W. J., & Harwood, D. L. (Eds.). (1986). *Music cognition.* New York: Academic Press.

Feld, S., & Keil, C. (1994). *Music grooves: Essays and dialogues.* Chicago: University of Chicago Press.

Francès, R. (1988). *The perception of music* (W. J. Dowling, Trans.). Hillsdale, NJ: Erlbaum. (Original work published 1958)

Gabriel, C. (1978). An experimental study of Deryck Cooke's theory of music and meaning. *Psychology of Music, 6,* 13–20.

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.

Gerardi, G. M., & Gerken, L. (1995). The development of affective responses to modality and melodic contour. *Music Perception, 12,* 279–290.

Gregory, A. H., & Varney, N. (1996). Cross-cultural comparisons in the affective response to music. *Psychology of Music, 24,* 47–52.

Haack, P. A. (1980). The behavior of music listeners. In D. Hodges (Ed.), *Handbook of music psychology* (pp. 141–182). Lawrence, KS: National Association for Music Therapy.

Hevner, K. (1935). Expression in music: A discussion of experimental studies and theories. *Psychological Review, 42,* 186–204.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48,* 246–268.

Holbrook, M. B., & Anand, P. (1990). Effects of tempo and situational arousal on the listener's perceptual and affective responses to music. *Psychology of Music, 18,* 150–162.

Hoshino, E. (1996). The feeling of musical mode and its emotional character in a melody. *Psychology of Music, 24,* 29–46.

Imberty, M. (1974). Perspectives nouvelles de la sémantique musicale expérimentale. *Musique en Jeu, 17,* 87–109

Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception, 8,* 189–202.

Kaufmann, W. (1968). *The ragas of north India.* Indianapolis: Indiana University Press.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception, 2,* 131–165.

Kivy, P. (1980). *The corded shell: Reflections on musical expression.* Princeton, NJ: Princeton University Press.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music, 21,* 3–19.

Krumhansl, C. L. (1995a). Effects of musical context on similarity and expectancy. *Systematische Musikwissenschaft. (Systematic Musicology), 3,* 211–250.

Krumhansl, C. L. (1995b). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum, 17,* 53–80.

Langer, S. K. (1957). *Philosophy in a new key (3rd ed.).* New York: New American Library.

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress.* New York: Norton.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures.* Chicago: University of Chicago Press.

Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception, 9,* 1–26.

Nielzén, S., & Cesarec, Z. (1982). Emotional experience of music as a function of musical structure. *Psychology of Music, 10,* 7–17.

Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research, 57,* 103–118.

Rigg, M. G. (1940). Speed as a determiner of musical mood. *Journal of Experimental Psychology, 27,* 566–571.

Robazza, C., Macluso, C., & D'Urso, V. (1994). Emotional reactions to music by gender, age, and expertise. *Perceptual and Motor Skills, 79,* 939–944.

Schellenberg, E. G. (1996). Expectancy in melody: Test of the implication-realization model. *Cognition, 58,* 75–125.

Schellenberg, E. G. (1997). Simplifying the implication-realization model. *Music Perception, 14,* 295–318.

Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion, 1,* 331–346.

Terwogt, M., & van Grinsven, F. (1991). Musical expression of moodstates. *Psychology of Music, 19,* 99–109.

Thompson, W. F., Cuddy, L. L., & Plaus, C. (1997). Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody production task. *Perception & Psychophysics. 59,* 1069–1076.

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts, 10,* 79–89.

Thompson, W. F., & Stainton, M. (1998). Expectancy in Bohemian folk song melodies: Evaluation of implicative principles for implicative and closural intervals. *Music Perception, 15,* 231–252.

Vitz, P. C. (1966a). Affect as a function of stimulus variation. *Journal of Experimental Psychology, 71,* 74–79.

Vitz, P. C. (1966b). Preference for different amounts of visual complexity. *Behavioral Science, 11,* 105–114.

Walker, R. (1996). Open peer commentary: Can we understand the music of another cul-
    ture? *Psychology of Music, 24,* 103–130
Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music.
    Scandinavian Journal of Psychology, 13, 241–257.