MEMORY AND MUSICAL EXPECTATION FOR TONES IN CULTURAL CONTEXT

MEAGAN E. CURTIS
Dartmouth College

JAMSHED J. BHARUCHA
Tufts University

WE EXPLORED HOW MUSICAL CULTURE SHAPES ONE'S listening experience. Western participants heard a series of tones drawn from either the Western major mode (culturally familiar) or the Indian *thaat* Bhairav (culturally unfamiliar) and then heard a test tone. They made a speeded judgment about whether the test tone was present in the prior series of tones. Interactions between mode (Western or Indian) and test tone type (congruous or incongruous) reflect the utilization of Western modal knowledge to make judgments about the test tones. False alarm rates were higher for test tones congruent with the major mode than for test tones congruent with Bhairav. In contrast, false alarm rates were lower for test tones incongruent with the major mode than for test tones incongruent with Bhairav. These findings suggest that one's internalized cultural knowledge may drive musical expectancies when listening to music of an unfamiliar modal system.

Received May 19, 2008, accepted December 18, 2008.

Key words: acculturation, cross cultural differences, false memory, music perception, neural networks

THE OFT-PURPORTED NOTION THAT MUSIC IS A universal language may reflect the ubiquity of the Western tonal system throughout the world, linking cultures when languages fail. However, when one encounters an unfamiliar tonal system, music may fail to attain the status of a universal language and may instead be perceived as strange and foreign. We investigated the musical expectations that develop as one listens to an unfamiliar modality. We hypothesized that contrary to the universal language viewpoint, culturally unfamiliar modalities are perceived through the framework of the cultural system with which one is already familiar.

The acquisition of musical knowledge begins at an early age. Infants have been shown to have relative pitch based representations of melodies by six months of age (Cohen, Thorpe, & Trehub, 1987; Trainor, Wu, & Tsang, 2004; Trainor & Trehub, 1992; Trehub, 2001; Trehub & Trainor, 1993). They exhibit knowledge of modal structure by 5 to 6 years of age (Cuddy & Badertscher, 1987; Dowling, 1999; Krumhansl & Keil, 1982; Trainor & Trehub, 1994) and exhibit knowledge of harmonic structure at the age of seven years (Trainor & Trehub, 1994). Listeners acquire knowledge of modal systems through passive perceptual learning (Justus & Bharucha, 2001; Tillmann, Bharucha, & Bigand, 2000). The implicit knowledge of a modal system drives a listener's expectations for future musical events (Bharucha & Stoeckig, 1986, 1987; Justus & Bharucha, 2001). These expectations stem from the learned associations between musical events. For instance, a listener who is familiar with the Western modal system knows implicitly that in the major mode, Do, Mi, Fa, Sol, and Ti co-occur with Re and La. Even when Re and La are not presented in the musical stimulus, top-down processing completes the well-known mode pattern and activates the representation of the missing tones (Tillmann, Bharucha, & Bigand, 2000).

When listening to music from a culturally unfamiliar system, listeners may be able to enjoy the music. However, they may perceive certain tones as sounding unexpected or off. These tones may sound strange because they violate the listener's cultural-specific modal expectations and may thus be perceived as wrong notes.

Similarities between modal systems may actually serve to perceptually exaggerate the disparities between the modal systems. If an unfamiliar system shares several scale degrees with a familiar system, the shared scale degrees may activate the representation of the familiar system and its accompanying expectations. When a tone in the unfamiliar system deviates from those in the familiar system, the listener may perceive the deviant tone as a violation of the familiar system, thus increasing the perceptual salience of that tone relative to others.

Several studies have looked at cross-cultural perception. Castellano, Bharucha, and Krumhansl (1984) played brief passages based on each of ten North Indian thaats and asked participants to rate a probe

Music Perception VOLUME 26, ISSUE 4, PP. 365–375, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2009.26.4.365

tone following each segment. Both Indian and Western listeners gave probe tone ratings that correlated highly with the distribution of tones in the prior context, suggesting that this distribution is integrated in the short term regardless of cultural background. However, there also was evidence that for the Indian participants, ratings were influenced by prior exposure over and above the proximal distribution of tones.

Krumhansl, Louhivuori, Toiviainen, Järvinen, and Eerola (1999) examined the melodic expectations generated by Finnish folk hymns in participants who either were familiar or unfamiliar with the hymns. Participants were asked to listen to short melodies that stopped in the middle of a phrase and to then rate how well a probe tone fit their musical expectations of what should follow. Although both groups of participants were sensitive to the general distribution of tones in the hymns, those familiar with the hymns made melodic continuation judgments that reflected a higher degree of knowledge about the transitional probability between tones.

Krumhansl et al. (2000) examined the melodic expectations generated by North Sami yoiks, which are vocal melodies composed of tones from the major pentatonic scale. Melodic continuation judgments were made by Sami participants who were highly familiar with the yoiks, Finnish participants who had studied the yoiks, and Western participants who were unfamiliar with the yoiks. Relatively high intergroup correlations revealed similarities in the ratings made across the three groups of participants. However, there were some subtle differences between groups, and the authors determined that the judgments of the Western participants were most strongly influenced by schematic Western tonal knowledge, while the judgments of the Samis were least influenced by schematic Western knowledge. Veridical melodic expectations were highest for the Sami group. The results were tested against four versions of the implication-realization model (originally proposed by Narmour, 1990), revealing that 3 of the 4 models had a Western tonal bias. A model proposed by Krumhansl (1995) best fit the data across cultures.

The previously described experiments revealed that participants who were unfamiliar with a particular modal system acquired short-term statistical knowledge of the tone distributions of the unfamiliar modal systems, resulting in probe tone judgments that were similar to those made by native listeners. In this paper, we are interested in discerning possible cultural divergences. In order to get at underlying mechanisms, we adopted memory and reaction time judgments rather than ratings.

Our paradigm is analogous to the recognition memory paradigm used by those studying false memories (e.g., Roediger & McDermott, 1995). In the memory paradigm, a participant is presented with a list of semantically related words, such as bed, rest, tired, dream, night, etc., and is given a subsequent recognition memory test. Participants have been shown to make false alarms to words that are semantically associated with the encoded list, even when those words weren't presented at encoding. For instance, participants who encoded the above list would be likely to false alarm on the word sleep, since the word would have been activated by the list of encoded words due to their semantic relationship.

Our paradigm was designed to test the strength of schematic musical knowledge by the same means that semantic knowledge is tested in the aforementioned paradigm. The internalized knowledge of modes and the automatic activation of tones in a mode can be assessed with a speeded memory probe task. Short modal tone sets were presented for encoding, but a musically related tone was missing. After a brief pause, the listener was presented with a test tone and was asked to judge, as quickly as possible, whether the test tone had been presented in the previous set. The test tone corresponded to either the tone that was musically related, but missing from the tone set, a tone that was present in the tone set, or a tone that was musically unrelated to the tone set. We hypothesized that participants would false alarm when the test tone was musically related to the tone set. However, our participants were not culturally familiar with all of the tone sets we used as stimuli. We predicted that while participants should be likely to false alarm in the condition described above, they would be unlikely to false alarm to test tones that were musically related to a tone set if the participants were unfamiliar with the musical culture from which the particular tone set was derived.

We constructed short tone sets using one exemplar mode from each of two different cultural systems. All of our participants were highly familiar with Western modal music and relatively unfamiliar with Indian modal music. We used the Western major mode, <Do, Re, Mi, Fa, Sol, La, Ti>, as the familiar modal set, and the Indian thaat Bhairav, <Do, Re-, Mi, Fa, Sol, La-, Ti>, as the unfamiliar modal set. (In our notation, we employ "-" to mean "flat"). The modes differ at two scale degrees: the 2nd and 6th scale degrees of Bhairav are a semitone lower than the corresponding degrees of the Western major mode. For example, if Do is the tonal center of both modes, Do, Mi, Fa, Sol, and Ti are common to both modes. The major mode also contains Re and La, while Bhairav contains Re- and La-.

If a musical stimulus from a culturally unfamiliar mode activates the listener's familiar system, judgments should reflect typical features of the familiar system. For instance, when presented with Do, Re-, Mi, Fa, Sol, and Ti from Bhairav, Western listeners should not have enough exposure to Bhairav to fill in the missing tone of the mode, which is La-. If presented with La- as a test tone, Western listeners should be likely to reject the tone, since it is inconsistent with their internalized cultural knowledge. However, if presented with La as a test tone, Western listeners should be likely to make false alarms, since La would co-occur with the notes Do, Mi, Fa, Sol, and Ti in their familiar system. Although Re- in the original stimulus is inconsistent with the familiar mode, this inconsistent input should not override the activation of the familiar system.

These predictions stem from a set of models that simulate the passive perceptual learning of modal patterns. This class of models simulates the acquisition of musical regularities by internalizing the statistical patterns in the corpus of music to which they are exposed. These models account for expectations or filling-in of typical but missing events by matching heard patterns with learned ones (Bharucha, 1991, 1999; Bharucha & Todd, 1991). We tested the experimental paradigm in neural network models of Western listeners. The results of the simulations are described below.

Neural Network Simulations

Two models with different architectures converge on similar predictions about the pattern of activation in response to the tone sets described above, pointing to the stability of these predictions. One is an auto-associative network (Bharucha & Olney, 1989), and the other a self-organizing network. In both cases, the predictions are driven by the principle that when selected sets of tones are used pervasively, the brain learns the set as a Gestalt; subsequently, when unique subsets are heard, the missing tones are filled in.

In the auto-associative network (Bharucha & Olney, 1989), tone units representing the 12 chromatic scale degrees are activated by the scale degrees that occur in the music. The twelve scale degrees are <Do, Re-, Re, Mi-, Mi, Fa, Fa+, Sol, La-, La, Ti-, Ti>, or in the parlance of Indian music, <Sa, Re-, Re, Ga-, Ga, Ma, Ma+, Pa, Dha-, Dha, Ni-, Ni>. These units are fed by a process of transformation from absolute pitch activations to pitch-invariant activation patterns, gated by the tonal center as described by Bharucha and Mencl (1996). These tone units in turn are fully connected to tone units that represent expectation levels for each of these twelve scale degrees. The strengths of the connections are set initially at random, mimicking a tabula rasa prior to learning. When a set of tones is presented to the network, activation spreads to the expectation units. These expectation activations are compared to the sounded set of tones, and the connection strengths are changed according to the Delta Rule to reduce the difference. The Delta Rule is essentially a formalization of Hebbian learning (Hebb, 1949), according to which the connection between neurons is strengthened in proportion to the activations of the connected neurons. Repeated exposure to a few tone sets results in the network automatically learning to generate expectations consistent with those sets. Subsequent to learning, hearing a subset of one of those sets causes the network to "fill in" the missing tones at the expectation level.

An auto-associative network was trained on tone sets that represent either the Western major and minor modes or ten Indian modes (thaats). The Western network was trained with two input patterns: major <Do, Re, Mi, Fa, Sol, La, Ti> and minor <Do, Re, Mi-, Fa, Sol, La-, Ti>. The network required 26 cycles of training to learn the pattern sets. Once trained, the Western network was able to differentiate between major and minor modes when presented with impoverished input. For instance, when presented with <Do, Mi, Fa, Sol>, the layer of output nodes completed the pattern, activating <Do, Re, Mi, Fa, Sol, La, Ti>. The network expects <Do, Mi, Fa, Sol> to co-occur with Re, La, and Ti because it has learned that these tones typically co-occur. Similarly, the Indian network, once trained, fills in incomplete thaat patterns. For example, when presented with <Do, Mi, Fa, Sol, La-, Ti>, it fills in Re- in accord with the thaat Bhairav, which consists of <Do, Re-, Mi, Fa, Sol, La-, Ti>.

While the Western network reliably identifies the major and minor modes when presented with impoverished input patterns, we were interested in the network's performance when presented with an unfamiliar pitch set, the thaat Bhairav. When presented with Bhairav in its entirety, the network attempted to make sense of the thaat in terms of the major and minor modes, but was unable to clearly identify the input as being a major or minor mode. The stimulus input tones <Do, Re-, Mi, Fa, Sol, La-, Ti> caused the network to distribute activation evenly between the tones that differentiate the major and minor modes, because the input contained both Mi and La-, which would not co-occur in either the major or minor mode. Since Mi would co-occur with La (but not La-), and La- would co-occur with Mi- (but not Mi), the network's expectation (output) showed a distribution of activation that was approximately equal between Mi- and Mi, as well as between La- and La. The network was clearly trying to make sense of this stimulus in terms of what it had learned: Western major and minor modes. This interpretation was additionally supported by the high level of output activation for Re, which wasn't in the input stimulus, but occurred in both the major and minor modes that had been learned by the network. Re-, which was in the input stimulus, had virtually no activation in the expectation layer. Rather than confusing the network, this tone was seemingly ignored by the network, which tried to categorize the stimulus as either a major or minor mode.

To model the experiment, the Western network was presented with impoverished input stimuli corresponding to the Western major mode and to the thaat Bhairav. For the major mode, the input stimulus consisted of <Do, Re, Mi, Fa, Sol, Ti> or <Do, Mi, Fa, Sol, La, Ti>. The goal of using these impoverished stimuli was to gauge whether participants would be likely to make false alarms when presented with modally congruous tones that weren't in the stimulus.

The network's activation levels for these tests are shown in Table 1. For both stimuli, the omitted tones (La in the first stimulus, Re in the second stimulus) were highly expected. When La was omitted from the stimulus (case 1), its output activation was .79. When Re was omitted from the stimulus (case 2), its output activation was .99. Thus, the model predicted that human participants would be likely to make false alarms to missing tones that occur typically in a culturally familiar context, and that these tones would be primed as measured by reaction time.

To simulate the perception of Bhairav by the Western listener, the network was presented with two sets of impoverished stimuli: <Do, Re-, Mi, Fa, Sol, Ti> and <Do, Mi, Fa, Sol, La-, Ti> shown in Table 1 as case 3 and case 4 respectively. As predicted, in both cases the model's output showed only minimal activation of the missing tones (La- in the first stimulus, and Re- in the second stimulus). In case 3, the stimulus elicited high activation of La (.78), and minimal activation of La- (.19), even though neither of them occurred in the stimulus. In case 4 the stimulus elicited high activation of Re (.99) and minimal activation of Re- (.03), even though neither occurred in the stimulus. Also of interest is that La was more highly activated (.54) than La- (.43), even though the latter occurred in the stimulus and the former did not. This is because the presence of Mi (suggesting the major mode) caused an expectation for La.

In the modeling described above, the tone sets were presented to the network as if they had occurred together rather than in sequence. In earlier work (Bharucha, 1999; Bharucha & Todd, 1991) we have shown how tones from a mode or key can be integrated over time such that at any given time they can be represented as a single object capturing the sounded tones as they have decayed over time. Variability in the sequential orderings of the tones of a mode—as well as variability in the durations and repetitions of tones—is best estimated by representing them as equally strong, which is what we have done.

We predict that our Western participants will show strong expectations for Re and La, regardless of whether they hear a tone set suggesting major or Bhairav, as measured by a memory judgment and reaction time. When presented with either impoverished Indian stimulus (cases 3 and 4), participants should be likely to easily reject the test tones Re- and La- (even though these tones are congruous with the Indian pitch set), since these tones do not co-occur in either the familiar Western major or minor mode. Additionally, when presented with Re or La as a test tone, the Western participants should be likely to make errors and judge that the

TABLE 1. Auto-Associative Network Tests and Activations.

		Do	Re–	Re	Mi–	Mi	Fa	Fa+	Sol	La–	La	Ti–	Ti
		Sa	Re–	Re	Ga–	Ga	Ma	Ma+	Pa	Dha–	Dha	Ni–	Ni
1	Stimulus	1	0	1	0	1	1	0	1	0	0	0	1
	Expectation	.99	.03	.99	.27	.78	.99	.03	1	.26	.79	.02	.99
2	Stimulus	1	0	0	0	1	1	0	1	0	1	0	1
	Expectation	.99	.04	.99	.09	.92	.99	.03	.99	.08	.91	.03	.99
3	Stimulus	1	1	0	0	1	1	0	1	0	0	0	1
	Expectation	.98	.03	.99	.25	.76	.98	.03	.99	.19	.78	.03	.99
4	Stimulus	1	0	0	0	1	1	0	1	1	0	0	1
	Expectation	.99	.03	.99	.52	.58	.99	.03	.99	.43	.54	.03	.99

tones were present in the stimulus, since these tones do co-occur according to a Western tonal framework.

Acculturation vs. Psychoacoustics

Claims of acculturation in the perception of pitch patterns should always be tested against an alternative hypothesis that is entirely stimulus driven. While we predict that the test tone ratings will conform to the cultural learning hypothesis described above, it is possible that the judgments of the test tones could reflect the influence of the psychoacoustic properties of the musical stimuli. We refer to this as the *harmonic partials hypothesis,* because the partials in harmonic (i.e., integer multiple) relationship to the fundamental frequencies of the stimulus tones could prime the missing test tones, whose fundamentals may correspond to harmonics of the stimulus tones.

Tekman and Bharucha (1998) demonstrated that chords can be primed in ways that contradict psychoacoustic similarity but conform to cultural regularities. For example, the C major chord primes the D major chord more strongly than the E major chord, even though the C and E major chords share more harmonics in common than do the C and D major chords. The D major chord is more likely to occur with the C major chord than is the E major chord; thus, even though there is priming at the level of harmonics, the cultural regularity can override this effect if it is strong enough.

Is it possible that the harmonic partials of the tone sets may reinforce certain test tones more than others, leading to biases in favor of Re and La that override culturally familiar patterns? The only differences between the Western and Indian tone sets occurred at the 2nd and 6th scale degrees, Re-/Re and La-/La. We examined the first seven harmonics of Re-, Re, La-, and La, as most of the energy in complex harmonic tones occur in the lower order partials. The pitch classes of the first seven partials for Re-, Re, La-, and La are shown in Table 2. If these sets of partials exert an influence on the judgments in this experiment, we can make the following predictions about the patterns of results that should emerge under such a hypothesis. When the tone set contains Re-, the test tone La- should be reinforced relative to when the tone set contains Re, as the 3rd and 6th partials of Re- are frequencies that correspond to La-. This should lead to slower reaction times for correct rejections and higher false alarm rates for test tone La- in the Re- tone set condition, relative to the same measures for La- in the Re tone set condition. Given the prediction that La- should reinforce Re-, it should be more difficult for participants to correctly reject the test tone Re- when it has not actually occurred in the tone set. However, when the tone set contains Re, the test tone La should be reinforced relative to when the tone set contains Re-, as the 3rd and 6th partials of Re are frequencies that correspond to La. This should lead to slower reaction times for correct rejections and higher false alarm rates for test tone La in the Re tone set condition, relative to the Re- tone set condition. The test tone Re is not reinforced by the harmonic partials of La- or La, so no differences should emerge between those two conditions, which is a prediction that also is consistent with the predictions of the cultural learning hypothesis. When the tone set contains La, the test tone Re- should be reinforced relative to when the tone set contains La-, as the 5th partial of La is a frequency that corresponds to Re-. This should lead to slower reaction times for correct rejections and higher false alarm rates for test tone Re- in the La tone set condition relative to the La- tone set condition. These predictions generally differ from those of the cultural learning hypothesis, which predicts slower reaction times for correct rejections and higher false alarms to test tones Re and La than to Re- and La-, regardless of the tones that comprise the stimulus tone set, as Re and La should be more strongly activated by both the Western and Indian tone sets than test tones Re- and La-.

Method

Participants

Twenty-one college students participated in the experiment and received compensation in the form of course credit. Background questionnaires were administered to each participant by computer, but the background

TABLE 2. Harmonic Partials of Test Tones.

Tone	Partial 1	Partial 2	Partial 3	Partial 4	Partial 5	Partial 6	Partial 7
Re–	Re–	Re–	La–	Re–	Fa	La–	Ti
Re	Re	Re	La	Re	Fa+	La	Do
La–	La–	La–	Mi–	La–	Do	Mi–	Fa+
La	La	La	Mi	La	Re–	Mi	Sol

data for three of the participants could not be retrieved, due to software error. The background information for the remaining 18 participants is as follows. Thirteen participants were female and five were male. Their mean age was 18.56 years (*SD* = 0.92). None of the participants reported that they were familiar with Indian music. All participants were familiar with Western music, and they had a mean of 3.67 years of formal training on a musical instrument (*SD* = 3.80). None of the participants reported having absolute pitch.

Stimuli

The stimuli consisted of 24 tone sets composed from the tones of the Western C major mode <Do, Re, Mi, Fa, Sol, La, Ti> and 24 tone sets composed from the tones of the Indian *thaat* Bhairav <Do, Re-, Mi, Fa, Sol, La-, Ti> with C as the tonal center. Given the difficulty of the memory task, we chose to use a single tonal center for all of the stimuli. Previous research has shown that when tonal stimuli are transposed, tonal judgments are highly consistent across keys (Krumhansl & Kessler, 1982), suggesting that the judgments obtained in one key should generalize well to other keys.

In each tone set, either the 2nd or 6th degree was omitted (omission was counterbalanced). The sets contained seven tones, restricted to a one octave range. The 1st degree of the mode occurred twice in each tone set, and each other scale degree (with the exception of the one that was omitted) occurred once. The tone set always began on either the 1st, 4th, or 5th scale degree and always ended on the 1st scale degree. Each Western tone set had an Indian counterpart, so that the tone-to-tone transitions were identical across the two sets of stimuli, with the exception of the 1 tone in each set that established whether it was a Western or Indian tone set. For the 1 tone that differed across tone set counterparts, the scale degree was the same, varying only according to whether it was flat or not. Implied harmony was effectively controlled across cultures, given that the tone-to-tone transitions were identical for each Western tone set and Indian counterpart, with the exception of the 1 tone that varied across cultures.

Each tone was 250 ms in duration. Each tone set was followed by 1000 ms of silence, and then a 1000 ms test tone (either Re-, Re, La-, or La). The tone sets and test tones were played with a midi flute timbre (Finale Notepad, MakeMusic, Inc.). Each tone set was played four times throughout the experiment, so that it could be paired once with each of the four test tones, for a total of 192 trials.

Apparatus and Procedure

The experiment was run on an Apple Macintosh computer using the stimulus presentation program PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). Auditory stimuli were presented over Sony MDR-V600 stereo headphones at a comfortable volume that was held constant for all participants.

Each participant was seated alone in a sound-attenuated room for the experiment. Participants were asked to listen carefully to each tone set, and upon hearing the test tone, to decide as quickly and accurately as possible whether the test tone occurred in the preceding tone set. Responses were made by pressing a key labeled “yes” or a key labeled “no” on a computer keyboard. Reaction times were measured from the onset of the test tone.

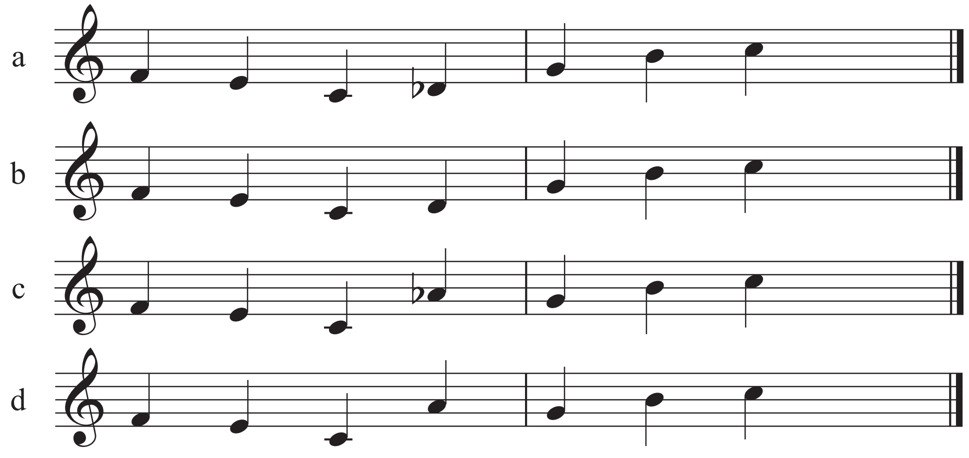

FIGURE 1. Notated examples of tone sets. Each Indian tone set (a and c) had a Western counterpart (b and d), differing only at the culture-distinguishing tone (the fourth tone in these examples).

TABLE 3. Responses to Test Tones Presented in Tone Sets.

	Test Tone Reaction Times	Test Tone Error Rates
Western Tone Set	1318.54 ms	30.95 %
Indian Tone Set	1295.68 ms	34.13 %
Mean	1307.11 ms	32.54 %

Results

We were particularly interested in the conditions in which the test tone did not occur in the tone set, as these are the conditions that should show the influence of cultural learning. However, we did examine the conditions in which the test tone *did* occur in the tone set, to establish that the inherent perceptual salience of the test tones did not differ across conditions. The reaction times for hits and error rates are shown in Table 3. There were no significant differences between the reaction times to the test tones occurring in the Western and Indian tone sets, $t(20) = -0.39$, $p > .50$, and there were no significant differences in the error rates (misses) across the Western and Indian tone sets, $t(20) = 0.99$, $p = .34$.

Tests of Cultural Learning Hypothesis

False alarm rates and reaction times for correct rejections were analyzed across four conditions: Western tone set/congruous test tones (La and Re), Western tone set/incongruous test tones (La- and Re-), Indian tone set/congruous test tones (La- and Re-), and Indian tone set/incongruous test tones (La and Re).

Mean false alarm rates are shown in Table 4. As predicted, the interaction between test tone congruity and stimulus tonality was significant, $F(1, 20) = 5.45$, $p = .03$, $\eta_p^2 = .21$. A higher percentage of false alarms occurred when test tones were culturally congruous with the Western melodies (La and Re, $M = 45.83\%$, $SD = 13.24$) than when congruous with the Indian melodies (La- and Re-, $M = 40.87\%$, $SD = 13.41$), $t(20) = -1.40$, $p = .18$ (two-tailed), but this difference failed to reach statistical significance. A significantly higher percentage of false alarms occurred when test tones were incongruous with the Indian melodies (La and Re, $M = 51.28\%$, $SD = 8.32$) than when incongruous with the Western melodies (La- and Re-, $M = 43.05\%$, $SD = 14.53$), $t(20) = 2.80$, $p = .01$ (two-tailed).

TABLE 4. Percentage of False Alarms to Test Tones Not Presented in Tone Sets.

	Congruous Test Tone	Incongruous Test Tone	*Mean*
Western Tone Set	45.83	43.05	44.44
Indian Tone Set	40.87	51.28	46.08
Mean	43.35	47.17	45.26

TABLE 5. Reaction Times (in ms) to Test Tones Not Presented in Tone Sets.

	Congruous Test Tone	Incongruous Test Tone	*Mean*
Western Tone Set	1588.65	1402.88	1495.76
Indian Tone Set	1386.15	1520.56	1453.36
Mean	1487.40	1461.72	1474.56

The main effect of test tone congruity was significant, $F(1, 20) = 6.27$, $p = .02$, $\eta_p^2 = .24$, with a higher percentage of errors occurring on culturally incongruous test tones ($M = 47.17\%$) than on congruous test tones ($M = 43.35\%$). The main effect of stimulus tonality (Western/Indian) was not significant, $F(1, 20) = 1.02$, $p = .32$, $\eta_p^2 = .05$.

Mean reaction times for correct rejections are shown in Table 5. The interaction between test tone congruity and tone set modality was significant, $F(1, 20) = 21.75$, $p < .001$, $\eta_p^2 = .52$. Test tones that were culturally congruous with the Indian tone sets (test tones La- and Re-, $M = 1386.15$ ms, $SD = 396.91$) were processed faster than test tones that were congruous with the Western tone sets (test tones La and Re, $M = 1588.65$ ms, $SD = 526.54$), $t(20) = -3.76$, $p = .001$ (two-tailed). Test tones that were culturally incongruous with the Western tone sets (La- and Re-, $M = 1402.88$ ms, $SD = 415.91$) were processed faster than test tones that were incongruous with the Indian tone sets (La and Re, $M = 1520.56$ ms, $SD = 466.98$), $t(20) = 3.26$, $p = .004$ (two-tailed).

The main effect of test tone congruity was not significant, $F(1, 20) = 0.80$, $p = .38$, $\eta_p^2 = .04$. The main effect of tone set modality (Western/Indian) was not significant, $F(1, 20) = 1.94$, $p = .18$, $\eta_p^2 = .09$.

The false alarm effects reported above suggest criterion shifts, which would be expected for a process of "filling in" or expectation: tones that are filled in or expected are activated by the context, thereby biasing the judgment in favor of saying that the test tone is present.

To see if there were also sensitivity effects generated by the context, *d'* scores were computed, as shown in Table 6. The main effect of test tone congruity was significant, $F(1, 20) = 5.44$, $p = .03$, $\eta_p^2 = .21$. The *d'* scores were lower for the culturally incongruous test tones ($M = 0.58$)

TABLE 6. d' Scores

	Congruous Test Tone	Incongruous Test Tone	*Mean*
Western Tone Set	0.67	0.63	0.65
Indian Tone Set	0.68	0.53	0.60
Mean	0.67	0.58	0.63

than for congruous test tones (M = 0.67), indicating that participants were less sensitive to whether the incongruous test tones were present in the tone sets than they were to the presence of the congruous test tones. The main effect of mode (Western/Indian) was not significant, $F(1, 20) = 1.11$, $p = .31$, $\eta_p^2 = .05$. The interaction between stimulus tonality and test tone congruity was not significant, $F(1, 20) = 0.26$, $p = .62$, $\eta_p^2 = .01$. Thus, although there was an overall sensitivity advantage for congruous tones, the interaction represents a criterion shift.

Tests of Harmonic Partials Hypothesis

Paired samples t-tests were conducted for each specific prediction discussed above and reiterated below, based on the harmonic partials hypothesis. All tests were two-tailed.

The harmonic partials hypothesis predicts that when the tone set contains Re-, the test tone La- should be reinforced relative to when the tone set contains Re. There were no significant differences between the reaction times for correct rejections, $t(20) = 1.30$, $p = .21$, or false alarms, $t(20) = 0.32$, $p > .50$, for these conditions.

The harmonic partials hypothesis predicts that when the tone set contains Re, the test tone La should be reinforced relative to when the tone set contains Re-. There were no significant differences between the reaction times for correct rejections, $t(20) = -0.34$, $p > .50$, or false alarms, $t(20) = -1.02$, $p = .32$, for these conditions.

The harmonic partials hypothesis predicts that the test tone Re should not be reinforced by the harmonic partials of La- or La, so no differences should emerge between those two conditions. There were no significant differences between the reaction times for correct rejections, $t(20) = -1.98$, $p = .06$, or false alarms, $t(20) = 0.71$, $p = .49$, for these conditions.

The harmonic partials hypothesis predicts that when the tone set contains La, the test tone Re- should be reinforced relative to when the tone set contains La-. There were no significant differences between the reaction times for correct rejections, $t(20) = -1.03$, $p = .32$, or false alarms, $t(20) = 1.50$, $p = .15$, for these conditions.

The harmonic partials hypothesis is thus disconfirmed, thereby supporting the cultural learning hypothesis.

Discussion

A culturally familiar musical context leads people to mistakenly believe that tones that are missing from that context were actually present. Tones that typically co-occur in a mode form a Gestalt representation, such that fragments of the mode can cause the missing tones to be cognitively filled in. The reaction time delay for missing tones that are congruous with the familiar mode shows that representations of these tones were primed or preactivated.

These results are consistent with the predictions of the neural network model: Western listeners were perceptually biased towards a Western modal framework, leading to the activation of the tones La and Re in their modal representations, and ultimately leading to slower rejections of these test tones and higher rates of false alarm. Although the Indian tone sets contained a tone (either Re- or La-) that would not co-occur with the rest of the melodic tone set <Do, Mi, Fa, Sol, and Ti> according to Western modal rules, the presence of the modally incongruous tone (that is, incongruous according to Western modal rules) did not cause the participants to adopt a different strategy for judging the test tones, such as shifting their response criterion to be more accepting of tones that did not fit their Western modal representations. Instead, the false alarm rates suggest that the participants were not biased by the influence of the La- or the Re- in the Indian melodies, and based their judgments on their Western modal knowledge. However, the hit rates and reaction times to hits (i.e., the trials in which the test tone did occur in the tone set) suggest that perceptual salience of Re- and La- did not differ significantly from the perceptual salience of Re and La. These findings suggest that the participants were able to recall the occurrence of the test tones with equal accuracy when the tones had actually occurred, but in the trials in which the test tone had not occurred, Western modal knowledge exerted a top-down influence on the judgments.

These findings are consistent with the cultural learning hypothesis. False alarm differences indicate that the effect of the context was to bias the perception of the Western listeners so that they believed that the missing congruous tones had been played. We did find a significant sensitivity main effect for test tone congruity, with participants being less able to discriminate the presence of modally incongruous test tones. We suspect that this effect is being driven by the test tones that were incongruous with the Indian tone sets (La and Re), as the mean *d'* for this condition ($M = 0.53$) was substantially lower than the mean *d'* for the test tones that were incongruous with the Western tone sets ($M = 0.63$). It is worth noting that the *d'* for the test tones that were congruous with the Western tone sets (La and Re) was substantially higher ($M = 0.67$) than when these test tones occurred in the context of the Indian tone sets ($M = 0.53$), indicating that the participants did not exhibit a generally lower sensitivity to the tones La and Re, but that this lower sensitivity occurs specifically in the context of the Indian pitch set. It is possible that the unfamiliar tone in the Indian tone set has cognitive implications for participants, making them less sensitive to the presence of test tones La and Re, which are likely to be of less perceptual salience than the unfamiliar tone in the Indian tone set, and also of less perceptual salience than test tones La- and Re-, as is seemingly indicated by the higher *d'* scores for La- and Re- in this condition.

Conclusions about cultural learning must always be made in the context of alternative hypotheses based on the physics or psychophysics of sound. The harmonic partials hypothesis predicted that the reaction times and false alarm rates would vary for specific test tones according to the tones comprising the stimulus tone sets. We found no evidence to support this hypothesis, as there were no significant differences in the reaction times and false alarm rates for each test tone across stimulus tone sets. The results of these analyses suggest that the influence of the harmonic partials on perception is not an adequate explanation for the results of the test tone task. Thus, it is likely that cultural learning accounts for the results.

Our stimuli were not fully representative of naturalistic musical stimuli, due to the necessity of controlling as many aspects of the stimuli as possible across modal cultures. Despite the diminished musicality of the controlled stimuli, the tone sets sufficiently activated the mental representations of modal structures in the participants, as evinced by our findings. It is possible that the use of more naturalistic musical stimuli may have produced a more robust effect.

The pattern of false alarms supports the theory that participants utilized a Western modal framework to perform the task. However, the overall rate of false alarms is somewhat high, and we can identify four factors that may explain the high false alarm rate. One reflects the diminished musicality of the stimuli; if the stimuli had been more naturalistic, modal expectations may have been stronger than those evinced in this study, leading to fewer false alarms in certain conditions (such as the Western tone set/incongruous test tone condition). Second, the task was designed to be extremely difficult, so as to ensure that participants would rely on tonal activation patterns, rather than on their veridical memory of each trial. All of the participants reported that it was very difficult to remember if the test tone had occurred in the tone set. The task difficulty may have led to high false alarm rates. Third, the participants grew accustomed to hearing Re- and La- in the context of Do, Mi, Fa, Sol, and Ti. This may have led to a higher level of expectations for Re- and La- to occur than one would expect in Western listeners, due to the probability of those tones occurring in the tone sets within the experiment. As previous research (Castellano, Bharucha, & Krumhansl, 1984; Krumhansl et al., 1999; Krumhansl et al., 2000) has shown, listeners are able to implicitly acquire the statistical probability of tones occurring in an unfamiliar modal system through passive exposure to that system. Perhaps our participants grew accustomed to hearing either Re- or La- occur in 50% of the tone sets, and their false alarms reflected this probabilistic knowledge. Fourth, the participants may have incorrectly assumed that their responses should be evenly distributed between "yes" (i.e., the test tone was present) and "no" (not present). The probability of encountering a trial in which the test tone was actually present in the tone set was 25%. The participants were not informed of this probability. The overall pattern of responses (57% "yes"/43% "no") is consistent with this interpretation. If the participants assumed that either answer was equally probable, it would certainly lead to high false alarm rates. However, this does not diminish the importance of the pattern of false alarms, which differed significantly across conditions, and was consistent with the cultural learning hypothesis.

Culture shapes our musical expectancies. The results of this experiment suggest that when listening to music from an unfamiliar modal system, we may impose our own cultural expectancies on that musical system. Thus, our experiences with an unfamiliar modality may be drastically different than the experiences of one who is familiar with the modality. For instance, when listening to music from an unfamiliar modality, one may have affective experiences that differ from those of a native listener. Expectancy violations are a source of affective responses to music (Meyer, 1956; Steinbeis, Koelsch, & Sloboda, 2006). Thus, tones in an unfamiliar modality that violate our cultural expectations may evoke emotional responses, and this source of emotional responses may be unique to the non-native listener. Of course, a native listener would have a unique set of affective responses evoked by culturally established expectancy violations that could not be experienced by a non-native listener.

Although a listener may gain short-term knowledge of the statistical regularities of an unfamiliar modality, the listener's long-term schematic knowledge of their own musical culture shapes the perception of music within and across cultures. Thus, the experience of music across cultures is likely to vary according to one's cultural background.

Author Note

Meagan E. Curtis is now at the Department of Psychology, Tufts University. The authors thank Kaivon Paroo for his assistance.

Correspondence concerning this article should be addressed to Meagan E. Curtis, Department of Psychology, Tufts University, 490 Boston Avenue, Medford, Massachusetts 02155. E-MAIL: meagan.curtis@tufts.edu

References

Bharucha, J. J. (1991). Pitch, harmony, and neural nets: A psychological approach. In P. M. Todd & D. G. Loy (Eds.), *Music and connectionism* (pp. 84-99). Cambridge, MA: MIT Press.

Bharucha, J. J. (1999). Neural nets, temporal composites and tonality. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 413-440). San Diego, CA: Academic Press.

Bharucha, J. J., & Mencl, W. E. (1996). Two issues in auditory cognition: Self-organization of categories and pitch-invariant pattern recognition. *Psychological Science, 7*, 142-149.

Bharucha, J. J., & Olney, K. L. (1989). Tonal cognition and artificial intelligence: Priming studies and connectionist modeling. *Contemporary Music Review, 4*, 341-356.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 403-410.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception and Psychophysics, 41*, 519-524.

Bharucha, J. J., & Todd, P. M. (1991). Modeling the perception of tonal structure with neural nets, *Computer Music Journal, 13,* 44-53.

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.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257-271.

Cohen, A. J., Thorpe, L. A., & Trehub, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology, 41*, 33–47.

Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception and Psychophysics, 41*, 609-620.

Dowling, W. J. (1999). The development of music cognition. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 603-625). San Diego, CA: Academic Press.

Hebb, D. (1949). *The organization of behavior.* New York: John Wiley and Sons.

Justus, T. C., & Bharucha, J. J. (2001). Modularity in music processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1000-1011.

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

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition, 10*, 243-251.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334-368.

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception, 17*, 151-195.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition, 76*, 13-58.

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

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

ROEDIGER, H. L., & MCDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

STEINBEIS, N., KOELSCH, S., & SLOBODA, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience, 18*, 1380-1393.

TEKMAN, H. G. & BHARUCHA, J. J. (1998). Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 252-260.

TILLMANN, B., BHARUCHA, J. J., & BIGAND, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*, 885-913.

TRAINOR, L. J., & TREHUB, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 394-402.

TRAINOR, L. J., & TREHUB, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics, 56*, 125-132.

TRAINOR, L. J., WU, L., & TSANG, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science, 7*, 289-296.

TREHUB, S. E. (2001). Musical predispositions in infancy. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (pp. 1-16). New York, NY: New York Academy of Sciences.

TREHUB, S. E., & TRAINOR, L. J. (1993). Listening strategies in infancy: The roots of language and musical development. In S. McAdams & E. Bigand (Eds.), *Cognitive aspects of human audition* (pp. 278-327). London: Oxford University Press.