Remembering Melodies from Another Culture: Turkish and American Listeners Demonstrate Implicit Knowledge of Musical Scales

Timothy Justus, Charles M. Yates, Nart Bedin Atalay, Nazike Mert, and Meagan E. Curtis

THE perception and cognition of music, especially those aspects concerning the processing of pitch and time within musical contexts, have been investigated using the tools of experimental psychology for well over a century. Among nineteenth-century German psychophysicists, there was some understanding that music cognition is historically and culturally contingent (Hui 2013). By the turn of the twentieth century, there was a clear dialogue between the experimental psychologists of the Berlin school and the nascent field of comparative musicology, which brought a global perspective to this work (Stumpf [1911] 2012). However, subsequent generations of European and American experimental psychologists tended to ground their study of music cognition firmly within the Western tonal-harmonic system, with a small number of notable exceptions (as described below). This overreliance on Western music and participants runs the risk of mistaking the peculiarities of one musical culture for human musical universals (cf. Nettl 1983). Fortunately, the twenty-first century has witnessed renewed interest in enculturation and musical diversity among experimental psychologists (Stevens 2012; Trehub, Becker, and Morley 2015).

Regarding rhythm and meter, for example, classic research concerning perceptual biases in metrical organization favoring 2:1 and 3:1 temporal ratios (Fraisse 1956) is no longer described as the result of a cognitive universal, but instead acknowledges the role of enculturation with Western music, in which said temporal ratios are pervasive. This shift is the result of programmatic developmental and cross-cultural research employing not only the simple, isochronous meters that characterize most Western music, but also the complex, non-isochronous meters that characterize other musical systems, such as Balkan (Rice 2004), Turkish (Bates 2011), and Hindustani music (Clayton 2000). Western infants, prior to a period of perceptual narrowing between six and twelve months of age, are sensitive to complex metrical distinctions that pose difficulty for Western adults (Hannon and Trehub 2005a, 2005b; Soley and Hannon 2010). Further, Balkan and Turkish adults retain this sensitivity to complex metrical structure into adulthood (Hannon, Soley, and Ullal 2012; Kalender, Trehub, and Schellenberg 2013). Turkish participants who are regular listeners of Turkish classical and folk music are especially sensitive to the complex meters found within these genres (Yates et al. 2017). Analogous results have been obtained using rhythmic production tasks with musicians from Germany, Bulgaria, and Mali (Polak et al. 2018).

Analytical Approaches to World Music, Vol. 7, No. 1. Published August 12, 2019.

Regarding musical pitch, researchers have employed probe-tone techniques to characterize listeners' implicit knowledge of tonality within Hindustani (Castellano, Bharucha, and Krumhansl 1984), Balinese (Kessler, Hansen, and Shepard 1984), Javanese (Perlman and Krumhansl 1996), Sami (Krumhansl et al. 2000), and Carnatic music (Raman and Dowling 2016). The corresponding infant work concerning pitch has tended to compare consonant versus dissonant intervals (Schellenberg and Trehub 1996) or Western versus artificial scales (Trehub, Schellenberg, and Kamenetsky 1999), rather than comparing Western pitch structures with those of another human musical system. The exceptions are studies of Western infants listening to Javanese scales (Lynch et al. 1990; Lynch and Eilers 1992). There is much systematic psychological work still to be done using this wealth of global musical scales and modes, many of which are well characterized by music theorists and ethnomusicologists.

Examination of global musical systems invites the realization that many supposedly universal and timeless aspects of musical pitch are Western European conventions, sometimes not more than a few hundred years old. These include the system of 12-tone equal temperament, the prominence of major and minor diatonic scales, and the importance of harmony based in chords constructed from these two common-practice modes. In Turkish, Arabic, Persian, and Hindustani music, among others, the number of possible pitch categories per octave differs from that of the Western system (Justus and Hutsler 2005). These systems each have a diverse set of scales and modes drawn from their respective tonal material, which give rise to dozens of distinct tonalities that do not employ triad-based harmony in the Western sense (Aydemir 2010; Muallem 2010; Jairazbhoy [1971] 1995).

The present work was inspired by a study by Curtis and Bharucha (2009), which, although relying on Western participants, developed musical stimuli based upon both the familiar major scale and unfamiliar *Bhairav thāt*. Bhairav is found within Hindustani classical music as one of many seven-tone *thāts*, which provide the pitch foundation for *rāgas* (Jairazbhoy [1971] 1995). The design of Curtis and Bharucha's (2009) experiment was analogous to the false-memory procedure of Deese (1959) and Roediger and McDermott (1995). Participants heard brief seven-tone melodies that were composed of all but one of the tones of the major scale or Bhairav thāt, and then were asked whether a subsequent probe tone had been presented in the preceding melody. The American participants, who as a group had only modest formal training in music, demonstrated implicit knowledge of Western tonality through an increased probability of incorrectly accepting the missing major scale tone as if it had been presented. Such perceptual inferences did not extend to the unfamiliar Bhairav scale. In a related study, Vuvan, Podolak, and Schmuckler (2014) found that whereas Western listeners falsely detected expected tones following major melodies, they did so to a lesser extent for minor melodies, and not at all for atonal melodies. In the experiments reported here, we apply this general approach within a cross-cultural study comparing Turkish and American listeners, who also varied systematically in their formal musical training.

The stimulus materials of the present study were based on the Western major scale, as well as the *rast makam*, a heptatonic scale thought to be of Persian origin that is also foundational to Turkish and Arabic classical music (Aydemir 2010; Muallem 2010; Marcus 2007). A first important difference between the two scales is their prevalence within their respective musical systems. The major scale is one of two modes (the other being minor) that have dominated Western tonal-harmonic music since the seventeenth century, having survived the other diatonic modes (Huron and Veltman 2006). In contrast, the rast makam is one of dozens of modes commonly encountered within Turkish and Arabic music, while nevertheless playing a central role within both systems. Whereas the tones of the Western major scale overlap with other makamlar in Turkish music (Aydemir 2010) and Arabic music (Muallem 2010), the rast makam has no equivalent in Western music.[1]

The rast makam differs from the major scale in that the third and seventh scale degrees are a quarter tone lower (1/24 of an octave) for rast compared to major (Muallem 2010). Therefore, whereas the major scale is diatonic, comprising seven steps of 1, 1, ½, 1, 1, 1, and ½ tone (one semitone being equal to 100 cents, or 1/12 of an octave), the rast makam is non-diatonic, comprising steps of 1, ¾, ¾, 1, 1, ¾, and ¾ tone, when in ascending form (Figure 1). Such microtonal scales require an underlying tonal material with more pitch categories than the 12 chromatic tones of Western music (see Burns 1999); however, theorists of Turkish and Arabic music have not agreed on what the underlying tonal material should be. Most theorists of Turkish music divide each whole tone into nine commas (Aydemir 2010; Signell 1977). More commonly within Arabic music theory, a 24-step equal-tempered division of the octave has been proposed (Marcus, 1993; Muallem, 2010; see Bozkurt et al. 2009 for a quantitative comparison). While acknowledging its limitations for Turkish music, we adopted the simpler equal-tempered approach in constructing our stimuli. For example, the tone E𝄳 was one quarter tone or 50 cents lower than E, and one quarter tone or 50 cents higher than E♭.

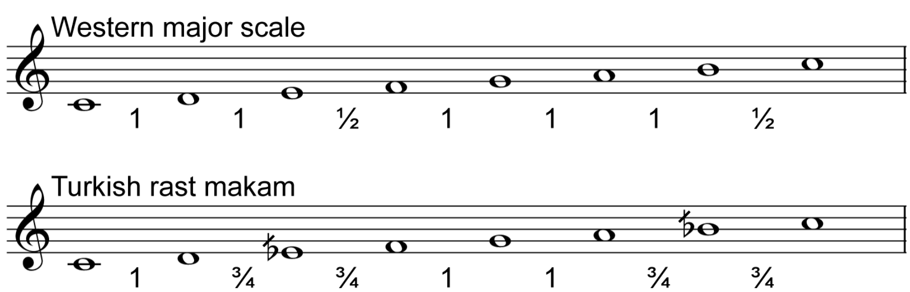

Figure 1. Major scale and rast makam. The accidentals on the third and seventh scale degree of the rast makam each indicate a half-flat, or a lowering by a quarter tone, relative to the major scale, while the numbers beneath the scales indicate the number of whole tones between consecutive scale degrees. As discussed in the text, we divided each octave into 24 equal-tempered tones (Muallem 2010), rather than dividing each whole tone into nine commas (Aydemir 2010).

1. We adopt the Turkish spelling of *makam* (plural *makamlar*) rather than the transcribed Arabic *maqām* (plural *maqāmāt*) in the present paper.

In addition to the underlying tonal material, another important difference between the major scale and the rast makam is that the latter has a characteristic absolute tonal center or tonic. Unlike in the Western system, in which major (or minor) tonality is based in relative pitch, and can be developed using any of the 12 chromatic pitches as the tonic, makamlar have characteristic tonal centers. Changing a makam's tonal center is likely to be accompanied by changes in scalar intervals, thereby changing the identity of the makam. According to Muallem (2010, 85), some theorists place great importance on the characteristic tonic, using this as a critical factor in classification. Some transpositions may occur in practice, but the characteristic tonal center is considered the "ideal position" and "essential form." For these reasons, the stimuli constructed for the present study, both major and rast, always used C as the tonic, rather than modulating among various tonal centers from trial to trial.[2]

Finally, it should be noted that the word *makam* refers to a modal system, and not merely a scale (Ayari and McAdams 2003; Akkoç, Sethares, and Karaosmanoğlu 2015). Each makam has not only a characteristic pitch structure (*perde*), but also characteristic melodic progressions and gestures (*seyir*). This is similar to how one might distinguish between *thāt* and *rāga* in Hindustani music theory (Jairazbhoy [1971] 1995). For example, a rast seyir would be expected to begin on middle C, first descending to G and ascending back to middle C before rising through the next octave, at least within Arabic music (Muallem 2010). Such development of the makam is as much a characteristic feature as are the component tones. In the present study, however, we wished to equate our major and rast "melodies" on all parameters other than the member tones (as described below). This was because prior work had demonstrated that listeners are sensitive to pitch distributions in stimulus sets (Oram and Cuddy 1995), including those derived from culturally unfamiliar music (e.g., Castellano, Bharucha, and Krumhansl 1984; Krumhansl et al. 1999; Stevens et al. 2013). Thus, any effects of an experimental manipulation of major and rast may be correctly attributed to listeners' prior knowledge of tonal material and scale membership, rather than a reflection of pitch distributions within the experiment itself. This initial study was not intended to address listeners' prior knowledge of the modal systems more broadly.

In the present experiment, we asked 60 participants, representing a range of familiarity with Western and Turkish musical scales, to listen to brief melodies, each composed of all but one of the member tones of the major scale or the rast makam, and then to decide whether a subsequent probe tone had been presented in the preceding melody. We first expected musicians to outperform nonmusicians in general, given a likely superiority in pitch discrimination and pitch short-term memory (cf. Halpern et al. 1996; Schellenberg and Moreno 2010; Weiss et al. 2015). We also expected enculturation effects when comparing the Turkish and American participants. Specifically, we expected the responses of the American participants to be influenced by implicit knowledge of the Western major scale, based on its

2. In Arabic music, the word *rast* refers to middle C. In Turkish music, the word *rast* refers to the written tone G. However, Turkish notation is transposed relative to Western and Arabic notation (Aydemir, 2010), such that G written in Turkish notation is closest to Western concert pitch D.

pervasiveness in the Western musical genres prevalent in the United States. In contrast, we expected the Turkish participants to be influenced by implicit knowledge of both the major scale and the rast makam, consistent with the use of both scale structures in traditional Turkish genres, as well as the prevalence of both Western and traditional Turkish music in Turkey (cf. Tekman and Hortaçsu 2002). Of the Turkish participants, we reckoned that musicians who were more familiar with Turkish classical and folk genres would be the most likely to demonstrate sensitivity to the rast makam (cf. Yates et al. 2017).

Methods

Participants

The 60 participants comprised the following five groups each with $n = 12$: American nonmusicians, American musicians, Turkish nonmusicians, Turkish musicians (studying both Western and Turkish genres), and Turkish classical and folk listeners (members of university clubs devoted to these genres). The American participants were recruited from the Claremont Colleges, and the Turkish participants were recruited from Middle East Technical University. All were young adults attending university, and thus the five groups were comparable in years of age and education.

Participants were considered nonmusicians if they had no more than two years of musical training (American: $M = 1$, $SD = 1$; Turkish: $M = 1$, $SD = 1$), whereas members of the musician groups had at least six years of musical training (American: $M = 12$, $SD = 5$; Turkish: $M = 9$, $SD = 6$). The Turkish classical and folk listeners were included as the fifth group given their years of systematic exposure to genres containing the rast makam ($M = 8$, $SD = 3$).

The instrument(s) on which the American musicians had been trained included piano ($n = 9$), voice (7), ukulele (4), guitar (3), violin (2), clarinet (2), saxophone (1), and percussion (1). All of the American participants reported listening to Western musical genres on a daily or weekly basis, and none reported any familiarity with Turkish or Arabic music. The primary language of all participants tested in the United States was English, with several reporting one or more additional languages, especially Spanish.

The instrument(s) on which the Turkish musician group had been trained included voice ($n = 12$), guitar (5), *bağlama/saz* (3), piano (3), violin (3), organ (2), and recorder (2). In contrast, the Turkish classical and folk group had studied the *oud* ($n = 4$), bağlama/saz (4), guitar (3), piano or keyboard (2), *cümbüş* (1), violin (1), clarinet (1), flute (1), and harmonica (1). Members of the Turkish nonmusician and musician groups reported listening to Western musical genres on a daily or weekly basis, but only two-thirds reported listening to Turkish musical genres with such regularity. In contrast, every member of the Turkish classical and folk group reported regular listening to traditional Turkish musical genres, with a subset also listening to Western genres. Specifically, this group indicated that they preferred Turkish classical ($n = 10$), Turkish folk (9), Turkish protest music (7), Western classical (6), Western pop (6), Rock (5), Arabesk (3), Turkish pop (2), Rap (1), Techno (1), and Underground (1). The primary language of all participants tested in Turkey was Turkish, and all reported one or more additional languages, especially English.

Stimuli

Melodies of seven tones each were written using Sibelius 7.5 software and synthesized with an oboe timbre, which we considered appropriate for both musical systems because it also sounds like the *zurna*, used in Turkish music. Initially, twelve melodies were composed, containing the tones C4 (262 Hz), D4 (294 Hz), F4 (349 Hz), G4 (392 Hz), A4 (440 Hz), and C5 (523 Hz), and a placeholder for a sixth critical context tone.[3] The melodies always began on C4, F4, G4, or C5, and always ended on C4 or C5. The placeholder for the critical context tone always occurred in the third, fourth, or fifth position, with equal probability. From the 12 initial melodies, 48 experimental melodies were created by adding one of four possible critical context tones: E𝄳 (320 Hz), E♮ (330 Hz), B𝄳 (480 Hz), or B♮ (494 Hz). Based on underlying tonal material of 24 equal-tempered divisions of the octave, the half-flats were one quarter tone or 50 cents lower than the corresponding natural, and one quarter tone or 50 cents higher than the corresponding flat. We refer to the 24 melodies containing either E♮ or B♮ as the "major contexts" and the 24 melodies containing either E𝄳 or B𝄳 as the "rast contexts" (Figure 2).[4]

Twelve probe tones were also created, again using the oboe-zurna timbre. Six of these correspond to the common scale tones that were presented in every context: C4, D4, F4, G4, A4, and C5. The remaining six probe tones were the most critical to our hypotheses: E♮, E𝄳, E♭, B♮, B𝄳, and B♭. We refer to E♮ and B♮, when following a major context, and E𝄳 and B𝄳 when following a rast context, as "same-mode probes." Similarly, we refer to E♮ and B♮, when following a *rast* context, and E𝄳 and B𝄳, when following a *major* context, as "other-mode probes." Finally, we refer to the E♭ and B♭ probes, regardless of which context preceded them, as "neither-mode probes," because these tunings of the third and seventh scale degrees do not occur in either the major scale or the rast makam, at least within ascending contexts.[5]

3. Hertz values are rounded and are for clarification only. The octave numberings here use American Standard Pitch Notation, in which C4 represents middle C, with a fundamental frequency of 262 Hz. In the text, we always refer to C4, D4, F4, G4, A4, and C5 by number. The remaining tones—E♮, E𝄳, E♭, B♮, B𝄳, and B♭—also always fell within octave four, but we omit the number 4 to draw attention to the tuning.
4. As acknowledged, the experimental melodies are limited by the fact that quarter-tonal rather than comma-based tuning was used for rast. Similarly, these "rast" melodies only attempt to capture the perde and not the seyir of the makam.
5. The neither-mode probes, E♭ (311 Hz) and B♭ (466 Hz) were included among the probe tones despite never being presented in the context melodies. We suspected that for listeners unfamiliar with the half-flat tones found in the rast makam, perceptual assimilation might occur in the direction of the flat probes. The inclusion of E♭ and B♭ also equated the number of common probe tones and critical probe tones (six each), and resulted in an overall relatedness proportion of seven out of twelve, or about 58 percent.

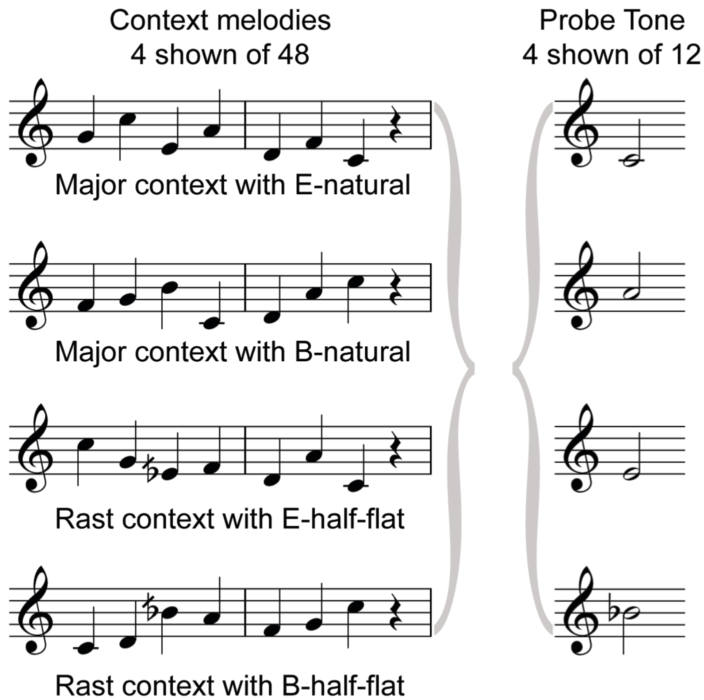

Figure 2. Stimulus melodies and probe tones. On the left, four examples of 48 context melodies are shown, with one of four critical context tones (E♮, B♮, E𝄳, B𝄳) as the third tone of the melody. (Critical tones could instead be presented as the fourth or fifth tone. All other aspects of the context melodies were counterbalanced between the major and rast conditions.) On the right, four examples of 12 probe tones are shown, with the first pair (C4 and A4) serving as examples of the six common probe tones (C4, D4, F4, G4, A4, C5), and the second pair (E♮ and B♭) serving as examples of the six critical probe tones (E♮, E𝄳, E♭, B♮, B𝄳, and B♭).

Further, if a probe tone corresponds to the same scale degree—third (E) or seventh (B)—that was presented in the corresponding context, regardless of whether it was flat, half-flat, or natural, we refer to it as a “same-degree” probe. If a probe tone corresponds to the scale degree that was not presented in the context, we refer to it as an “other-degree” probe. It should be noted of the six critical probe tones E♮, E𝄳, E♭, B♮, B𝄳, and B♭ that only those that are both “same mode” and “same scale degree” actually occurred in the context.

Over the course of the experiment, participants heard each of the 48 melodies exactly once. All were written with C as the tonic, rather than modulating among various pitch centers from trial to trial. Each probe tone was heard four times, two following a major context (one with E♮ and one with B♮) and two following a rast context (one with E𝄳 and one with B𝄳). The pairings between context melody and probe tone were counterbalanced using 12 stimulus sets (one for each participant in a group), to avoid any confounds between specific melodies and probe tones. Further, each of the 12 stimulus sets was associated with a different pseudorandom trial order, to avoid any confounds due to carry over from one trial to the next.

Procedure

Consent forms, questionnaires, and experimental instructions were provided in English for the American participants and in Turkish for the Turkish participants. The experiment was implemented in Presentation (Neurobehavioral Systems) and was carried out in a sound-attenuated environment with Sennheiser closed headphones. Each trial began with a fixation cross in the center of the screen for 2000 ms. Participants then heard a seven-tone melody lasting 4200 ms, followed by 5000 ms of silence (to avoid contributions of echoic memory), and then a probe tone lasting 1200 ms. They were asked to indicate whether they thought the probe tone had been presented in the melody, using a confidence scale of 1 (yes, quite confident that the tone was part of the melody) to 6 (no, quite confident the tone was not part of the melody).[6] Because participants were to input their responses over six response keys, speeded answers were not required. The fixation cross for the next trial appeared as soon as a response was input. After completing four practice trials, which could be repeated until participants understood the task and response scale, and the 48 trials of the main experiment, the participants were debriefed and received either a small payment or course credit. The Institutional Review Boards of Pitzer College and Middle East Technical University approved all procedures.

Results

Our analysis focused on the 24 ratings given by each participant to the six critical probe tones: E♮, E𝄳, E♭, B♮, B𝄳, and B♭, which could follow either a major context (one containing E♮ or B♮) or a rast context (one containing E𝄳 or B𝄳), each doing so twice over the course of the experiment. For the analysis, these tones are best thought of not as their absolute pitch, but rather in terms of the relationships between context and probe; they form a Context Mode (two levels: major, rast) by Probe Mode (three levels: same, other, neither) by Probe Scale Degree (two levels: same, other) design, as shown in Figure 3.

For example, a melody containing E♮ (or B♮) and followed by probe tone E♮ (or B♮) is an example of a major context with a same-mode, same-degree probe. Similarly, a melody containing E𝄳 (or B𝄳) and followed by probe tone E𝄳 (or B𝄳) is an example of a rast context with a same-mode, same-degree probe. The same-mode, same-degree probes are shown in the left-most position of Figure 3; these are the only probe tones in the analysis for which the objectively correct answer is "yes, the tone was part of the melody." None of the other probe tones were presented in the preceding melodies.

6. It should be noted that although we adopted continuous ratings, instead of the binary decision asked by Curtis and Bharucha (2009), the present task is also distinct from that of the classic probe-tone technique (summarized in Krumhansl 1990), in which participants were asked how well the probe tone completed the context, not whether it had actually occurred.

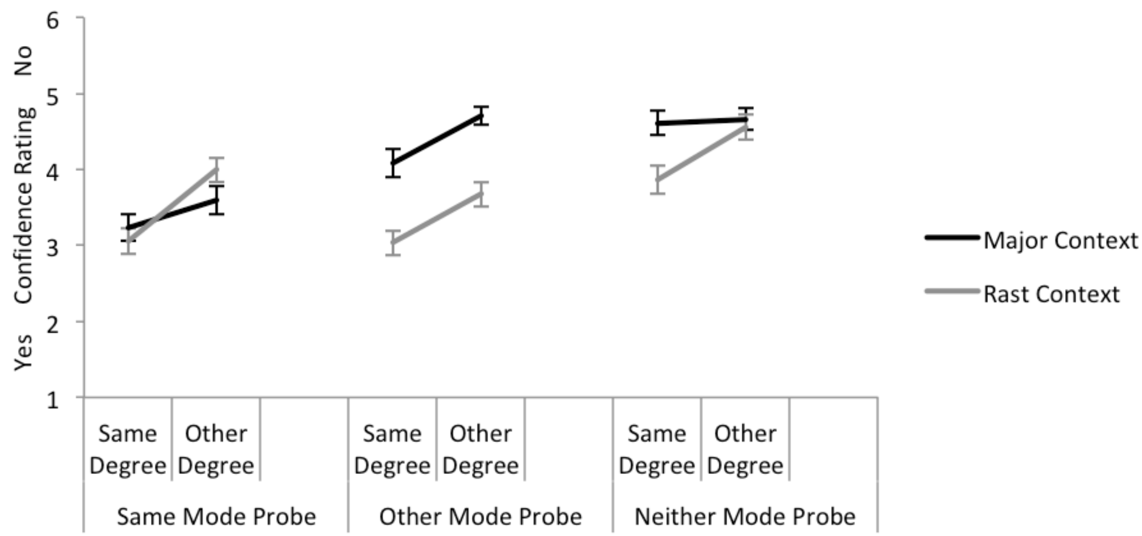

Figure 3. Experimental results overall. Mean ratings given to the critical probe tones (E♮, E𝄳, E♭, B♮, B𝄳, and B♭) as a function of Context Mode (major, rast), Probe Mode (same, other, neither), and Probe Scale Degree (same, other). Participants used a confidence scale of 1 (yes, quite confident that the tone was part of the melody) to 6 (no, quite confident the tone was not part of the melody). Among these tones, only the same mode, same degree probes (at the left) had actually occurred. Error bars indicate standard error of the mean.

An omnibus ANOVA including all five participant groups revealed significant main effects of all three within-groups factors. First, there was a main effect of Context Mode (major, rast), $F(1, 55) = 27.4$, $p < .001$, $\eta_p^2 = .33$, such that participants generally accepted (i.e., gave lower ratings to) probes following rast contexts ($M = 3.7$, $SD = 0.4$) over those following major contexts ($M = 4.1$, $SD = 0.3$). Second, there was a main effect of Probe Mode (same, other, neither), $F(2, 54) = 32.3$, $p < .001$, $\eta_p^2 = .54$, such that participants generally accepted same-mode probes ($M = 3.5$, $SD = 0.4$) over other-mode probes ($M = 3.9$, $SD = 0.4$), and other-mode probes over neither-mode probes ($M = 4.4$, $SD = 0.4$). Third, there was a main effect of Probe Scale Degree (same, other), $F(1, 55) = 36.6$, $p < .001$, $\eta_p^2 = .40$, such that participants generally accepted same-degree probes ($M = 3.6$, $SD = 0.3$) over other-degree probes ($M = 4.2$, $SD = 0.4$).

The analysis also revealed two significant interactions. First, Context Mode interacted with Probe Mode, $F(2, 54) = 12.2$, $p < .001$, $\eta_p^2 = .31$, such that the largest effects of major and rast context were observed for other-mode probes, followed by neither-mode probes, and with no difference overall between major and rast context for the same-mode probes (Figure 4). Second, Context Mode interacted with Probe Scale Degree, $F(1, 55) = 6.4$, $p = .01$, $\eta_p^2 = .10$, such that larger effects of major and rast context were observed for same-degree probes than for other-degree probes (Figure 5). The interaction between Probe Mode and Probe Scale Degree was not significant, $p = .27$, nor was the three-way interaction among these factors, $p = .16$.

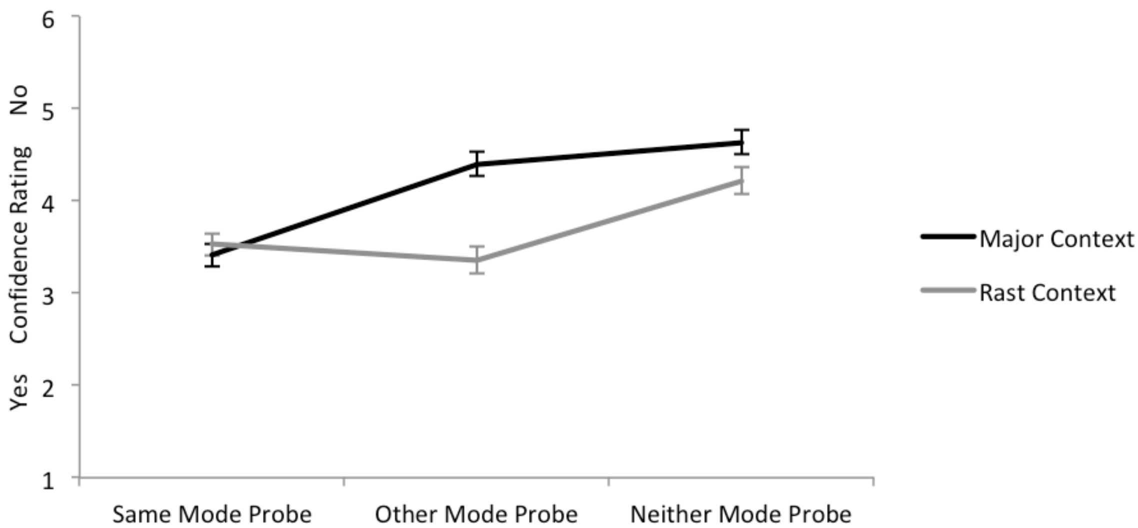

Figure 4. Context mode by probe mode interaction. Mean ratings given to the critical probe tones illustrating the significant interaction ($p < .001$) between Context Mode (major, rast) and Probe Mode (same, other, neither). Error bars indicate standard error of the mean.

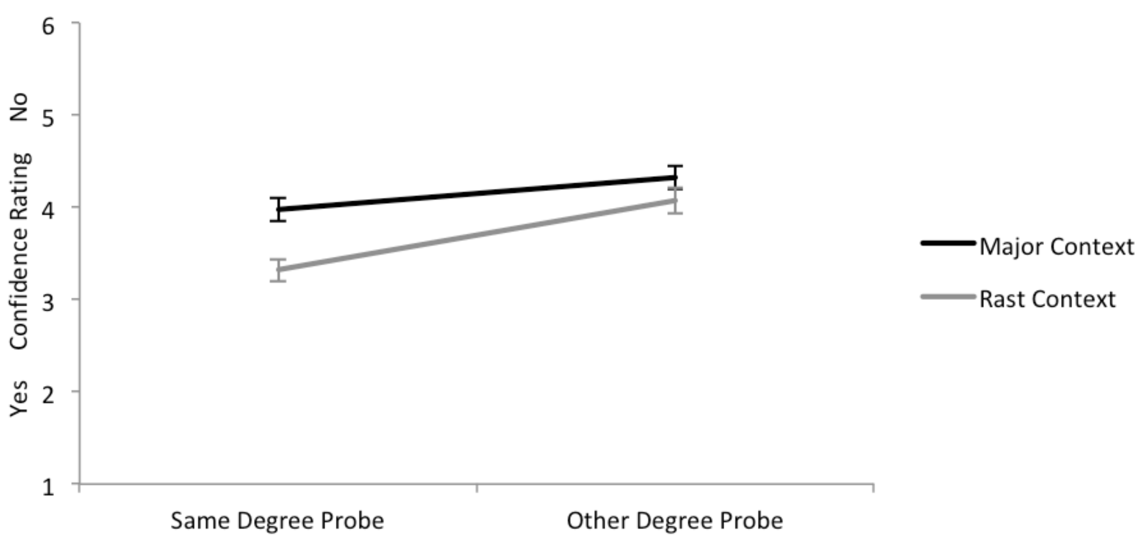

Figure 5. Context mode by probe scale degree interaction. Mean ratings given to the critical probe tones illustrating the significant interaction ($p = .01$) between Context Mode (major, rast) and Probe Scale Degree (same, other). Error bars indicate standard error of the mean.

No significant interactions involving the Group variable were observed in the omnibus ANOVA; participants generally demonstrated the above effects regardless of whether they were American or Turkish, and regardless of whether they were nonmusicians, musicians, or members of the Turkish classical and folk clubs. Follow-up analyses with the same stimulus factors, but for each group separately, generally revealed the same pattern.

To better understand the observed stimulus interactions, we performed an analysis restricted to the same-mode probes (shown at the left of Figure 3). This analysis revealed one main effect: Probe Scale Degree, $F(1, 55) = 14.2$, $p < .001$, $\eta_p^2 = .21$, such that participants generally accepted same-degree probes over other-degree probes, and an interaction trend between Context Mode and Probe Scale Degree, $F(1, 55) = 3.1$, $p = .08$, $\eta_p^2 = .05$, such that larger effects of scale degree were observed for rast contexts compared to major contexts. No significant interactions involving the Group variable were observed, and inspection of the data suggested a similar interaction pattern regardless of musical training or culture.

We next performed an analysis restricted to the other-mode probes (shown in the center of Figure 3). This analysis revealed both main effects: Context Mode, $F(1, 55) = 38.2$, $p < .001$, $\eta_p^2 = .41$, and Probe Scale Degree, $F(1, 55) = 23.7$, $p < .001$, $\eta_p^2 = .30$, but no interaction between the two. As before, no significant interactions involving the Group variable were observed.

Finally, we performed an analysis restricted to the neither-mode probes (shown at the right of Figure 3). This analysis revealed both main effects: Context Mode, $F(1, 55) = 9.9$, $p = .003$, $\eta_p^2 = .15$, and Probe Scale Degree, $F(1, 55) = 8.5$, $p = .005$, $\eta_p^2 = .13$, as well as an interaction, $F(1, 55) = 6.5$, $p = .01$, $\eta_p^2 = .11$. No significant interactions involving the Group variable were observed.

Discussion

The present experiment examined implicit knowledge of tonal material and musical scales as a function of cultural background and musical training. Specifically, American musicians, American nonmusicians, Turkish musicians, Turkish nonmusicians, and Turkish classical and folk club members listened to seven-tone melodies composed of the member tones of either the major scale or the rast makam, and were then asked to identify whether a probe tone had been presented in the melody. To our knowledge, this is the first cross-cultural study in which experimental stimuli were designed to isolate the differences in tonal material and scale between Western and Turkish makam music, or other modal systems (Arabic, Persian) employing the rast makam. In general, the results suggest that the participants' short-term memory for the stimuli was influenced by prior musical knowledge, especially of the major scale, but contrary to our predictions, no unequivocal effects of musical training or culture were observed.

First consider the same-mode probes shown at the left of Figure 3. In these trials, either major contexts (containing E♮ or B♮) or rast contexts (containing E𝄳 or B𝄳) were followed by a probe tone of the same mode. However, the probe tone could be either of the same scale degree as in the context (i.e., the presented tone) or the other scale degree (the missing tone). Correct "yes" responses (here, lower ratings) for same-mode, same-degree trials are likely driven by accurate short-term pitch memory, and reinforced by unconscious top-down inference based on schematic musical knowledge. However, for the same-mode, other-degree probes, such bottom-up and top-down processes are in opposition; these probe tones did not occur in the context, but schematic musical knowledge might suggest that they had, leading to an inaccurate unconscious inference, similar to a false memory (cf. Curtis and Bharucha 2009; Vuvan et al. 2014).

The differences between same-mode, same-degree probes and same-mode, other-degree probes differed as a function of whether the mode in question was the major scale or rast makam. Participants generally showed a smaller effect of probe degree for major compared to rast, suggesting that their top-down inferences were more consistent with knowledge of the major scale than with knowledge of the rast makam. In contrast, the larger effect for the rast contexts, and particularly the somewhat more confident "no" ratings given to rast context, same mode, other degree probes, suggests that participants' responses were based on veridical short-term memory, and less clouded by illusory unconscious inference.

Next consider the other-mode probes shown at the center of Figure 3. In these trials, either major contexts or rast contexts were followed by a probe tone of the *other* mode. As before, the probe tone could either be the same scale degree as in the context or the other scale degree, but they were always tones that had not been presented in the melody. Here, participants demonstrated the largest effect of context, generally being more likely to accept other-mode probes following rast contexts (that is, the E♮s and B♮s found in the major scale) over other-mode probes following major contexts (that is, the E𝄳s and B𝄳s found in the rast makam). This pattern is also suggestive of a greater facility with the major scale over the rast makam; recall that none of these tones were presented, and thus a higher rating (a more confident "no") is the more veridical response. As before, participants were more likely to accept same-degree over other-degree probes, but this effect did not interact with context.

Finally consider the neither-mode probes shown at the right of Figure 3. In contrast to the other-mode probes, which were not presented in the immediately preceding context, but were presented among the experimental melodies as a whole, the neither-mode probes (E♭ and B♭) are not members of either the major scale or the ascending rast makam, and were not part of any experimental melodies. Participants generally rejected (i.e., gave higher ratings to) these tones, with one exception: there was some tendency to accept the rast-context, same-degree probes relative to the other three types of neither-mode probe. This result is suggestive of perceptual assimilation of the half-flat tones presented in rast contexts to the nearby flat equivalents (cf. Krumhansl and Shepard 1979), as might be expected of someone more familiar with the Western tonal material of 12 equal-tempered chroma per octave than with the microtonal tuning characteristic of Turkish music.

While the present work revealed clear evidence of long-term musical knowledge influencing short-term pitch memory, it did not demonstrate any significant differences among the American nonmusicians, American musicians, Turkish nonmusicians, Turkish musicians, and Turkish classical and folk listeners, suggesting that all five groups based their responses primarily on knowledge of the Western major scale. This contrasts with our previous study concerning temporal organization in music, in which musicians generally performed better on the tasks, while only the members of the Turkish classical and folk clubs were sensitive to the complex meters characteristic of these genres (Yates et al. 2017).

There are at least two reasons why effects of musical training and culture were more readily observed in the prior study of meter compared to the present study of scale. One concerns the nature of the tasks. In Yates et al. (2017; also see Kalender et al. 2013), listeners rated the extent to which the rhythm of a repeating melody changed in four conditions: no change, meter-preserving change, meter-violating change, and obvious change. This was done for both simple and complex meters, with the difference in rating between the meter-preserving and meter-violating conditions serving as the measure of metrical sensitivity. While that study also employed a continuous behavioral response, it is rather different to the present task in which short-term memory and prior musical knowledge were pitted against each other, for example, when prior knowledge of mode makes it difficult to reject a probe tone that did not actually occur.

A second reason for the difference between studies concerns the asymmetries in knowledge of the Western and Turkish musical systems, both in general and as reflected in our participant groups. Most Western listeners, including our American participants, are not familiar with Turkish classical and folk genres or the diverse modes and complex meters found within them. In contrast, most Turkish listeners, including our Turkish participants, are in a sense “bimusical,” being familiar with both musical traditions (cf. Wong, Roy, and Margulis 2009; Wong et al. 2011). Considering the earlier study of meter, a participant experienced with simple as well as complex meters could demonstrate knowledge of both, given the experimental design. However, in the present study, robust effects of the major scale would obscure any influences of the rast makam.

It is possible that Turkish musicians, especially Turkish classical and folk musicians, would demonstrate knowledge of the rast makam in other experimental designs. Considering the bimusicality of this group, future studies might recruit only musicians with extensive listening and performance experience within the Turkish system, especially on instruments such as the bağlama (or saz), which permits the microtonal tuning required for the rast makam. For similar reasons, such studies might exclude any musician who had studied the piano, as such training specifically reinforces Western tonal material. In general, empirical demonstrations of implicit knowledge of Turkish makamlar may require the recruitment of makam masters and the use of richer stimulus materials that capture both the perde and seyir of these modes (as in Akkoç et al. 2015; Ayari and McAdams 2003).

We are aware of one prior study that used a probe-tone technique with Turkish makamlar and participant groups of varying familiarity with this system. Karaelma (2008) compared four groups of undergraduate Turkish participants: 17 students of Turkish makam music (studying Turkish music only), 20 students of music education (studying both Turkish and Western music), 20 students in a military brass band (studying Western music only), and 20 students of visual arts education (not studying music). Eight makamlar were used in the study, including rast. These contexts and the subsequent probe tones were performed by a *ney* or flute. The results generally suggested that probe tone ratings were correlated among the groups, and while inferential statistics were not reported, inspection of the probe tone ratings following rast contexts suggests the clearest differentiation of the tones among the Turkish makam students.

The present work also complements prior studies illustrating that American and Turkish participants are better able to encode new samples of music from their own culture as opposed to another, a finding that has been observed with both adults (Demorest et al. 2008) and fifth-grade children (Morrison, Demorest, and Stambaugh 2008). In these initial studies, real but unfamiliar musical excerpts were taken from Western, Turkish, and Chinese genres. Thus, the experimental materials differed over a variety of formal dimensions including aspects of pitch (tonal material, scale, mode), temporal organization (tempo, rhythm, meter), and instrumental timbre, all of which likely influenced participants' ability to best discriminate old and new melodies within culturally familiar genres. In a subsequent study, the within-culture advantage was preserved for American and Canadian participants even when such cues were reduced (Demorest et al. 2016).

Future studies comparing the Turkish and Western musical systems may wish to consider listeners' knowledge not only of the formal elements that characterize the modes (such as perde and seyir) but also their extra-musical meanings. To an experienced listener, each makam likely evokes distinct semantic and emotional associations, not unlike the distinct moods and connotations of major versus minor tonality among Western listeners (Parncutt 2014; Justus, Gabriel, and Pfaff 2018). It is not uncommon for theorists to use cross-modal metaphors in describing the distinct flavor (Aydemir 2010) or color (Muallem 2010) of each makam, just as one might speak of the *rāsa* associated with each rāga in Hindustani music. Indeed, Turkish classical theorists use the word *çeşni* (flavor):

> I describe the *çeşnis* to my students in this way: Think about spices. Every spice has its own particular taste, aroma, and color. What if we were to read all about a spice, learn its composition and examine its color but never tasted it? . . . Only after learning its flavor would we be able to decide ourselves where to use it, and in what quantity. This is precisely the case with the *çeşnis* of Turkish music . . . Only when we hear these pitches from a reliable source will we have "tasted" the *çeşnis*, and only by listening to and imitating them time and time again may we internalize them. After this process the musician will be able to use the *çeşnis* wherever and however he wants, and identify the makam of a piece that he hears. (Aydemir 2010, 8)

Such connotations in the minds of musicians and experienced listeners suggest not only implicit knowledge of the modal forms themselves, but also learned associations between these gestalten and their extra-musical, often emotional, meanings. It may be through implicit learning of these form-meaning associations that expert listeners come to be able to recognize and name each makam. In turn, being able to name the makam underlying a presented melody may influence judgments concerning its component tones in experiments like ours. In Karaelma's (2008) study, for example, participants who were able to identify the rast makam by name provided more clearly differentiated responses compared to those who could not.

Prior cross-cultural work on emotion in music has primarily focused on whether certain emotional connotations are available to the culturally unfamiliar listener on the basis of non-arbitrary psychoacoustic cues such as tempo, loudness, or dissonance (Balkwill and Thompson 1999; Balkwill, Thompson, and Matsunaga 2004; Fritz et al. 2009). However, in the context of the present discussion, it might be equally interesting to focus instead on the culture-specific, learned associations of expert listeners. Although, to our knowledge, the emotional connotations of the Turkish modes have not yet been studied using psychological methods, the predicted associations would be relevant to how listeners select and use music in their daily lives (cf. Boer et al. 2012), including how culture-specific music is employed in building group affiliation and identity (cf. Tekman and Hortaçsu 2002).

A final point of speculation concerns the relation between language and (musical) thought. An intriguing difference between the primary languages of our participant groups is that Turkish, among other languages including Farsi, uses a thick (*kalın*) to thin (*ince*) spatial metaphor in describing pitch, in addition to the low to high metaphor used in Germanic and Romance languages (Shayan, Ozturk, and Sicoli 2011). While both mappings appear to be equally "natural" and learnable (Dolscheid et al. 2014; Shayan et al. 2014), differences in the spatial metaphors of participants' native languages (e.g., Farsi vs. Dutch) have been shown to influence performance in psychophysical tasks involving pitch and space in which no language is required (Dolscheid et al. 2013). In the context of the present experiment, the native languages of the participants may have resulted in qualitatively different spatial imagery as they tracked the pitch height (or width) of the melodies and probe tones. Because such spatial imagery can also be considered a form of extra-musical meaning, future cross-cultural work may wish to examine how language, culture, and thought interact in determining musical understanding.

ACKNOWLEDGMENTS

This research was supported by a faculty development grant from Pitzer College's Institute for Global Local Action and Study. The authors would also like to thank Mine Misirlisoy, Samantha Bromley-Coolidge, and Gary So for their assistance. Preliminary findings were presented at the annual meeting of the Psychonomic Society (2014) and the Auditory Perception, Cognition, and Action Meeting (2015).

REFERENCES

Akkoç, Can, William A. Sethares, and M. Kemal Karaosmanoğlu. 2015. "Experiments on the Relationship between Perde and Seyir in Turkish Makam Music." *Music Perception* 32(4): 322–43. https://doi.org/10.1525/mp.2015.32.4.322.

Ayari, Mondher, and Stephen McAdams. 2003. "Aural Analysis of Arabic Improvised Instrumental Music (Taqsīm)." *Music Perception* 21(2): 159–216. https://doi.org/10.1525/mp.2003.21.2.159.

Aydemir, Murat. 2010. *Turkish Music Makam Guide.* Translated by Erman Dirikcan. Istanbul: Pan Yayıncılık.

Balkwill, Laura-Lee, and William Forde Thompson. 1999. "A Cross-Cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues." *Music Perception* 17(1): 43–64. https://doi.org/10.2307/40285811.

Balkwill, Laura-Lee, William Forde Thompson, and Rie Matsunaga. 2004. "Recognition of Emotion in Japanese, Western, and Hindustani Music by Japanese Listeners." *Japanese Psychological Research* 46(4): 337–47. https://doi.org/10.1111/j.1468-5584.2004.00265.x.

Bates, Eliot. 2011. *Music in Turkey: Experiencing Music, Expressing Culture*. Oxford: Oxford University Press.

Boer, Diana, Ronald Fischer, Hasan Gürkan Tekman, Amina Abubakar, Jane Njenga, and Markus Zenger. 2012. "Young People's Topography of Musical Functions: Personal, Social and Cultural Experiences with Music Across Genders and Six Societies." *International Journal of Psychology* 47(5): 355–69. https://doi.org/10.1080/00207594.2012.656128.

Bozkurt, Barış, Ozan Yarman, M. Kemal Karaosmanoğlu, and Can Akkoç. 2009. "Weighing Diverse Theoretical Models on Turkish *Maqam* Music against Pitch Measurements: A Comparison of Peaks Automatically Derived from Frequency Histograms with Proposed Scale Tones." *Journal of New Music Research* 38(1): 45–70. https://doi.org/10.1080/09298210903147673.

Burns, Edward M. 1999. "Intervals, Scales, and Tuning." In *The Psychology of Music*, edited by Diana Deutsch, 2nd ed., 215–64. San Diego: Academic Press.

Castellano, Mary A., Jamshed J. Bharucha, and Carol L. Krumhansl. 1984. "Tonal Hierarchies in the Music of North India." *Journal of Experimental Psychology: General* 113(3): 394–412. https://doi.org/10.1037/0096-3445.113.3.394.

Clayton, Martin. 2000. *Time in Indian Music: Rhythm, Metre, and Form in North Indian Rag Performance*. Oxford: Oxford University Press.

Curtis, Meagan E., and Jamshed J. Bharucha. 2009. "Memory and Musical Expectation for Tones in Cultural Context." *Music Perception* 26(4): 365–75. https://doi.org/10.1525/mp.2009.26.4.365.

Deese, James. 1959. "On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate Recall." *Journal of Experimental Psychology* 58(1): 17–22. https://doi.org/10.1037/h0046671.

Demorest, Steven M., Steven J. Morrison, Münir N. Beken, and Denise Jungbluth. 2008. "Lost in Translation: An Enculturation Effect in Music Memory Performance." *Music Perception* 25(3): 213–23. https://doi.org/10.1525/mp.2008.25.3.213.

Demorest, Steven M., Steven J. Morrison, Vu Q. Nguyen, and Erin N. Bodnar. 2016. "The Influence of Contextual Cues on Cultural Bias in Music Memory." *Music Perception* 33(5): 590–600. https://doi.org/10.1525/mp.2016.33.5.590.

Dolscheid, Sarah, Sabine Hunnius, Daniel Casasanto, and Asifa Majid. 2014. "Prelinguistic Infants Are Sensitive to Space-Pitch Associations Found Across Cultures." *Psychological Science* 25(6): 1256–61. https://doi.org/10.1177/0956797614528521.

Dolscheid, Sarah, Shakila Shayan, Asifa Majid, and Daniel Casasanto. 2013. "The Thickness of Musical Pitch: Psychophysical Evidence for Linguistic Relativity." *Psychological Science* 24(5): 613–21. https://doi.org/10.1177/0956797612457374.

Fraisse, Paul. 1956. *Les structures rythmiques: Etude psychologique*. Louvain: Publications universitaires de Louvain.

Fritz, Thomas, Sebastian Jentschke, Nathalie Gosselin, Daniela Sammler, Isabelle Peretz, Robert Turner, Angela D. Friederici, and Stefan Koelsch. 2009. "Universal Recognition of Three Basic Emotions in Music." *Current Biology* 19(7): 573–76. https://doi.org/10.1016/j.cub.2009.02.058.

Halpern, Andrea R., SeYeul Kwak, James C. Bartlett, and W. Jay Dowling. 1996. "Effects of Aging and Musical Experience on the Representation of Tonal Hierarchies." *Psychology and Aging* 11(2): 235–46. https://doi.org/10.1037/0882-7974.11.2.235.

Hannon, Erin E., and Sandra E. Trehub. 2005a. "Metrical Categories in Infancy and Adulthood." *Psychological Science* 16(1): 48–55. https://doi.org/10.1111/j.0956-7976.2005.00779.x.

———. 2005b. "Tuning In to Musical Rhythms: Infants Learn More Readily Than Adults." *Proceedings of the National Academy of Sciences* 102(35): 12639–43. https://doi.org/10.1073/pnas.0504254102.

Hannon, Erin E., Gaye Soley, and Sangeeta Ullal. 2012. "Familiarity Overrides Complexity in Rhythm Perception: A Cross-Cultural Comparison of American and Turkish Listeners." *Journal of Experimental Psychology: Human Perception and Performance* 38(3): 543–48. https://doi.org/10.1037/a0027225.

Hui, Alexandra. 2013. *The Psychophysical Ear: Musical Experiments, Experimental Sounds, 1840–1910*. Cambridge, MA: MIT Press.

Huron, David, and Joshua Veltman. 2006. "A Cognitive Approach to Medieval Mode: Evidence for an Historical Antecedent to the Major/Minor System." *Empirical Musicology Review* 1(1): 33–55. https://doi.org/10.18061/1811/24072.

Jairazbhoy, Nazir Ali. (1971) 1995. *The Rāgs of North Indian Music: Their Structure and Evolution*. First revised Indian edition. Mumbai: Popular Prakashan.

Justus, Timothy, Laura Gabriel, and Adela Pfaff. 2018. "Form and Meaning in Music: Revisiting the Affective Character of the Major and Minor Modes." *Auditory Perception and Cognition* 1(3–4): 229–47. https://doi.org/10.1080/25742442.2019.1597578.

Justus, Timothy, and Jeffrey J. Hutsler. 2005. "Fundamental Issues in the Evolutionary Psychology of Music: Assessing Innateness and Domain-Specificity." *Music Perception* 23(1): 1–27. https://doi.org/10.1525/mp.2005.23.1.1.

Kalender, Beste, Sandra E. Trehub, and E. Glenn Schellenberg. 2013. "Cross-Cultural Differences in Meter Perception." *Psychological Research* 77(2): 196–203. https://doi.org/10.1007/s00426-012-0427-y.

Karaelma, Barış. 2008. "Makamsal işitmede algısal perde hiyerarşisinin dinleyicilerin eğitim düzeyleriyle ilişkisi" [The relation between the perceptual tonal hierarchy and listeners' level of education in makam perception]. PhD diss., Gazi University.

Kessler, Edward J., Christa Hansen, and Roger N. Shepard. 1984. "Tonal Schemata in the Perception of Music in Bali and in the West." *Music Perception* 2(2): 131–65. https://doi.org/10.2307/40285289.

Krumhansl, Carol L. 1990. *Cognitive Foundations of Musical Pitch*. Oxford: Oxford University Press.

Krumhansl, Carol L., Jukka Louhivuori, Petri Toiviainen, Topi Järvinen, and Tuomas Eerola. 1999. "Melodic Expectation in Finnish Spiritual Folk Hymns: Convergence of Statistical, Behavioral, and Computational Approaches." *Music Perception* 17(2): 151–95. https://doi.org/10.2307/40285890.

Krumhansl, Carol L., and Roger N. Shepard. 1979. "Quantification of the Hierarchy of Tonal Functions within a Diatonic Context." *Journal of Experimental Psychology: Human Perception and Performance* 5(4): 579–94. https://doi.org/10.1037/0096-1523.5.4.579.

Krumhansl, Carol L., Pekka Toivanen, Tuomas Eerola, Petri Toiviainen, Topi Järvinen, and Jukka Louhivuori. 2000. "Cross-Cultural Music Cognition: Cognitive Methodology Applied to North Sami Yoiks." *Cognition* 76(1): 13–58. https://doi.org/10.1016/S0010-0277(00)00068-8.

Lynch, Michael P., and Rebecca E. Eilers. 1992. "A Study of Perceptual Development for Musical Tuning." *Perception and Psychophysics* 52(6): 599–608. https://doi.org/10.3758/BF03211696.

Lynch, Michael P., Rebecca E. Eilers, D. Kimbrough Oller, and Richard C. Urbano. 1990. "Innateness, Experience, and Music Perception." *Psychological Science* 1(4): 272–76. https://doi.org/10.1111/j.1467-9280.1990.tb00213.x.

Marcus, Scott. 1993. "The Interface between Theory and Practice: Intonation in Arab Music." *Asian Music* 24(2): 39–58. https://doi.org/10.2307/834466.

———. 2007. *Music in Egypt: Experiencing Music, Expressing Culture*. Oxford: Oxford University Press.

Morrison, Steven J., Steven M. Demorest, and Laura A. Stambaugh. 2008. "Enculturation Effects in Music Cognition: The Role of Age and Music Complexity." *Journal of Research in Music Education* 56(2): 118–29. https://doi.org/10.1177/0022429408322854.

Muallem, David. 2010. *The Maqām Book: A Doorway to Arab Scales and Modes.* Translated by Yoram Arnon. Kfar Sava, Israel: Or-Tav Music Publications.

Nettl, Bruno. 1983. *The Study of Ethnomusicology: Twenty-Nine Issues and Concepts*. Champaign: University of Illinois Press.

Oram, Nicholas, and Lola L. Cuddy. 1995. "Responsiveness of Western Adults to Pitch-Distributional Information in Melodic Sequences." *Psychological Research* 57(2): 103–18.

Parncutt, Richard. 2014. "The Emotional Connotations of Major Versus Minor Tonality: One or More Origins?" *Musicae Scientiae* 18(3): 324–353. https://doi.org/10.1177/1029864914542842.

Perlman, Marc, and Carol L. Krumhansl. 1996. "An Experimental Study of Internal Interval Standards in Javanese and Western Musicians." *Music Perception* 14(2): 95–116. https://doi.org/10.2307/40285714.

Polak, Rainer, Nori Jacoby, Timo Fischinger, Daniel Goldberg, Andre Holzapfel, and Justin London. 2018. "Rhythmic Prototypes Across Cultures: A Comparative Study of Tapping Synchronization." *Music Perception* 36(1): 1–23. https://doi.org/10.1525/mp.2018.36.1.1.

Raman, Rachna, and W. Jay Dowling. 2016. "Real-Time Probing of Modulations in South Indian Classical (Carnatic) Music by Indian and Western Musicians." *Music Perception* 33(3): 367–93. https://doi.org/10.1525/mp.2016.33.3.367.

Rice, Timothy. 2004. *Music in Bulgaria: Experiencing Music, Expressing Culture*. Oxford: Oxford University Press.

Roediger, Henry L., III, and Kathleen B. McDermott. 1995. "Creating False Memories: Remembering Words Not Presented in Lists." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21(4): 803–14. https://doi.org/10.1037/0278-7393.21.4.803.

Schellenberg, E. Glenn, and Sylvain Moreno. 2010. "Music Lessons, Pitch Processing, and *g*." *Psychology of Music* 38(s): 209–21. https://doi.org/10.1177/0305735609339473.

Schellenberg, E. Glenn, and Sandra E. Trehub. 1996. "Natural Musical Intervals: Evidence from Infant Listeners." *Psychological Science* 7(5): 272–77. https://doi.org/10.1111/j.1467-9280.1996.tb00373.x.

Shayan, Shakila, Ozge Ozturk, Melissa Bowerman, and Asifa Majid. 2014. "Spatial Metaphor in Language Can Promote the Development of Cross-Modal Mappings in Children." *Developmental Science* 17(4): 636–43. https://doi.org/10.1111/desc.12157.

Shayan, Shakila, Ozge Ozturk, and Mark A. Sicoli. 2011. "The Thickness of Pitch: Crossmodal Metaphors in Farsi, Turkish, and Zapotec." *Senses and Society* 6(1): 96–106. https://doi.org/10.2752/174589311X12893982233911.

Signell, Karl L. 1977. *Makam: Modal Practice in Turkish Art Music*. Seattle, WA: Asian Music Publications.

Soley, Gaye, and Erin E. Hannon. 2010. "Infants Prefer the Musical Meter of Their Own Culture: A Cross-Cultural Comparison." *Developmental Psychology* 46(1): 286–92. https://doi.org/10.1037/a0017555.

Stevens, Catherine J. 2012. "Music Perception and Cognition: A Review of Recent Cross-Cultural Research." *Topics in Cognitive Science* 4(4): 653–67. https://doi.org/10.1111/j.1756-8765.2012.01215.x.

Stevens, Catherine J., Julien Tardieu, Peter Dunbar-Hall, Catherine T. Best, and Barbara Tillmann. 2013. Expectations in Culturally Unfamiliar Music: Influences of Proximal and Distal Cues and Timbral Characteristics." *Frontiers in Psychology* 4: 789. https://doi.org/10.3389/fpsyg.2013.00789.

Stumpf, Carl. (1911) 2012. *The Origins of Music*. Translated by David Trippett. Oxford: Oxford University Press.

Tekman, Hasan Gürkan, and Nuran Hortaçsu. 2002. "Music and Social Identity: Stylistic Identification as a Response to Musical Style." *International Journal of Psychology* 37(5): 277–85. https://doi.org/10.1080/00207590244000043.

Tillmann, Barbara, Jamshed J. Bharucha, and Emmanuel Bigand. 2000. "Implicit Learning of Tonality: A Self-Organizing Approach." *Psychological Review* 107(4): 885–913. https://doi.org/10.1037/0033-295X.107.4.885.

Trehub, Sandra E., Judith Becker, and Iain Morley. 2015. "Cross-Cultural Perspectives on Music and Musicality." *Philosophical Transactions of the Royal Society B* 370(1664): 20140096. https://doi.org/10.1098/rstb.2014.0096.

Trehub, Sandra E., E. Glenn Schellenberg, and Stuart B. Kamenetsky. 1999. "Infants' and Adults' Perception of Scale Structure." *Journal of Experimental Psychology: Human Perception and Performance* 25(4): 965–75. https://doi.org/10.1037/0096-1523.25.4.965.

Vuvan, Dominique T., Olivia M. Podolak, and Mark A. Schmuckler. 2014. "Memory for Musical Tones: The Impact of Tonality and the Creation of False Memories." *Frontiers in Psychology* 5: 582. https://doi.org/10.3389/fpsyg.2014.00582.

Weiss, Michael W., Patrícia Vanzella, E. Glenn Schellenberg, and Sandra E. Trehub. 2015. "Pianists Exhibit Enhanced Memory for Vocal Melodies but not Piano Melodies." *Quarterly Journal of Experimental Psychology* 68(5): 866–77. https://doi.org/10.1080/17470218.2015.1020818.

Wong, Patrick C. M., Alice H. D. Chan, Anil Roy, and Elizabeth H. Margulis. 2011. "The Bimusical Brain Is Not Two Monomusical Brains in One: Evidence from Musical Affective Processing." *Journal of Cognitive Neuroscience* 23(12): 4082–93. https://doi.org/10.1162/jocn_a_00105.

Wong, Patrick C. M., Anil K. Roy, and Elizabeth Hellmuth Margulis. 2009. "Bimusicalism: The Implicit Dual Enculturation of Cognitive and Affective Systems." *Music Perception* 27(2): 81–88. https://doi.org/10.1525/mp.2009.27.2.81.

Yates, Charles M., Timothy Justus, Nart Bedin Atalay, Nazike Mert, and Sandra E. Trehub. 2017. "Effects of Musical Training and Culture on Meter Perception." *Psychology of Music* 45(2): 231–45. https://doi.org/10.1177/0305735616657407.