Cognition 235 (2023) 105405

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Worlds apart? Testing the cultural distance hypothesis in music perception of Chinese and Western listeners

Mathias Klarlund [a,b,c,*], Elvira Brattico [c,d], Marcus Pearce [c,i], Yiyang Wu [a,e], Peter Vuust [c], Morten Overgaard [h], Yi Du [a,e,f,g,**]

[a] *CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China*
[b] *Sino-Danish College, University of Chinese Academy of Sciences, Beijing, China*
[c] *Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & Royal Academy of Music Aarhus/Aalborg, Aarhus, Denmark*
[d] *Department of Education, Psychology, Communication, University of Bari Aldo Moro, Italy*
[e] *Department of Psychology, University of Chinese Academy of Sciences, Beijing, China*
[f] *CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai, China*
[g] *Chinese Institute for Brain Research, Beijing, China*
[h] *Center for Functionally Integrative Neuroscience, Dept of Clinical Medicine, Aarhus University, Aarhus, Denmark*
[i] *Music Cognition Lab, Queen Mary University of London, London, England, UK*

ARTICLE INFO

Keywords:
Music perception
Cultural distance
Predictive coding
Enculturation

ABSTRACT

According to the *cultural distance hypothesis (CDH)*, individuals learn culture-specific statistical structures in music as internal stylistic models and use these models in predictive processing of music, with musical structures closer to their home culture being easier to predict. This cultural distance effect may be affected by domain-specific (musical ability) and domain-general individual characteristics (openness, implicit cultural bias). To test the CDH and its modulation by individual characteristics, we recruited Chinese and Western adults to categorize stylistically ambiguous and unambiguous Chinese and Western melodies by cultural origin. Categorization performance was better for unambiguous (low CD) than ambiguous melodies (high CD), and for in-culture melodies regardless of ambiguity for both groups, providing evidence for CDH. Musical ability, but not other traits, correlated positively with melody categorization, suggesting that musical ability refines internal stylistic models. Therefore, both cultures show musical enculturation in their home culture with a modulatory effect of individual musical ability.

1. Introduction

Music provides unique opportunities to study and understand both culture and mind (Huron, 2008). Despite the famous poetic notion that "music is the universal language of mankind" (Longfellow, 1835) culture-specific variations in music are widely acknowledged. Throughout a lifetime of exposure to music containing these patterns of variations, listeners become enculturated to the structural regularities of the musical cultures within which they develop (Hannon & Trainor, 2007; Hansen, Kragness, Vuust, Trainor, & Pearce, 2021; Matsunaga, Hartono, & Abe, 2015; McDermott, Lehr, & Oxenham, 2010; Pearce, 2018), and this inevitably promotes an implicitly and culturally biased perception of music (Demorest, Morrison, Jungbluth, & Beken, 2008; Demorest, Morrison, Nguyen, & Bodnar, 2016) and the world in general (Yoon, Langrehr, & Ong, 2011). While globalization has homogenized cultures and Western music is influencing all others, we can still count as many musical cultures as there are musical styles (Holton, 2000; Huron, 2008; Nettl, 2005), with a musical culture operationally considered to consist of music created within a society that shares particular characteristics distinguishing it from other musical styles.

Demorest and Morrison (2015) attempted to formalize the notion of musical enculturation with their '*cultural distance hypothesis*' (CDH) which states that "the degree to which the music of any two cultures differ in the statistical patterns of pitch and rhythm will predict how well a person from one of the cultures can process the music of the other" (p. 189). This hypothesis has been expressed in quantitative terms using the

* Correspondence to: M Klarlund, Universitetsbyen 3, Building 1710, 8000 Aarhus C, Denmark.
** Correspondence to: Y Du, 16 Lincui Road, Chaoyang district, Beijing 100101, China.
E-mail addresses: mklarlund@clin.au.dk (M. Klarlund), duyi@psych.ac.cn (Y. Du).

https://doi.org/10.1016/j.cognition.2023.105405
Received 9 April 2022; Received in revised form 2 February 2023; Accepted 8 February 2023
Available online 18 February 2023
0010-0277/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

'Information Dynamics of Music' (IDyOM) model (Morrison, Demorest, & Pearce, 2019; Pearce, 2018) IDyOM is one of a class of models that account for musical perception in terms of probabilistic prediction. The *Predictive Coding Theory* (PCT) (Clark, 2013; Heilbron & Chait, 2018; Walsh, McGovern, Clark, & O'Connell, 2020) *of music* (Koelsch, Vuust, & Friston, 2019; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009) offers a quantitative account for how musical culture might constrain perception. According to the theory, music perception is a process of reducing sensory prediction errors by continuously updating an internal model of the causes of the sensory world (for a review, see Shipp, 2016) ensured by heightened precision-weighted predictions (Foster Vander Elst, Vuust, & Kringelbach, 2021; Koelsch et al., 2019; Land, 2014; Vuvan & Hughes, 2019). Musical variability across cultures provides contextual cues that helps promote the acquisition and application of distinct internal models specific for each culture, separated by 'cognitive firewalls' (Huron, 2006; Stewart, 2007): parallel 'families' of rules, specific to different musical styles, genres and cultures (Haumann, Vuust, Bertelsen, & Garza-Villarreal, 2018; Huron, 2006). Appropriate model selection and application ensures appropriate predictions are generated, reducing cognitive load and facilitating processing for the listener (Huron, 2006 - ch: 11). However, an empirical psychological validation of the CDH as quantified by IDyOM is still lacking.

Musical culture is not the only domain-specific factor influencing internal predictive models during music listening. Musical ability and training lead to more sophisticated internal models that can predict music with greater precision and accuracy (Hansen & Pearce, 2014; Hansen, Vuust, & Pearce, 2016; Vuust & Witek, 2014). Additionally, individual variance in the precision of internal predictive models of music might also derive from domain-general implicit biases or attitudes. Implicit attitudes are typical guiders of social judgements and behavior (Greenwald & Lai, 2019), and specifically racial attitude behavior is thought to be more reliably assessed through implicit measures (Fazio, Jackson, Dunton, & Williams, 1995), than explicit measures, such as individual effortful inferential processing (Blair, Dasgupta, & Glaser, 2015). Although variability of musical cultures does not necessarily track consistently or exclusively along racial or ethnic lines, implicit cultural bias is posited as a potential modulator of perceptual sensitivity to musical cultures. Meanwhile, it is possible that individuals high in *openness to experience* (hereafter, openness), a personality trait, may experience (musical) cultures differently to people who are less open as a result of their "breadth, depth, and permeability of consciousness, and . . .[their]. . . recurrent need to enlarge and examine experience" (McCrae & Costa, 1997, p. 826). Furthermore, a recurring and long-term update of internal models has been linked to heightened openness, as it suggests an interest in exploring 'the new' (Hashkes, Van Rooij, & Kwisthout, 2017).

The primary aim of the current study was to demonstrate perceptual sensitivity to musical culture and empirically assess the CDH as embodied in IDyOM under a predictive coding framework. Specifically, we investigated whether and how cultural background, musical ability and domain-general attitudes affect the internal predictive models guiding music perception, whilst validating with human participants the cultural distance predictions resulting from IDyOM simulations. In two consecutive studies, we asked Chinese (study 1) and Western (study 2) listeners varying in musical ability to listen to Chinese and Western melodies varying in cultural distance and categorize them according to cultural origin. We used the IDyOM model to select Chinese and Western melodies that were either culturally *unambiguous* (high cultural distance) or *ambiguous* (low cultural distance). We hypothesized that 1) unambiguous melodies would be more stylistically predictable for enculturated listeners than ambiguous melodies and therefore be easier to categorize; 2) listeners would categorize both ambiguous and unambiguous in-culture melodies more successfully than out-of-culture melodies; and 3) an individual's ability to categorize melodies according to cultural origin would correlate positively with individual domain-general and domain-specific attitudes, namely implicit cultural bias, the personality trait of openness and musical ability.

2. Methods

2.1. Participants

2.1.1. Study 1 (Chinese participants)

A total number of 106 young, native, and Beijing residing Chinese participants took part in the study. Six participants reported some level of hearing loss and were excluded from analyses. The sample size was determined from a power analysis (G*Power), which indicated that a sample of 100 participants would yield a power >99% for a paired *t*-test (two-tailed) using a medium effect size (d = 0.5) or 50% power in the case of a weak effect size (d = 0.2), based on an a priori Type I error rate (α) of 0.05. All the remaining 100 participants (69 female, M = 21.8 years, SD = 2.2) had normal hearing (250–8000 Hz, pure-tone threshold ≤20 dB HL). A total of 8 participants had received >3 years of formal training on a musical instrument. Musical background was assessed to ensure high variability according to the Goldsmith Sophistication Index (see below). All participants gave written informed consent approved by the Institute of Psychology, Chinese Academy of Sciences and received 100 CNY ¥ (~13.80 EUR €) in monetary compensation for their efforts.

2.1.2. Study 2 (German participants)

A total number of 103 native German participants passed the initial screening of eligibility to partake in the study. German participants were chosen for our sample of Western participants, as the stimuli used consisted of German folk melodies. Participant requirements for eligibility were German citizenship, little to no musical training, age between 16 and 35 years, and self-reported normal hearing capacities. Participants were recruited using Prolific (www.prolific.co) [11/092022–30/112022]. Four participants did not complete the study in its entirety and were therefore excluded from analyses. A total of 14 participants had received 4–5 years of formal musical training with no participants exceeding this. All the remaining 99 participants completed the experiment successfully after having agreed to an informed consent approved by the Department of Education, Psychology and Communication, University of Bari, Aldo Moro, Italy. They all received 15 USD $ (~14 EUR €) in monetary compensation for their efforts.

2.2. Stimuli

Melodies for the categorization test were selected by cultural distance based on an IDyOM analysis (Morrison et al., 2019; Pearce, 2018). Two IDyOM models were trained: a *Chinese model* trained on a large corpus of Chinese folk music; and a *Western mod*el trained on a large corpus of German folk music, thus approximating a standard for a Western music tradition. Specifically, the melodies used to train the two IDyOM models were selected from a collection of 769 German hymns and folk melodies from the Essen Folk Song Collection (EFSC; Schaffrath, 1995, data sets *fink* and *erk*) for the Western model and 858 Chinese traditional folk songs from the EFSC (data sets *han* and *shanxi*) for the Chinese model (see supplementary materials for further details).

When applied to a set of musical stimuli, each of these models estimates the conditional probability of each note in each stimulus, given the preceding musical context and the prior training of the model. Taking the average negative log probability (or *information content, IC*) provides an overall measure of the unpredictability of the stimulus. Following Pearce (2018), the divergence in unpredictability between the two models yields a quantitative measure of cultural distance (Morrison et al., 2019). Stimuli which are more predictable for the Chinese than the Western model are more strongly characteristic of the idiosyncrasies of Chinese musical culture and vice versa. Stimuli which are equally predictable for the two models are stylistically ambiguous. IDyOM was configured to use the trained long-term model to predict the pitch of each note in each melody based on a pitch-interval representation.

As shown in Fig. 1A, 80 novel melodies not previously used for IDyOM model training were selected as stimuli from the same EFSC datasets; 40 Chinese melodies (20 ambiguous and 20 unambiguous) and 40 Western melodies (20 ambiguous and 20 unambiguous) based on cultural distance. Melodies with a cultural distance approaching zero were considered 'ambiguous' (M = 0.16, SD = 0.12) while melodies with a cultural distance of relatively high absolute value (M = ± 0.83, SD = ± 0.13) were regarded as 'unambiguous' (see supplementary materials for further details). The melodies were converted from *MIDI* format (timbre-less datafiles) into *wav* files and played on a digital orchestral harp to give them a culturally ambiguous timbre. All melodies were kept monophonic with original tonality and rhythm and featured unvaried dynamic levels and an unvaried internal tempo. The duration of the melodies ranged from 6 to 19 s.

2.3. Procedure

2.3.1. Study 1 (Chinese participants)

Each participant was placed by a desk with a lab-computer (Macbook Pro vers. 2018) on which all tests were administered. The participants were left alone in the experimental room to avoid any type of distraction and were only accompanied by the experimenter when a new test had to be presented or when in need of help to clarify potential misunderstandings or computational complications. The procedure was structurally identical for all participants and took between 80 and 90 min to complete. A schematic visualization of the procedure is shown in Fig. 1B. The Implicit Association Test (IAT) was purposely positioned as the first test to circumvent any influence from other tests on the reaction times recorded during the IAT. This was followed by the Goldsmiths Musical Sophistication Index (Gold-MSI) questionnaire, the melodic categorization test, a personality questionnaire and the Musical Ear Test (MET). Each of these tests and questionnaires will be presented in more detail below.

2.3.2. Study 2 (German participants)

For the Western (i.e. German) participants, we created a project in the Gorilla platform (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020), which included online versions of the exact tasks the Chinese participants were exposed to. Specifically, we first created a 'musicianship screening node' only available to eligible participants as screened through Prolific (www.prolific.co) (screened for age and nationality), followed by an 'information consent node'. The procedure from there mirrored the paradigm for the Chinese participants, with IAT being the first task, followed by (in order) Gold-MSI, the melodic categorization task, a personality questionnaire, and MET. Standardized German version of Gold-MSI, the personality questionnaire and the valenced words of IAT were used, while all remaining tasks including information content were translated to the German language by a native German researcher to encourage the most natural responses from the participants and to circumvent any interpretation issues. All tasks were identical to the Chinese counterpart, including ratio of practice/test-trials, visual and melodic stimuli, randomizations, and instructions. Notably, the MET, Gold-MSI and the personality questionnaire did ask participants to answer using a computer mouse, instead of answering on paper.

2.4. Melodic categorization task

A Python-based application was developed and used to administer the categorization task. Participants listened to the melodies at a comfortable listening level (approx. 75 dB SPL) through Sennheiser HD650 headphones. Preceding the categorization test, a training trial using four culturally unambiguous melodic soundbites (two Western, two Chinese) played either on harp or piano was included. Participants were asked to select which instrument was featured in the soundbite, provide a confidence rating, and state if they had heard the soundbite before. The design of the training trial screen was identical to the main trial screen and participants were made aware that the training trial would not be included in the formal task.

The categorization test was structured into four blocks with twenty melodies occupying each. Blocks were evenly assigned to either ambiguous or unambiguous melodies, such that two blocks contained unambiguous melodies exclusively while the other two blocks contained

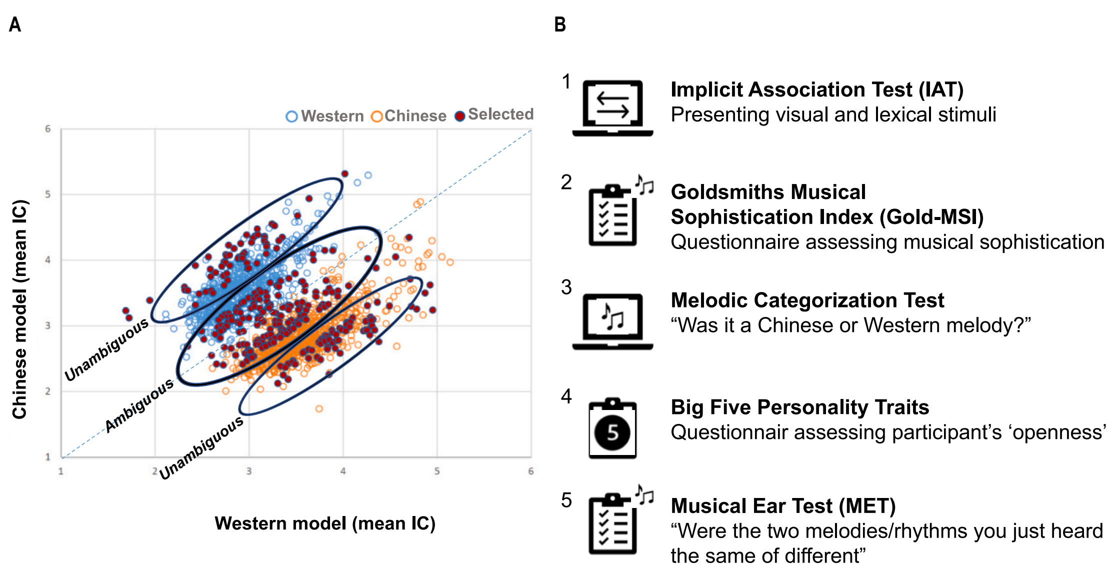

Fig. 1. Stimuli and experimental procedure. (A) Selected melodies in a plot of averaged IC values from two IDyOM models: the Chinese model and the Western model. The blue dashed line represents equivalence between the two models, while the cultural distance of a melody is calculated as it's perpendicular distance from the line of equivalence. From a total of 381 melodies, a selection of 80 melodies (40 Chinese and 40 Western) was chosen based on cultural distance extremes (ambiguous or unambiguous) while ensuring equal variance in culture-agnostic melodic predictability (indicated by the line of equivalence with melodies further to the top right being more unpredictable). Blue and orange-outlined circles depict Western and Chinese melodies respectively. Red-filled circles represent melodies chosen for the experiment. (B) Structure of the experimental procedure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

only ambiguous melodies. For each participant, the ambiguous melodies (both Western and Chinese) were played first, but in a randomized order. Participants were asked to categorize each melody as either Chinese or Western, provide a confidence rating on a five-point Likert scale (1 = very uncertain, 5 = very certain) and declare if they knew the melody beforehand. Written instructions were presented in native languages for both participant groups.

2.5. Questionnaires and tests for personal traits

2.5.1. Implicit association test

Participants' implicit attitudes (i.e., bias) towards Chinese and Western cultures were measured using an open-source version of the IAT (Greenwald, Nosek, & Banaji, 2003; Nosek, Greenwald, & Banaji, 2007) through PsychoPy (version 3.2.3, Meade, 2009) The IAT measures participants' response times in sorting pictures of ethnically Western and Chinese faces with attributing positive and negative words. Facial pictures were acquired from NimStim Set of Facial Expressions database (Tottenham et al., 2009) Ten pictures were used, five for each ethnicity, three males and two females. All facial expressions were neutral. Pictures were converted into grayscale and cropped from eyebrows to chin and from ear to ear (350 × 225 pixels). Ten non-ambiguous valenced Chinese words were used with five positive words (欢乐/freude: joy, 爱/liebe: love, 和平/freiden: peace, 愉悦/vergnügen: pleasure, 欢笑/lachen: laughter) and five negative words (痛苦/qual: agony, 邪恶/übel: evil, 可怕/schrecklich: terrible, 糟糕/scheußlich awful, 刺痛/verletzt: hurt). Results from the IAT were obtained through an accompanying R-script as 'D-values' ($D = (mean\_latency(Chinese + positive; Western + negative) - mean\_latency(Chinese + negative; Western + positive))/SD$), ranging from −2 to 2 with break points indicating: 'slight' (0.15), 'moderate' (0.35) and 'strong' (0.65) bias. Negative values describe a Western bias and positive values a Chinese bias.

2.5.2. Personality questionnaires

Big Five Personality questionnaire (Study 1). To examine each participant's openness to experience, a Chinese version of the Revised NEO Five-Factor Inventory (NEO-FFI-R) was administered in paper form. The NEO-FFI-R contains 60 items (12 per domain) with answers given on the five-point Likert scale (McCrae & Costa, 2004). While only the items associated with openness were used for the ensuing analysis, participants were asked to answer the entire inventory to avoid biased responses.

Ten item personality inventory (TIPI) (Study 2). An executive decision was made to substitute the longer NEO-FFI-R with the shorter TIPI (Gosling, Rentfrow, & Swann, 2003) for the online version of the current paradigm. This was done to lower the estimated session duration for an online methodology where fatigue and loss of focus are more often reported (Savage & Waldman, 2008). Furthermore, TIPI has been shown to reach comparable results to other 'big-five-inventories', such as the NEO-FFI and NEO-PI-R (Gosling et al., 2003). By substituting the NEO-FFI-R with TIPI, the overall length of the paradigm went from approx. 90 min to ~80 min.

2.5.3. Goldsmiths musical sophistication index questionnaire

The Gold-MSI was used to obtain a self-reported assessment of musical ability (Müllensiefen, Gingras, Musil, & Stewart, 2014). Although a standardized German version (Schaal, Bauer, & Müllensiefen, 2014) exists, the English version of the Gold-MSI questionnaire (https://www.gold.ac.uk/music-mind-brain/gold-msi/download/) was translated into Chinese by the experimenters, as there was no standardized Mandarin version. Participants were instructed to answer the questionnaire as accurately as possible with no time limit. All items were scored on the same 7-point scale and received equal weights for scoring.

2.5.4. Musical ear test

The MET was used for measuring participants' musical ability since it is short (~20 mins) and applicable to individuals regardless of musical training experience (Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). To accommodate Chinese participants, the original English instruction was translated into Chinese and recorded by a native Chinese speaker (Zhang, Xie, Li, Shu, & Zhang, 2020). German participants were presented with written instructions to the MET in German, whilst the counting between trials during the task was kept in English. MET consists of a melodic subtest with 52 pairs of melodic phrases, and a rhythm subtest with 52 trials of rhythmic phrases. Participants were instructed to judge whether the two musical phrases were identical or different. Scores for the melodic subtest, rhythm subtest and both were used.

2.6. Statistical analyses

Responses to melodies that participants reported having heard prior to the experiment amounted to a total of 330 trials and were excluded from all analyses. To test the first two hypotheses regarding stylistic categorization, we applied two and three-way mixed ANOVAs to the classification performance for each melody with the within-subject factors *cultural origin* (Chinese vs. Western) and *cultural distance* (unambiguous vs. ambiguous), and the between-subjects factor *participant group* (Chinese vs. Western), followed by pairwise comparisons with Bonferroni correction. We measured categorization performance by *confidence-weighted accuracy,* a more sophisticated measure than simple ratios of correct/incorrect responses, which includes confidence as a weighting to the responses:

$$Confidence - weighted\ accuracy = (N._{correct} \times Confidence_{correct}) - (N._{incorrect} \times Confidence_{incorrect})$$

where N represents number of responses.

To test the third hypothesis, Pearson correlations were calculated between performance and features of interest (bias, openness, musical ability).

3. Results

3.1. Omnibus categorical analysis

A three-way mixed ANOVA showed a significant interaction between cultural origin of melodies (Chinese, Western), cultural distance (ambiguous, unambiguous), and participant group (Chinese, Western) ($F(1, 198) = 7.14$, $p = .008$, $\eta2 = 0.04$), motivating four further 2-way ANOVAs to understand the nature of the interaction.

3.2. Overall cultural distance effect

For testing *Hypothesis 1* (Fig. 2A), a two-way mixed ANOVA showed a significant main effect of cultural distance with better performance for unambiguous melodies than ambiguous melodies ($F(1,198) = 465.99$, $p < .001$, $\eta2 = 0.70$). There was neither a significant interaction between participant group (Chinese, Western) and cultural distance (ambiguous, unambiguous) ($F(1, 198) = 2.647$, $p = .105$, $\eta2 = 0.01$), nor a significant main effect of group ($F(1, 198) = 1.91$, $p = .168$, $\eta2 = 0.01$). This provides empirical evidence for the cultural distance effect, which was

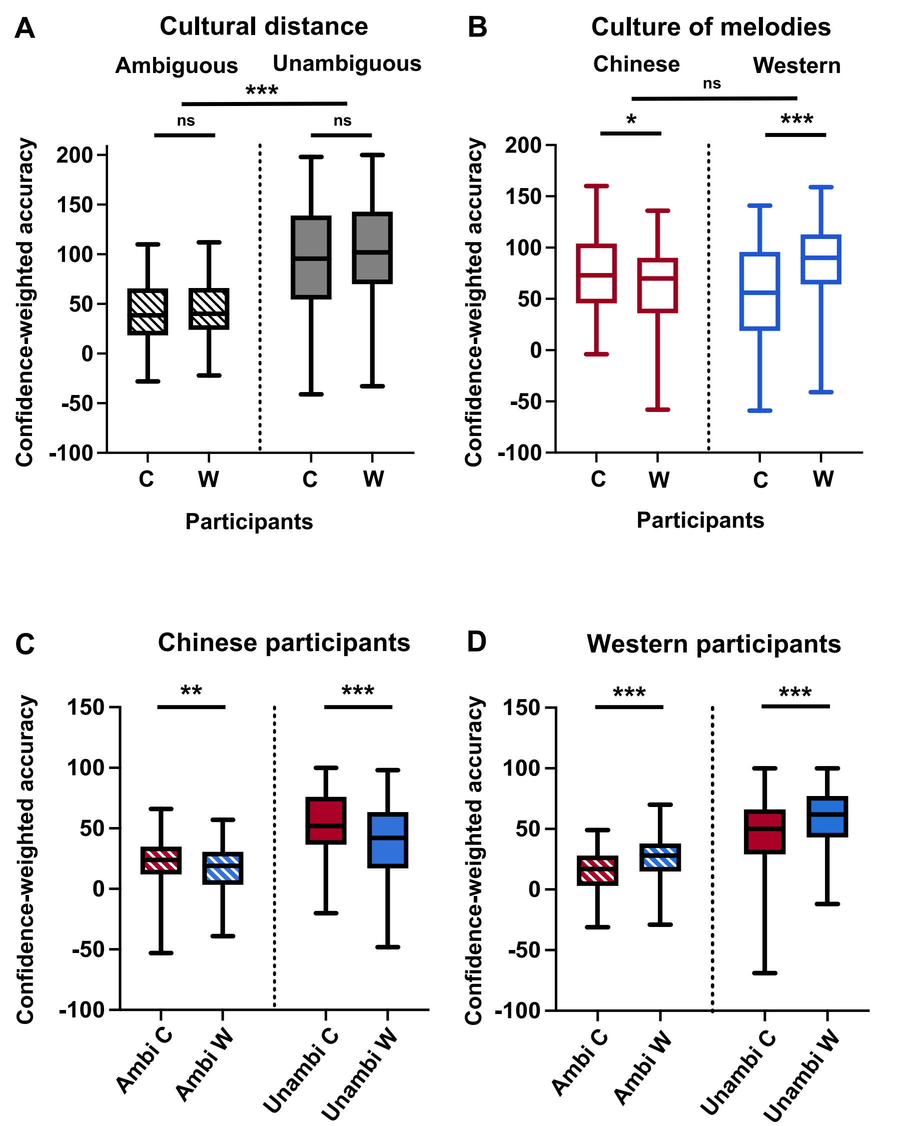

Fig. 2. Cultural distance effect and in-culture advantage on categorization performance for Chinese and Western participants. (A) Confidence-weighted accuracy for categorizing ambiguous (cross-hatching) and unambiguous (filled) melodies regardless of melodic cultural origin for Chinese and Western participants. (B) Confidence-weighted accuracy for categorizing Chinese (red) and Western (blue) melodies regardless of cultural distance for both participant groups. (C) and (D) Confidence-weighted accuracy for categorizing ambiguous (cross-hatching) and unambiguous (filled) melodies from both Chinese (red) and Western (blue) cultures in Chinese participants and Western participants, respectively. *** $p < .001$, ** $p < .01$, * $p < .05$, ns > 0.05. C, Chinese; W, Western; Ambi, ambiguous; Unambi, unambiguous. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

equally strong between participant groups.

3.3. General in-culture advantage

For testing *Hypothesis 2* (Fig. 2B), a two-way mixed ANOVA revealed a significant interaction between participant group (Chinese, Western) and cultural origin of melodies ($F(1, 198) = 71.09$, $p < .001$, $\eta2 = 0.27$), but did not show a significant main effect of cultural origin of melodies ($F(1, 198) = 0.06$, $p = .813$, $\eta2 = 0.01$). Furthermore, there was a significant difference in confidence-weighted accuracy between groups for Chinese melodies ($F(1,198) = 7.10$, $p = .008$, $\eta2 = 0.04$) and for Western melodies ($F(1, 198) = 21.94$, $p < .001$, $\eta2 = 0.10$). Pairwise comparisons showed significant differences between participant groups regarding the in-culture advantage for both Chinese melodies (confidence-weighted accuracy was 14.68 percentage points higher for Chinese participants, $p = .008$) and Western melodies (confidence-weighted accuracy was 29.28 points higher for Western participants, $p < .001$). This suggests an in-culture advantage for both participant groups with respect to the cultural origin of melodies.

3.4. Interaction between cultural distance and cultural origin

Two-way repeated measures ANOVAs with independent variables cultural distance (ambiguous, unambiguous) and cultural origin (Chinese, Western) were conducted separately for each participant group. The results revealed a significant interaction for Chinese participants between cultural distance (ambiguous, unambiguous) and cultural origin of melodies (Western, Chinese) present ($F(1, 99) = 10.92$, $p = .002$, $\eta2 = 0.12$) as well as significant main effects for both cultural origin of melodies ($F(1, 99) = 28.35$, $p < .001$, $\eta2 = 0.22$) and cultural distance ($F(1, 99) = 200.40$, $p < .001$, $\eta2 = 0.67$). Although no significant interaction effect was found for Western participants ($F(1, 98) = 0.59$, $p < .442$, $\eta2 = 0.01$), the analyses did reveal significant main effects of both cultural origin of melodies ($F(1, 98) = 33.36$, $p < .001$, $\eta2 = 0.25$) and cultural distance ($F(1, 98) = 271.88$, $p < .001$, $\eta2 = 0.74$).

For Chinese participants (Fig. 2 C), pairwise comparisons showed significantly higher confidence-weighted accuracy in categorizing Chinese (in-culture) melodies than Western (out-of-culture) melodies for both stylistically ambiguous ($p = .004$, *Cohen's d* $= 0.33$), and unambiguous melodies ($p < .001$, *Cohen's d* $= 0.63$). A paired two-tailed *t*-test revealed significantly larger difference between ambiguous and unambiguous melody categorization for Chinese than Western stimuli ($t(99) = 3.21$, $p = .002$, *Cohen's d* = 0.33).

For Western participants (Fig. 2D), pairwise comparisons showed significantly higher confidence-weighted accuracy in categorizing Western (in-culture) melodies than Chinese (out-of-culture) melodies for both stylistically ambiguous ($p < .001$, *Cohen's d* = 0.60), and unambiguous melodies ($p < .001$, *Cohen's d* = 0.46). Unlike for Chinese participants, a paired two-tailed *t*-test between ambiguous and unambiguous melody categorization for Chinese than Western stimuli did not reveal a significant difference ($t(98) = 0.77$, $p = .442$, *Cohen's d* = 0.09).

3.5. Correlations between individual traits

Table 1 displays summary statistics of individual traits for Chinese and Western participants. For Chinese participants, IAT results showed a significant positive bias in implicit attitudes towards in-culture Chinese faces ($M = 0.31$, $SD = 0.37$, one-sample t-test: $t(99) = 8.56$, $p < .001$, *Cohen's d* = 0.41). Similarly, for Western participants, IAT results showed a significant bias in implicit attitudes towards in-culture Western faces ($M = -0.19$, $SD = 0.40$, one-sample t-test: $t(98) = 4.68$, $p < .001$, *Cohen's d* = 0.29).

For Chinese participants (Fig. 3 top), the melodic subpart and total MET scores correlated positively with all subscales of the self-reported musical sophistication questionnaire (Gold-MSI) (all $r > 0.32$, $p < .001$), except for 'emotion' (all $r < 0.16$, $p > .131$). The rhythmic subpart of MET also positively correlated with all Gold-MSI subscales, but to a slightly lesser degree (all $r > 0.22$, $p < .230$), except for 'emotion' ($r = 0.13$, $p = .203$). Different from Chinese participants, MET and Gold-MSI correlations were more sporadic in Western participants (Fig. 3 bottom); MET-melody correlated with all Gold-MSI subscales (all $r > 0.20$, $p < .043$), except for 'active engagement' ($r = 0.07$, $p = .480$), while MET-rhythm only correlated positively with 'perceptive ability' ($r = 0.23$, $p = .020$). MET-total correlated with subscales of 'perceptual ability', 'singing ability' and 'general sophistication' (all $r > 0.25$, $p < .010$), while 'emotion', 'musical training' and 'active engagement' failed to show significance in the correlation analyses (all $r < 0.20$, $p > .054$).

Additionally, for Chinese participants, all Gold-MSI subscales correlated positively with openness (all $r > 0.21$, $p < .029$), while only MET-melody correlated with openness ($r = 0.27$, $p = .006$), but not the rhythmic nor the total MET (all $r < 0.17$, $p > .090$). IAT did not correlate with any other trait in Chinese participants (all $|r| < 0.14$, $p > .205$). However, for Western participants, IAT and openness correlated positively with the 'emotion' subscale of Gold-MSI as the only considered feature (*IAT:* $r = 0.22$, $p = .030$; *Openness:* $r = 0.20$, $p = .045$). Notably, a positive correlation between IAT and emotion suggests that emotional experience of music may influence the cultural bias away from home-culture, as the in-cultural bias for Western was associated with negative values in IAT. Openness did not correlate with other features in Western participants (all $|r| < 0.19$, $p > .071$).

Pearson r values and significance indications (*** $p < .001$, ** $p < .01$, * $p < .05$, ns > 0.05) of individual trait correlations. The bold *** outside the matrix indicates that everything within the square is strongly correlated ($p < .001$), unless stated otherwise.

3.6. Correlations between categorization performance and individual traits

For testing *Hypothesis 3*, as shown in Table 2, MET positively correlated with categorization performance for both ambiguous and unambiguous melodies in Chinese participants (all $r > 0.21$, $p < .005$), and for both ambiguous Western and unambiguous melodies in Western participants (all $r > 0.26$, $p < .025$). To exemplify, the correlations between MET-total and confidence-weighted accuracy are visualized in Fig. 4., In addition, Gold-MSI subcategories 'perceptual ability', 'musical training', 'singing ability' and 'general sophistication' correlated with categorization performance only for unambiguous melodies in Chinese participants (all $r > 0.21$, $p < .025$). For Western participants, Gold-MSI subcategories 'perceptual ability', 'musical training', 'emotion' and 'general sophistication' correlated with categorization performance for ambiguous Western and unambiguous melodies (all $r > 0.20$, $p < .046$). Neither implicit bias nor openness yielded significant results in both groups (all $r < 0.15$, $p > .135$). These results suggest that musical ability rather than implicit bias or openness is the critical personal trait that mediates musical categorization performance according to cultural origin.

4. Discussion

The experiments reported here aimed to produce the first empirical validation of the cultural distance hypothesis by examining the ability of Chinese and Western listeners to identify the cultural origin of Chinese and Western melodies varying in cultural distance as predicted by IDyOM, a computational model of musical expectation. We also investigated how categorization performance was modulated by domain-general individual traits (implicit cultural bias, openness) and domain-specific abilities (musical ability). The results corroborate the cultural distance hypothesis, showing better categorization performance for stylistically unambiguous than ambiguous melodies in both participant groups, and an in-culture advantage for both unambiguous and ambiguous melodies, such that in-culture melodies were better categorized than out-of-culture melodies for both Chinese and Western participant groups.

Moreover, in both groups, of the three individual traits (implicit bias, openness, and musical ability) hypothesized to modulate enculturated individuals' ability to correctly classify melodies, only musical ability assessed by MET significantly correlated with categorization performance for both ambiguous and unambiguous melodies. In contrast, musical ability assessed through the self-reported musical sophistication index (Gold-MSI) showed significant correlations only for unambiguous melodies in both groups and unambiguous melodies in Western (but not Chinese) participants. One potential reason for this is the way musical ability is assessed by Gold-MSI, which is a self-report questionnaire, and by MET, which is a performance test, and thus their proposed influence on musical categorization performance.

Table 1
Summary statistics for individual traits.

	Chinese participants				Western participants			
	Max	Min	Mean	Std	Max	Min	Mean	Std
Gold-MSI (GS)	110	28	66.2	15.7	88	12	40.5	15.8
MET total (%)	89.4	48.1	71	8.43	93.3	31.7	69.8	12.1
IAT	1.26	−0.69	0.31	0.37	0.86	−1.04	−0.19	0.40
Openness(NEO-FFI-R/TIPI)	49	27	38.9	4.55	7	1.5	5.10	1.17

GS = general sophistication.
NEO-FFI-R (Revised NEO Five Factor Inventory) – Chinese Participants.
TIPI (Ten-Item Personality Inventory) – Western participants.

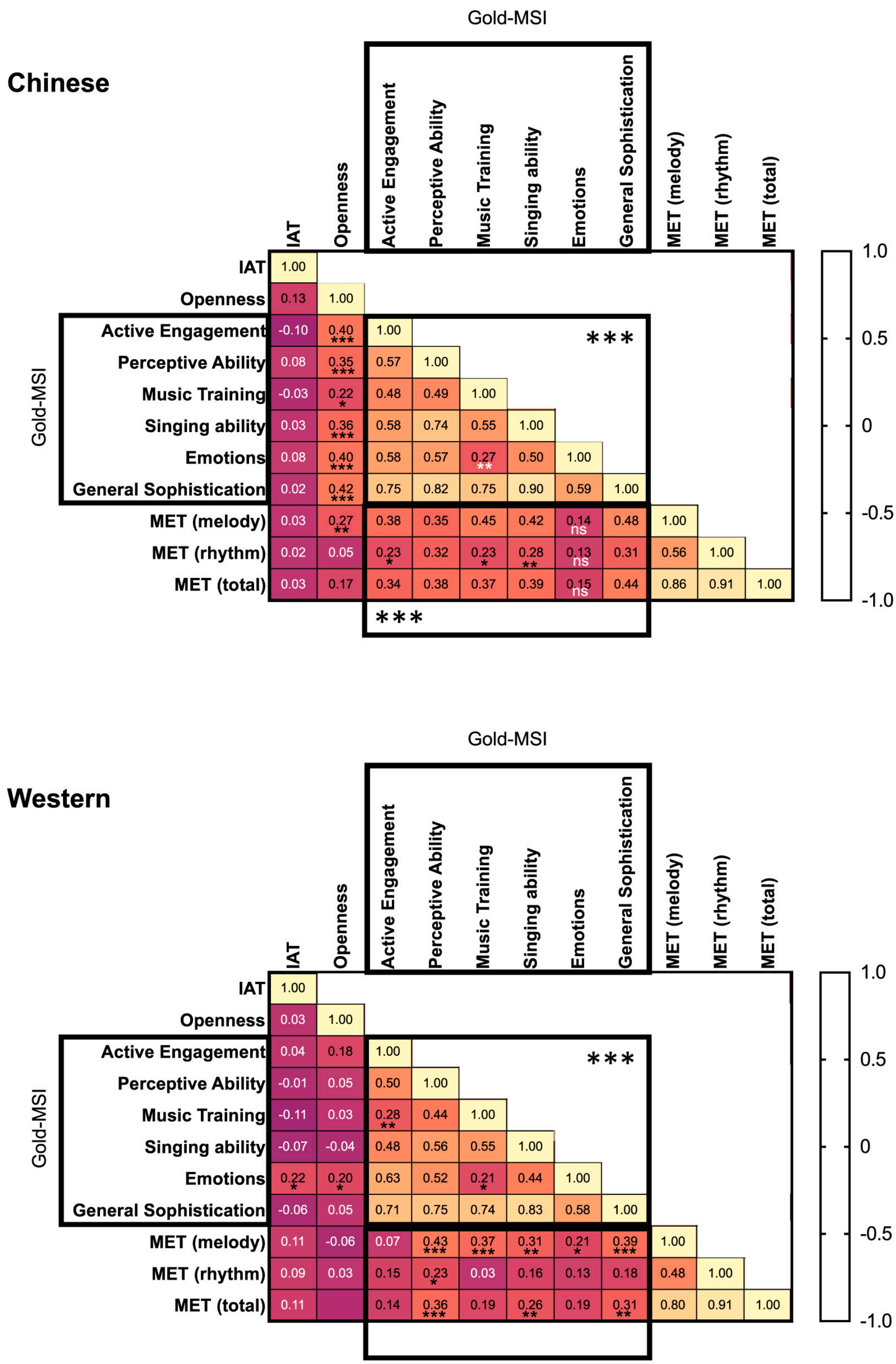

Fig. 3. Correlation matrix of domain-general and domain-specific traits.

4.1. *In-culture advantage*

The in-culture advantage presented in this paper mirrors prior observations in which listeners from specific musical cultures have shown processing advantages for the music of their own culture compared to a comparison group from a different musical culture (see Morrison et al., 2019 for a review). For example, when listening to North Indian music, the expectations of North Americans failed to capture certain characteristics of North Indian musical culture, which was reflected in the expectations of Indian listeners (Castellano, Bharucha, & Krumhansl, 1984); further, Turkish listeners showed better recognition memory performance for Turkish than Western music, whereas listeners from the USA showed the converse pattern (Demorest et al., 2008). Presumably the stimuli in many of these studies were selected to correspond

Table 2
Pearson correlation coefficients and FDR-corrected *p*-values (in parentheses) between various individual traits and categorization performance for ambiguous and unambiguous melodies.

	Chinese participants				Western participants			
	Performance for ambiguous melodies		Performance for unambiguous melodies		Performance for ambiguous melodies		Performance for unambiguous melodies	
	Chinese	Western	Chinese	Western	Chinese	Western	Chinese	Western
IAT	−0.07 (0.407)	0.04 (0.600)	−0.06 (0.446)	−0.01 (0.449)	−0.07 (0.886)	0.07 (0.402)	−0.06 (0.301)	0.10 (0.156)
Openness	0.12 (0.291)	0.06 (0.560)	0.15 (0.129)	0.10 (0.195)	−0.05 (0.886)	−0.04 (0.430)	−0.15 (0.098)	−0.10 (0.156)
Gold-MSI								
Active engagement	0.11 (0.291)	0.04 (0.600)	0.15 (0.129)	0.15 (0.087)	−0.11 (0.886)	−0.04 (0.430)	−0.02 (0.406)	0.01 (0.359)
Perceptual ability	0.20 (0.107)	0.20 (0.118)	**0.28 (0.038)***	**0.28 (0.009)****	−0.04 (0.886)	**0.30 (0.004)****	**0.34 (0.001)****	**0.38 (0.001)****
Musical training	0.10 (0.421)	0.13 (0.261)	0.21 (0.058)	**0.26 (0.009)****	0.09 (0.886)	**0.28 (0.006)****	0.17 (0.071)	**0.22 (0.017)***
Emotions	0.10 (0.338)	0.10 (0.409)	0.15 (0.129)	0.15 (0.087)	0.09 (0.886)	−0.01 (0.533)	**0.24 (0.020)***	**0.24 (0.015)***
Singing ability	0.14 (0.252)	0.16 (0.180)	0.17 (0.117)	**0.20 (0.037)***	−0.23 (0.231)	0.13 (0.187)	−0.01 (0.268)	0.05 (0.252)
General sophistication	0.15 (0.250)	0.16 (0.180)	**0.22 (0.049)***	**0.26 (0.009)****	0.02 (0.886)	**0.20 (0.045)***	**0.23 (0.022)***	**0.25 (0.009)****
MET								
Melody	**0.27 (0.017)***	**0.27 (0.028)***	**0.25 (0.041)***	**0.27 (0.009)****	−0.04 (0.886)	**0.48 (0.003)****	**0.43 (0.001)****	**0.56 (0.001)****
Rhythm	**0.28 (0.017)***	**0.24 (0.050)***	**0.22 (0.049)***	**0.26 (0.009)****	−0.03 (0.886)	**0.29 (0.006)****	**0.34 (0.001)****	**0.27 (0.006)****
Total	**0.31 (0.017)***	**0.29 (0.028)***	**0.26 (0.038)***	**0.30 (0.009)****	−0.04 (0.886)	**0.42 (0.003)****	**0.44 (0.001)****	**0.45 (0.001)****

[*] FDR-corrected $p < .05$.

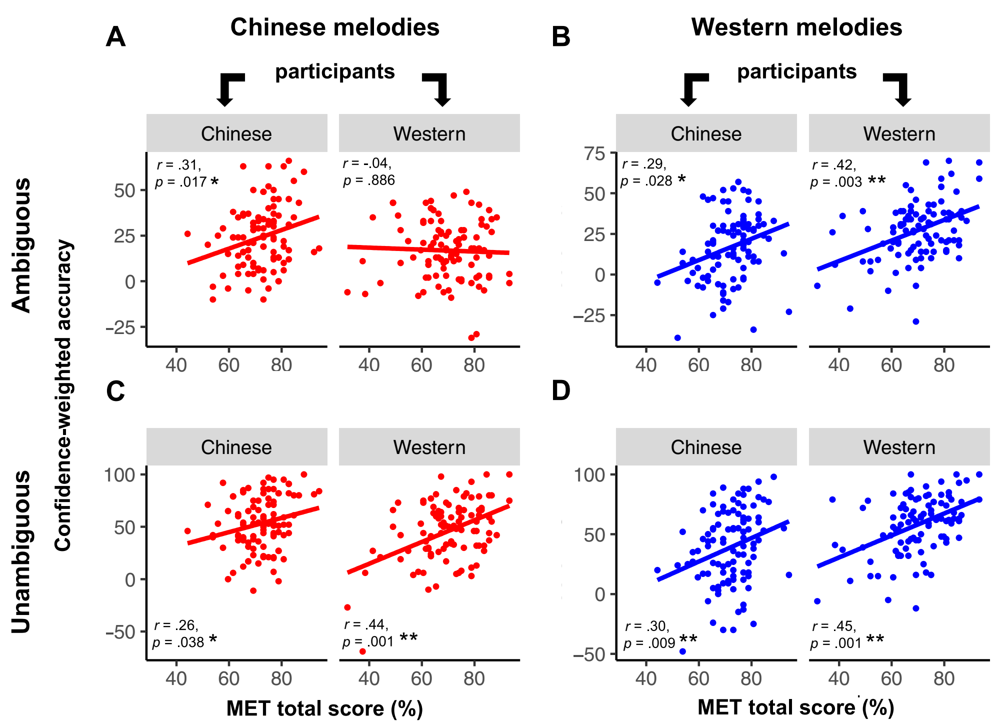

Fig. 4. Correlations between MET-total score and confidence-weighted accuracy. Significant correlations were found for (A) Chinese ambiguous melodies, (B) Western ambiguous melodies, (C) Chinese unambiguous melodies and (D) Western unambiguous melodies in both groups. Note that Western participants did not show any correlation between MET total and categorization performance for Chinese ambiguous melodies. ** $p < .01$, * $p < .05$ by Pearson's correlations.

conceptually to our 'unambiguous' stimuli in which the stylistic differences between stimuli from each culture were readily apparent. Therefore, while these previous findings demonstrate an in-culture advantage they do not show an effect of cultural distance. By systematically varying cultural distance (ambiguous vs unambiguous stimuli), the present results contribute to this literature, the first empirical evidence for a computational model of the psychological processes underlying such enculturation effects.

4.2. Culture distance effect

The construct of cultural distance and the significant effects presented here between two extreme points (high: unambiguous, low: ambiguous) have both theoretical and methodological implications. Since culture as a phenomenon is notoriously difficult to conceptualize and measure (Chp 4 - Sparrow, 2009), establishing a transparent means to capture and quantify cultural differences has been sought for a long time. Cultural distance quantified by IDyOM offers a tool for quantitative assessment that can potentially bypass some of the intangible and complex structures that make up cultural variability. While having limitations (as discussed below), the present results suggest that musical cultures, and the distances between them, can be conceived in terms of patterns or regularities in sequential pitch structure which are acquired by listeners through statistical learning of the music to which they are exposed during enculturation. Specifically for the present study, cultural distance offers a means to quantify musical culture in a way that captures a particular melody's cultural belongingness relative to statistical regularities capturing the norms of the musical culture from which the melody was taken. The fact that listeners responded consistently with this conception of cultural distance suggests that they perceive music using psychological mechanisms of statistical learning and probabilistic prediction, akin to those implemented in the IDyOM model (Pearce, 2018). Cross-cultural differences have previously been shown to exist for rhythmic priors (Jacoby et al., 2021), perception of musical meter (Hannon & Trehub, 2005a, 2005b), tonality perception (Raman & Dowling, 2016, 2017), perception of grouping structure (Ayari & McAdams, 2003; Iversen, Patel, & Ohgushi, 2008; Nan, Knösche, & Friederici, 2006, 2009; Nan, Knösche, Zysset, & Friederici, 2008), memory for music (Demorest et al., 2008) and emotion recognition (Balkwill & Thompson, 1999; Fritz et al., 2009), without accounting for quantitative differences in strength of in-culture advantage, that are provided by cultural distance. We speculate that cultural distance based on statistical learning and probabilistic prediction provides a psychological mechanism that can both shed new light both on existing findings and inspire hypothesis-based research to understand the intricate psychological factors underlying the differences between musical cultures.

4.3. Modulatory effects of musical ability

Musical ability is thought to be related to a wide range of factors such as general intelligence, domain specific skills and deliberate practice (Burgoyne, Harris, & Hambrick, 2019; Detterman & Ruthsatz, 1999; Ruthsatz, Detterman, Griscom, & Cirullo, 2008), and is known to correlate with music-related (e.g., Demorest et al., 2008) and nonmusical (i.e., speech and language) tasks (e.g., Anvari, Trainor, Woodside, & Levy, 2002; Y. Du & Zatorre, 2017). It is likely that domain-specific musical ability, rather than traits such as openness and implicit cultural bias, are associated with the possession of more detailed and sophisticated models of the musical styles which, in turn, facilitates categorization performance.

Consistent with this proposal, the MET measure was overall more sensitive than the Gold-MSI self-report questionnaire in predicting ambiguous melody categorization, and especially so for Chinese participants, but also more consistent across conditions for Western participants. This may be because musical ability performance tests are less susceptible to subjective biases, which has important implications for the assessment of musical ability in general (Seashore, 1915). Another (not mutually exclusive) possibility is that some components of the Gold-MSI, such as active engagement or emotional experience of music, may not influence stylistic categorization.

MET has been found to correlate with the expertise of musicians and enable differentiation of musicians and non-musicians (Wallentin et al., 2010), but our findings further support previous indications that MET scores also provide an index of expertise within non-musicians (Swaminathan, Kragness, & Schellenberg, 2021). Krumhansl (2000) found that differences in expectations for a non-Western style between expert and non-expert listeners could be characterized in terms of statistical models. In a follow up study, Eerola, Louhivuori, and Lebaka (2009) found that non-musician listeners rated the fitness of probe-tones in out-of-culture music in accordance with data-driven bottom-up predictions, whereas listeners of higher musical proficiency exhibited schema-driven top-down predictions, regardless of the cultural background of the listener (Western or African). In this regard, our results contribute to a literature showing optimization of predictive processing as a function of musical ability (Hansen et al., 2016; Hansen & Pearce, 2014). We suggest that listeners with higher MET score possess more sophisticated and robust stylistic models of musical structure acquired through statistical learning and enculturation, and this leads to an improved ability to distinguish different musical styles.

4.4. Openness and implicit bias

Musical ability is widely considered a domain-specific characteristic, while personality (focusing here on openness to experience) and implicit attitudes (bias for in-culture faces) are domain-general dispositional attributes. A cross-cultural effect of these distinctive characteristics was found when seeing how Gold-MSI scores (Fig. 3) correlated more strongly with domain-general feature (openness) for the Chinese participants, but for domain-specific scores among Western participants. One possible explanation could be that music sophistication in China is acquired more widely through primary and secondary education and is more noticeably seen as a way to promote not only artistic skills but also moral and intellectual development, including curiosity and broadened perspective (Ho, 2010; Law & Ho, 2011). In Germany, this link may be less pronounced, leading to the correlation of Gold-MSI with domain-specific musicality. Pending further confirmatory research, this remains speculative given that we didn't hypothesise any difference in this respect.

Openness has been previously found to predict musical ability (Swaminathan & Schellenberg, 2018), and positive attitudes towards a given culture have been shown to increase after listening to music from that culture (Vuoskoski, Clarke, & DeNora, 2017). In the present experiment, for Chinese participants, openness was correlated with all the subcategories and the general score of the self-assessment inventory of Gold-MSI, but not musical ability assessed through the performance test of MET. This was not the case for Western participants, where only the emotion subcategory from Gold-MSI correlated with openness. Openness has previously been demonstrated to correlate positively with formal musical training and thus helped to explain how openness predicts auditory discrimination tasks (Swaminathan & Schellenberg, 2018; Thomas, Silvia, Nusbaum, Beaty, & Hodges, 2016), but never discriminations between extreme levels of cultural ambiguity. One could speculate that while being open to new experiences might influence attitudes to unknown cultural customs, the task of differentiating cultural distances across musical cultures requires more enhanced domain-specific prediction sensitivity than the domain-general capabilities offered by traits such as openness.

Implicit bias did show in-group cultural favouritism, thus adding to the literature of similar findings for both Chinese (K. Du, Hunter, Scarf, & Ruffman, 2021) and Western (Rutland, 1999) individuals. The non-significant correlations reported in this study are indicative of implicit cultural biases that are distinguishable between cross-race face identification and cross-cultural musical norms.

4.5. Limitations

To fully highlight the influence of the enculturation effect on music perception, we suggest further research on neural correlates of musical categorization and its modulation by musical ability and cultural distance. Specifically, we acknowledge the cultural situatedness inherent in the melodic material of the MET as an instrument of musical ability. Although the tool has been validated across multiple cultures, including Chinese (Zhang et al., 2020), an underlying potential of measurement bias is not of irrelevant concern. Future cross-cultural research will benefit from the development of musical ability tests specifically targeted at music material relevant to the cultures of interest. Furthermore, current influential theories of perception and cognition (including PCT) have been conceived in a Western setting, and evaluation in a cross-cultural context is therefore imperative to the validation of these theories (Stevens, 2012). We acknowledge the apparent limitation of only one cultural comparison (Chinese and Western) and are convinced that future research addressing enculturation effects would benefit by adding multiple cultures in categorization tasks (see e.g., Demorest et al., 2008). Additionally, the IDyOM model used to generate measures of cultural difference was configured to predict only pitch. It would undeniably be beneficial to widen the range of parameters to include rhythm and polyphonic music (including harmony for musical cultures in which this would be appropriate), to fathom the musical nuances of cultural specificity when determining the cultural distance of the stimuli. We trained IDyOM on the Essen Folksong Collection as a best effort to approximate the overall stylistic experience of listeners within a given musical culture with available electronically-encoded musical corpora. The results suggest that this approximation holds but it would be interesting to assess the benefits of using training sets that more accurately represent the listening history of particular groups or individuals. Furthermore, given the null correlation results reported, it would be interesting to assess the influence of other individual traits such as empathy, conscientiousness, and musical preference. Finally, a more detailed assessment of intra-cultural background among participants could enable a fine-grained distinction of cultural differences. Especially so, when working categorically with Chinese and Western cultures, which undoubtably include extensive cultural diversity within their respective borders.

5. Conclusions

In sum, we found (1) a main effect of cultural distance in categorizing melodies, i.e., better performance for culturally unambiguous than culturally ambiguous melodies; (2) a significant in-culture advantage for both unambiguous and ambiguous melodies, i.e., higher performance for Chinese than Western melodies for Chinese participants, and vice versa for Western participants; (3) among the different individual factors, only domain-specific musical ability (MET) correlated positively with categorization performance for both ambiguous and unambiguous melodies, possibly reflecting more sophisticated cognitive models of stylistic structure in musically capable individuals.

This study is the first to provide an empirical validation of the cultural distance hypothesis and the findings presented here also suggest that musical ability modulates the precision of internal model selection in both an ambiguous and unambiguous musical context across Chinese and Western participants. We are hopeful that cultural distance will prevail as a variable in future studies addressing cognitive differences in music perception across cultures and that researchers will profit from the reliability of IDyOM categorizations. We feel confident that just as music perception in a predictive coding framework emphasizes the importance of cross-cultural music studies, cross-cultural studies also contribute to the generalizability of predictive coding as a theory of human cognition.

Funding

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (grant XDB32010300). The Center for Music in the Brain (MIB) is funded by the Danish National Research Foundation (DNRF, project number 117).

Open practices statement

'The study reported in the current manuscript was not preregistered. Deidentified data, processing scripts and stimuli files have been made publicly accessible at 'https://github.com/mathiasklarlund/worlds_apart'.

Author contributions

E. Brattico and M. Overgaard developed the study concept. M. Klarlund, E. Brattico and Y. Du designed the study. M. Pearce provided the stimuli. M. Klarlund and Y.Y. Wu performed testing and data collection. M. Klarlund performed data analysis and interpretation under the supervision of E. Brattico and Y. Du. M. Klarlund drafted the manuscript, and E. Brattico, Y. Du, Y.Y. Wu, M. Pearce, P. Vuust and M. Overgaard provided critical revisions. All authors approved the final version of the manuscript for submission.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105405.

References

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*(2), 111–130. https://doi.org/10.1016/S0022-0965(02)00124-8

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Ayari, M., & McAdams, S. (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

Balkwill, L.-L., & Thompson, W. F. (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

Blair, I. V., Dasgupta, N., & Glaser, J. (2015). Implicit attitudes. In *, vol. 1. APA handbook of personality and social psychology, Vol. 1. Attitudes and social cognition* (pp. 665–691). https://doi.org/10.4135/9781483386874.n236

Burgoyne, A. P., Harris, L. J., & Hambrick, D. Z. (2019). Predicting piano skill acquisition in beginners: The role of general intelligence, music aptitude, and mindset. *Intelligence, 76*, Article 101383. https://doi.org/10.1016/j.intell.2019.101383

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. https://doi.org/10.1037/0096-3445.113.3.394

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Demorest, S. M., & Morrison, S. J. (2015). *Quantifying culture: The cultural distance hypothesis of melodic expectancy* (p. 13).

Demorest, S. M., Morrison, S. J., Jungbluth, D., & Beken, M. N. (2008). Lost in translation: An enculturation effect in music memory performance. *Music Perception, 25*(3), 213–223. https://doi.org/10.1525/mp.2008.25.3.213

Demorest, S. M., Morrison, S. J., Nguyen, V. Q., & Bodnar, E. N. (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

Detterman, D. K., & Ruthsatz, J. (1999). Toward a more comprehensive theory of exceptional abilities. *Journal for the Education of the Gifted, 22*(2), 148–158. https://doi.org/10.1177/016235329902200204

Du, K., Hunter, J. A., Scarf, D., & Ruffman, T. (2021). Chinese children's in-group favoritism is affected by age and gender. *Journal of Applied Developmental Psychology, 72*, Article 101232. https://doi.org/10.1016/j.appdev.2020.101232

Du, Y., & Zatorre, R. J. (2017). Musical training sharpens and bonds ears and tongue to hear speech better. *Proceedings of the National Academy of Sciences, 114*(51), 13579–13584. https://doi.org/10.1073/pnas.1712223114

Eerola, T., Louhivuori, J., & Lebaka, E. (2009). Expectancy in Sami Yoiks revisited: The role of data-driven and schema-driven knowledge in the formation of melodic expectations. *Musicae Scientiae, 13*(2), 231–272. https://doi.org/10.1177/102986490901300203

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013–1027. https://doi.org/10.1037/0022-3514.69.6.1013

Foster Vander Elst, O., Vuust, P., & Kringelbach, M. L. (2021). Sweet anticipation and positive emotions in music, groove, and dance. *Current Opinion in Behavioral Sciences, 39*, 79–84. https://doi.org/10.1016/j.cobeha.2021.02.016

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., … Koelsch, S. (2009). Universal Recognition of Three Basic Emotions in Music. *Current Biology, 19* (7), 573–576. https://doi.org/10.1016/j.cub.2009.02.058

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Greenwald, A. G., & Lai, C. K. (2019). *Implicit social cognition* (p. 29).

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*(11), 466–472. https://doi.org/10.1016/j.tics.2007.08.008

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48–55. https://doi.org/10.1111/j.0956-7976.2005.00779.x

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, 102*(35), 12639–12643. https://doi.org/10.1073/pnas.0504254102

Hansen, N. C., Kragness, H. E., Vuust, P., Trainor, L., & Pearce, M. T. (2021). Predictive uncertainty underlies auditory boundary perception. *Psychological Science, 32*(9), 1416–1425. https://doi.org/10.1177/0956797621997349

Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01052

Hansen, N. C., Vuust, P., & Pearce, M. (2016). 'If you have to ask, You'll never know': Effects of specialised stylistic expertise on predictive processing of music. *PLoS One, 11*(10), Article e0163584. https://doi.org/10.1371/journal.pone.0163584

Hashkes, S., Van Rooij, I., & Kwisthout, J. (2017). Perception is in the details: A predictive coding account of the psychedelic phenomenon. In *Proceedings of the 39th annual meeting of the cognitive science society (CogSci 2017)* (pp. 2907–2912).

Haumann, N. T., Vuust, P., Bertelsen, F., & Garza-Villarreal, E. A. (2018). Influence of musical enculturation on brain responses to metric deviants. *Frontiers in Neuroscience, 12*, 218. https://doi.org/10.3389/fnins.2018.00218

Heilbron, M., & Chait, M. (2018). Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience, 389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Ho, W.-C. (2010). Moral education in China's music education: Development and challenges. *International Journal of Music Education, 28*(1), 71–87. https://doi.org/10.1177/0255761409351351

Holton, R. (2000). Globalization's Cultural Consequences. *The Annals of the American Academy of Political and Social Science*, 570–583.

Huron, D. (2008). Lost in music. *Nature, 453*, 456–457.

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

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *The Journal of the Acoustical Society of America, 124* (4), 2263–2271. https://doi.org/10.1121/1.2973189

Jacoby, N., Polak, R., Grahn, J., Cameron, D. J., Lee, K. M., Godoy, R., … Mcdermott, J. (2021). Universality and cross-cultural variation in mental representations of music revealed by global comparison of rhythm priors [preprint]. *PsyArXiv*.. https://doi.org/10.31234/osf.io/b879v

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences, 23*(1), 63–77. https://doi.org/10.1016/j.tics.2018.10.006

Krumhansl, C. (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

Land, M. F. (2014). Do we have an internal model of the outside world? *Philosophical Transactions of the Royal Society, B: Biological Sciences, 369*(1636), 20130045. https://doi.org/10.1098/rstb.2013.0045

Law, W.-W., & Ho, W.-C. (2011). Music education in China: In search of social harmony and Chinese nationalism. *British Journal of Music Education, 28*(3), 371–388. https://doi.org/10.1017/S0265051711000258

Longfellow, H. W. (1835). *Outre-mer: A pilgrimage beyond the sea*. Harper.

Matsunaga, R., Hartono, P., & Abe, J. (2015). The acquisition process of musical tonal schema: Implications from connectionist modeling. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01348

McCrae, R. R., & Costa, P. T. (1997). Conceptions and correlates of openness to experience. In *Handbook of personality psychology* (pp. 825–847). Elsevier. https://doi.org/10.1016/B978-012134645-4/50032-9.

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences, 36*(3), 587–596. https://doi.org/10.1016/S0191-8869(03)00118-1

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology, 20*(11), 1035–1041. https://doi.org/10.1016/j.cub.2010.04.019

Meade, A. W. (2009). FreeIAT: An open-source program to administer the implicit association test. *Applied Psychological Measurement, 33*(8), 643. https://doi.org/10.1177/0146621608327803

Morrison, S. J., Demorest, S. M., & Pearce, M. T. (2019). Cultural distance: A computational approach to exploring cultural influences on music cognition. In M. H. Thaut, & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 41–65). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198804123.013.3.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One, 9*(2), Article e89642. https://doi.org/10.1371/journal.pone.0089642

Nan, Y., Knösche, T. R., & Friederici, A. D. (2006). The perception of musical phrase structure: A cross-cultural ERP study. *Brain Research, 1094*(1), 179–191. https://doi.org/10.1016/j.brainres.2006.03.115

Nan, Y., Knösche, T. R., & Friederici, A. D. (2009). Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biological Psychology, 82*(1), 70–81. https://doi.org/10.1016/j.biopsycho.2009.06.002

Nan, Y., Knösche, T. R., Zysset, S., & Friederici, A. D. (2008). Cross-cultural music phrase processing: An fMRI study. *Human Brain Mapping, 29*(3), 312–328. https://doi.org/10.1002/hbm.20390

Nettl, B. (2005). *The study of ethnomusicology: Thirty-one issues and concepts* (New ed.). University of Illinois Press.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The implicit association test at age 7: A methodological and conceptual review. In *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). Psychology Press.

Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation: Enculturation: Statistical learning and prediction. *Annals of the New York Academy of Sciences, 1423*(1), 378–395. https://doi.org/10.1111/nyas.13654

Raman, R., & Dowling, W. J. (2016). Real-time probing of modulations in South Indian classical (Carnātic) music by Indian and Western musicians. *Music Perception, 33*(3), 367–393. https://doi.org/10.1525/mp.2016.33.3.367

Raman, R., & Dowling, W. J. (2017). Perception of modulations in South Indian classical (Carnātic) music by student and teacher musicians. A cross-cultural study. *Music Perception, 34*(4), 424–437. https://doi.org/10.1525/mp.2017.34.4.424

Ruthsatz, J., Detterman, D., Griscom, W. S., & Cirullo, B. A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence, 36*(4), 330–338. https://doi.org/10.1016/j.intell.2007.08.003

Rutland, A. (1999). The development of national prejudice, in-group favouritism and self-stereotypes in British children. *British Journal of Social Psychology, 38*(1), 55–70. https://doi.org/10.1348/014466699164031

Savage, S. J., & Waldman, D. M. (2008). Learning and fatigue during choice experiments: A comparison of online and mail survey modes. *Journal of Applied Econometrics, 23* (3), 351–371. https://doi.org/10.1002/jae.984

Schaal, N. K., Bauer, A.-K. R., & Müllensiefen, D. (2014). Der Gold-MSI: Replikation und Validierung eines Fragebogeninstrumentes zur Messung Musikalischer Erfahrenheit anhand einer deutschen Stichprobe. *Musicae Scientiae, 18*(4), 423–447. https://doi.org/10.1177/1029864914541851

Schaffrath, H. (1995). The Essen folksong collection. In *, 6. Database containing*.

Seashore, C. E. (1915). The measurement of musical talent. *The Musical Quarterly, I*(1), 129–148. https://doi.org/10.1093/mq/I.1.129

Shipp, S. (2016). Neural elements for predictive coding. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01792

Sparrow, P. R. (2009). Handbook of international human resource management: Integrating people, process, and context. *Human Resource Management, 544*.

Stevens, C. J. (2012). Music perception and cognition: A review of recent cross-cultural research. *Topics in Cognitive Science, 4*(4), 653–667. https://doi.org/10.1111/j.1756-8765.2012.01215.x

Stewart, L. (2007). Musical thrills and chills. *Trends in Cognitive Sciences, 11*(1), 5–6. https://doi.org/10.1016/j.tics.2006.09.013

Swaminathan, S., Kragness, H. E., & Schellenberg, E. G. (2021). The musical ear test: Norms and correlates from a large sample of Canadian undergraduates. *Behavior Research Methods, 53*(5), 2007–2024. https://doi.org/10.3758/s13428-020-01528-8

Swaminathan, S., & Schellenberg, E. G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Scientific Reports, 8*(1), 9223. https://doi.org/10.1038/s41598-018-27571-2

Thomas, K. S., Silvia, P. J., Nusbaum, E. C., Beaty, R. E., & Hodges, D. A. (2016). Openness to experience and auditory discrimination ability in music: An investment approach. *Psychology of Music, 44*(4), 792–801. https://doi.org/10.1177/0305735615592013

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Vuoskoski, J. K., Clarke, E. F., & DeNora, T. (2017). Music listening evokes implicit affiliation. *Psychology of Music, 45*(4), 584–599. https://doi.org/10.1177/0305735616680289

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music – Brain responses to rhythmic incongruity. *Cortex, 45*(1), 80–92. https://doi.org/10.1016/j.cortex.2008.05.014

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01111

Vuvan, D. T., & Hughes, B. (2019). Musical style affects the strength of harmonic expectancy. *Music & Science, 2*. https://doi.org/10.1177/2059204318816066, 2059204318816066.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The musical ear test, a new reliable test for measuring musical competence. *Learning and Individual Differences, 20*(3), 188–196. https://doi.org/10.1016/j.lindif.2010.02.004

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences, 1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Yoon, E., Langrehr, K., & Ong, L. Z. (2011). Content analysis of acculturation research in counseling and counseling psychology: A 22-year review. *Journal of Counseling Psychology, 58*(1), 83–96. https://doi.org/10.1037/a0021128

Zhang, L., Xie, S., Li, Y., Shu, H., & Zhang, Y. (2020). Perception of musical melody and rhythm as influenced by native language experience. *The Journal of the Acoustical Society of America, 147*(5), EL385–EL390. https://doi.org/10.1121/10.0001179