The effect of enculturation and implicit attitudes on musical categorization

Pætur Zachariasson
Student number: 20155938
Supervisor: Laura Petrini
Characters: 121.443

Aalborg Universitet
Psychology, 10th semester
May 29th, 2020

Abstract

This study examines whether human listeners perform similarly in a musical categorization task as the computational model IDyOM, which is designed to analyze statistical regularities of musical styles. The study also investigates whether participants' implicit cultural attitudes, as measured by the implicit association test, can predict performance in the categorization task. In addition, the study investigates whether there is a relationship between implicit attitudes and the categorization of ambiguous out-of-culture melodies. Ten Danish participants were presented with traditional Chinese and German melodies and asked to categorize them according to the perceived cultural origin. The melodies have been analyzed by the computational model IDyOM and found to be stylistically ambiguous. The participants also completed an implicit association test that measures implicit attitudes toward Chinese and Danish culture. The results found no statistically significant difference between the participants' and the computational model's categorization accuracy. However, a statistically significant difference was found in the categorization accuracy of distinct and ambiguous melodies. These findings demonstrate that the mechanisms underlying the IDyOM model can plausibly predict the effect that musical enculturation has on human listeners' ability to categorize melodies according to the cultural origin. The results found no statistically significant difference in categorization accuracy between participants with higher and lower implicit cultural bias. No relationship was detected between implicit attitudes and categorization accuracy of ambiguous out-of-culture melodies.

Keywords

categorization, statistical learning, music enculturation, cultural distance, IDyOM, implicit association test, perceptual assimilation model

Table of Contents

Appendix

1. Test for normality
2. Participant data and analyses output
3. Power analyses
4. Flyer

Introduction

When it comes to communicating cultural information, there are arguably not many societal practices as widespread and versatile as music (Volgsten, 2014). Music is commonly used to convey political campaigns, distribute consumer products, proclaim religious doctrines, and countless other cultural operations. Music is also central to many cultural identities (Volgsten, 2014). Musical styles are often named after the geographical area from which they originate (e.g., K-pop, Cajun music) or the cultural group associated with the style (e.g., Celtic folk, Latin music). Exposure to different musical environments is believed to have implications on the psychological processing of music (Haumann, Vuust, Bertelsen, & Garza-Villarreal, 2018). However, in a world that is continually globalizing and where all conceivable musical genres are just one click away on internet sites like YouTube and applications like Spotify and iTunes, it has been argued by leading researchers in the field that this can have a homogenizing effect on the musical mind (Huron, 2008).

Therefore, the neuropsychological study of music can be considered an essential and relatively urgent endeavor. In this study, some of the mechanisms and effects of the so-called enculturation process of music will be examined.

Due to social restrictions initiated by the Danish Government in response to Covid-19, all experiments planned for this study were postponed indefinitely. The initial plan was to conduct experiments on around 100 participants in Aarhus with a more elaborated paradigm than the one used in this study. In addition to the musical categorization task and the implicit association test, the planned paradigm would include the musicality test Musical Ear Test (MET; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010) and the questionnaire Goldsmith's MSI (Müllensiefen, Gingras, Musil, & Stewart, 2014). After discussing the situation with the supervisor, it was decided to use data collected in a pilot study some months earlier in order to maintain the timeline for assignment hand-in, set by Aalborg University.

It should be noted that plans to conduct the extended experiments have not been dropped but are expected to take place in the future when it is deemed more appropriate. The paradigm for the extended experiments was initially outlined in collaboration with the Center for Music in the Brain in Aarhus, and a parallel study with Chinese participants is planned to take place in China later this year, administered by a fellow student.

Disposition

Here is a general summary of the paper's main sections.

In the section *Principals of predictive processing*, the theoretical foundation for this study will be established. Predictive coding theories are arguably the most common approach to neuropsychological research in music at present, and here some of the central principles of these theories will be examined, along with the neurological underpinnings that substantiate this theoretical approach.

The section *Musical perception* will look closer at how the cognitive system processes music and how this processing generates perceptual expectations, which subsequently induce different psychological states.

In the section *Enculturation*, aspects that pertain to the acquisition of a primary musical culture will be described, along with thoughts on how such cognitive templates can be understood as perceptual schemas.

In *Music and language*, the focus will be on syntactic similarities between the two systems, music and language, in order to provide a foundation for the fourth hypothesis in the study. A linguistic model that makes predictions about perception of nonnative sound categories will also be presented in this section.

Next, the computational model that was used to analyze the melodies in the experiment will be introduced and explained in the section *IDyOM*.

The theoretic part of this assignment will be concluded with the section *Implicit attitudes* that will introduce a theoretical model through which the differentiation of implicit and explicit attitudes can be conceptualized. The implications that this model has for the interpretation of the implicit association test scores will then be discussed.

The project's four hypotheses will be presented in the section *The present study*, and a detailed description of the experiment's procedures will be in the section *Method.* The section *Results* will present the results from the statistical analyses used to investigate the hypotheses.

The discussion of the study's results will be done in two main sections, in which the first section, *Implications of results,* deals with the hypotheses separately, and the second section, *General discussion,* will look more generally at the results as well as consider other perspectives.

Finally, a *Conclusion* will summarize the study's findings.

Principals of predictive processing

In the following sections, the central theoretical principles of musical cognition will be examined. First, predictive coding will be introduced as a theoretical framework through which the cognitive processing of sensory information is understood. Statistical learning and probabilistic prediction will then be proposed as the cognitive mechanisms underlying this process by detailing experimental evidence from music-related research. The neural substrates of musical cognition and some relevant neural components will also be explored.

Predictive coding

Human perception appears to be graduated to the statistical properties of the sensory environment (Knill & Pouget, 2004). This has been demonstrated by research in various modalities, suggesting that attributes of both early visual processing and early auditory processing are most comprehensively explained by statistical learning algorithms, which will provide highly efficient representations at later stages of processing (Olshausen & Field, 1996; E. C. Smith & Lewicki, 2006). Such findings are consistent with the theoretical framework often referred to as *predictive coding*, which in short can be summarized as the proposition that the brain relies on generative "top-down" models formed by prior experience when processing and interpreting incoming "bottom-up" sensory information (Friston, 2010).

According to theories of predictive coding, top-down expectations generated by the cognitive system are confirmed or violated as bottom-up evidence from the sensory environment is being processed. The electrophysiological response *prediction error* is elicited when there is an incongruence between the top-down expectations and the bottom-up evidence (Friston, 2010). Therefore, prediction error functions as an adaptive neurological feedback mechanism that subsequently will update and revise the top-down perceptual model in order to promote cognitive parsimony (Badcock, Friston, Ramstead, Ploeger, & Hohwy, 2019). It is thought of as a hierarchical system where bottom-up sensory information advances upwards through feed-forward processes, while top-down layers in the hierarchy attempt to predict the substance of the information through the use of feedback processes, such as prediction error. In the case of successful predictions, the upward propagation of information is canceled out, while in the case of failed predictions, there will be an upward propagation of prediction error (Badcock et al., 2019). The dynamic interplay between top-down expectations and bottom-up evidence will gradually and continuously adjust the cognitive model used in the given modality in order to minimize prediction error (Friston, 2005).

Predictive coding theories have also been used to describe and conceptualize the brain's processing of musical material (Narmour, 1991; Pearce & Wiggins, 2006). According to such theories, listeners generate expectations about forthcoming auditory events as a musical composition unfolds. These top-down expectations are based on listeners' prior musical exposure and accumulated knowledge of a particular musical style (Pearce & Wiggins, 2006), and the expectations influence the brain's processing of musical material (Haumann et al., 2018).

Statistical learning and probabilistic prediction

Statistical learning is the implicit capacity of the brain to extract statistical regularities from the sensory environment (Perruchet & Pacton, 2006). It is considered a fundamental learning mechanism in various domains of cognitive neuroscience, including visual processing (Summerfield & Egner, 2009), motor-sequencing (D. M. Wolpert & Flanagan, 2001), language (DeLong, Urbach, & Kutas, 2005), and both nonmusical (Furl et al., 2011) as well as musical processing (Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). Based on cognitive models established through statistical learning, the brain generates likelihood estimations of sensory events via the process of *probabilistic prediction*, which enables the organization of the sensory environment encountered (Barsalou, 2009). In terms of music processing, it is hypothesized that people implicitly and automatically acquire knowledge of the statistical structure of a given musical environment through statistical learning and are enabled to perceptually organize and process this knowledge through probabilistic prediction (Pearce, 2018).

Statistical learning can account for much of the effects that exposure to different musical environments has on musical perception, including differences between listeners of different cultural styles (Demorest, Morrison, Nguyen, & Bodnar, 2016) and age (Hannon & Trehub, 2005a). For example, one study explored perceptual differences in terms of expectancy patterns between subjects of different cultural backgrounds (Carlsen, 1981). A vocal continuation-paradigm was applied to three groups of participating musicians: American, German, and Hungarian. The participants were presented with melodic phrases that were interrupted abruptly. They were then asked to sing the notes that they expected to come next in the melodic phrase. Significant differences were found between the groups when analyzing the different groups' expected continuations, suggesting that cultural background influences listeners' perceptual expectations of musical events (Carlsen, 1981).

The cognitive processing of musical material also involves several aspects of the memory system, for example, episodic memory, semantic memory, and working memory (Platel, Baron, Desgranges, Bernard, & Eustache, 2003; Schulze & Koelsch, 2012). In a study where subjects were introduced to an unconventional musical system and asked to make force choice decisions concerning the recognition of tonal sequences, it was demonstrated that the participants showed superior recognition memory for previously heard sequences, regardless whether the sequences appeared in melodies that occurred more than once in the experiment or in newly introduced exemplars (Loui & Wessel, 2008). This finding indicates that listeners internalize the syntactical structure of a musical style after exposure to a sufficient set of exemplars. It also demonstrates that retrieval is more successful in case of more exposure, presumably because the internalized structures are more readily represented cognitively (Loui & Wessel, 2007). In other words, exposure to musical sequences results in the internalization of statistical tonal structures that comprise a musical style, and this, in turn, facilitates retrieval.

Following this vein, it can be assumed that the process of encoding musical material is related to the perceived predictability of the sequence in question. From an information-theoretic perspective, a melodic sequence that is perceived to be predictable does not need to be encoded in full since an appropriate predictive system will be able to reconstruct it successfully without much strain (MacKay, 2004, p. 74). Neural responses that are associated with predictive processes have been demonstrated to correspond to the probability of different musical continuations (Omigie, Pearce, Williamson, & Stewart, 2013). For example, studies using neuroimaging techniques have demonstrated correlations between electrophysiological components that reflect surprisal and the information-theoretic measuring-unit *information content* (IC), thereby demonstrating a close relationship between such neural responses and computational probability (Omigie et al., 2013). IC represents the inversed probability of an event occurring in a particular context, meaning that the lower IC is prescribed to an event, for example a musical note in a melodic sequence, the more predictable it is considered (MacKay, 2004, p. 67). This means that a predictable melodic sequence has low IC and is relatively compressible, while an unpredictable sequence has high IC and requires more storage memory in the encoding process (MacKay, 2004, p. 74).

Neural substrates of musical cognition

When a sound is heard, neural signals representing the soundwaves are transmitted from the ear to the brainstem, and from the brainstem through subcortical pathways to the thalamus, and then from the thalamus to the auditory cortex, where the primary auditory processing takes place (Large, 2017, p. 4). The auditory cortex is part of the temporal lobe and consists of three main sections: the core area, the belt region, and the parabelt region (Pickles, 2012, pp. 211-214). The auditory cortex is hierarchically organized in such a way that the sensory core area, constituting the lower layers of this organization, has a simpler receptive field that will project information to the higher layers like the belt, the parabelt, and the prefrontal cortex, which have more complex response properties (Rauschecker & Scott, 2009).

Two separate neural pathways originate in the core area. One travels dorsally and posteriorly and is directed towards parietal areas, while the other travels ventrally and anteriorly within the temporal lobe (Medalla & Barbas, 2014). Both pathways eventually target separate areas in the frontal cortex and are considered bidirectional, thereby creating functional feed-forward and feedback loops between the higher and lower layers in the hierarchical organization (Zatorrea & Salimpoor, 2013). These loops are believed to enable interaction between the auditory system and the memory system, as well as the auditory system and the motor system regarding planning and organizing action response (Kumar et al., 2016). Such interactions are hypothesized to lay the basis for predictive ability in music cognition by generating expectations derived from syntactic regularities of prior musical experience (Zatorrea & Salimpoor, 2013). In other words, exposure to the sequential regularities that characterize a particular musical environment in terms of pitch and rhythm shapes the cognitive templates necessary to perceptually differentiating between musical styles.

Neural components of musical processing

Changes in brain activity during a music-listening session can be observed with high spatial and temporal precision with the use of functional neuroimaging techniques like *electroencephalography* (EEG) and *magnetoencephalography* (MEG) (Baillet, 2011). Such techniques are capable of mapping out neural responses to stimuli called *event-related potentials* (ERP).

One particular ERP that prompts interest regarding auditory processing is the *mismatch negativity* (MMN). It is an ERP with an increased amplitude that peaks around 100 to 250 milliseconds after a deviant auditory event in the auditory cortex and the inferior frontal cortex (Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN can occur in response to both linguistic deviations and musical irregularities, such as unexpected changes in repetitive sequences, pitch frequency, duration, volume, and timbre, as well as in the absence of an expected sound (Näätänen et al., 2007). Attention influences the magnitude of the MMN, but generally, it is considered to be preattentive, meaning it occurs even if the subject is not paying attention to the deviant sound (Brattico, Tervaniemi, Näätänen, & Peretz, 2006).

Another ERP that is relevant for music processing is the *early right anterior negativity* (ERAN), which is elicited in response to rule-violations based on the listener's long-term musical knowledge, rather than the short-term violations that elicit the MMN (Omigie et al., 2013). ERAN has an increased amplitude that peaks around 180 to 400 milliseconds after a deviant sound (Koelsch, Gunter, Wittfoth, & Sammler, 2005). Experiments have associated the ERAN with violations of harmonic rules, and violations of the expected melodic context (Koelsch, Jentschke, Sammler, & Mietchen, 2007; Miranda & Ullman, 2007).

The MMN and the ERAN are considered to be closely related since low-probability deviant sounds elicit both, and they have been suggested to be neural signatures of the same mechanism, namely statistical learning (Omigie et al., 2013).

The ERPs discussed here are relevant for this study since these are the neural components hypothesized to underlie the cognition and perceptual processing of music, especially in the case of hearing deviant sounds and sound sequences, as is the case when listening to ambiguous and out-of-culture music.

Musical perception

In the sections below, aspects of musical perception will be explored by focusing on the earlier stages of auditory processing as well as the organizing principles associated with each stage. A simplified version of the Neurocognitive Model of Music Perception (Koelsch & Siebel, 2005) will be depicted. Following this, the focus will be set on how expectation facilitates emotional and other psychological responses to music.

Stages of processing

The perception of music starts with the extraction of auditory features of the acoustic information, which is being perceived. Sound frequencies are mainly processed in the auditory brainstem, thalamus, and auditory cortex (Large, 2017, p. 4). The initial analysis provides information about features like location, timbre, periodicity, as well as pitch height, pitch chroma, and loudness (Koelsch & Siebel, 2005). These processes are reflected in neural activity patterns like the auditory brainstem response (Picton, Durieux-Smith, & Moran, 1994), the frequency following response (J. C. Smith, Marsh, & Brown, 1975), and mid-latency components like P1 and N1 (Näätänen & Picton, 1987).

Following this early analysis, the auditory information enters the echoic sensory memory and is organized by principles similar to those of Gestalt theory (Moelants & Leman, 1997, p. 61). This produces an auditory "scene" consisting of the interactions between various mental representations of different acoustic elements (Bregman & McAdams, 1994). Some of the most crucial organizational principles are *similarity*, referring to elements that form coherent structures based on the fulfillment of particular conditions; *proximity*, referring to a clustering effect of elements in close proximity; and *grouping*, referring to segregation of groups of elements (Bregman & McAdams, 1994).

Organizational principles at this stage of processing have been convincingly described in a connectionist frame of theoretical reference (Bharucha, 1987). Take, for example, the principle of perceptual grouping, which is an algorithmic strategy of identifying temporal boundaries in a sequence of stimuli. It is also considered a fundamental part of perceptual processing in other modalities, such as language comprehension (Brent, 1999), and memory (Kurby & Zacks, 2008). From a connectionist perspective, temporal sequencing is based on the effects of task processing rather than explicit spatial representations (Elman, 1990). In other words, the identification of a segment's boundary will occur when an event in the sequence is deemed unpredictable in the given context since the process of generating expectations is disrupted and fails to compute a probable continuation (Pearce, 2018). A grouping effect will, therefore, occur at this point, resulting in the perceptual segmentation of musical phrases.

Since unpredictable events generally demand more processing capacity, these organizational principles can be neurophysiologically measured (Pearce, Müllensiefen, et al., 2010). Processes at this stage are partly reflected by the neural signature *mismatch negativity* (MMN) and are believed to be mediated primarily by the auditory cortex, the superior temporal gyrus, and the planum temporale (Näätänen et al., 2007). As mentioned above, the MMN is a neural component that is elicited by a change in stimulus at approximately 100-250 milliseconds, primarily in the auditory cortex and the inferior frontal cortex. It is believed to reflect the neural detection of a change in a constant auditory environment (Näätänen et al., 2007).

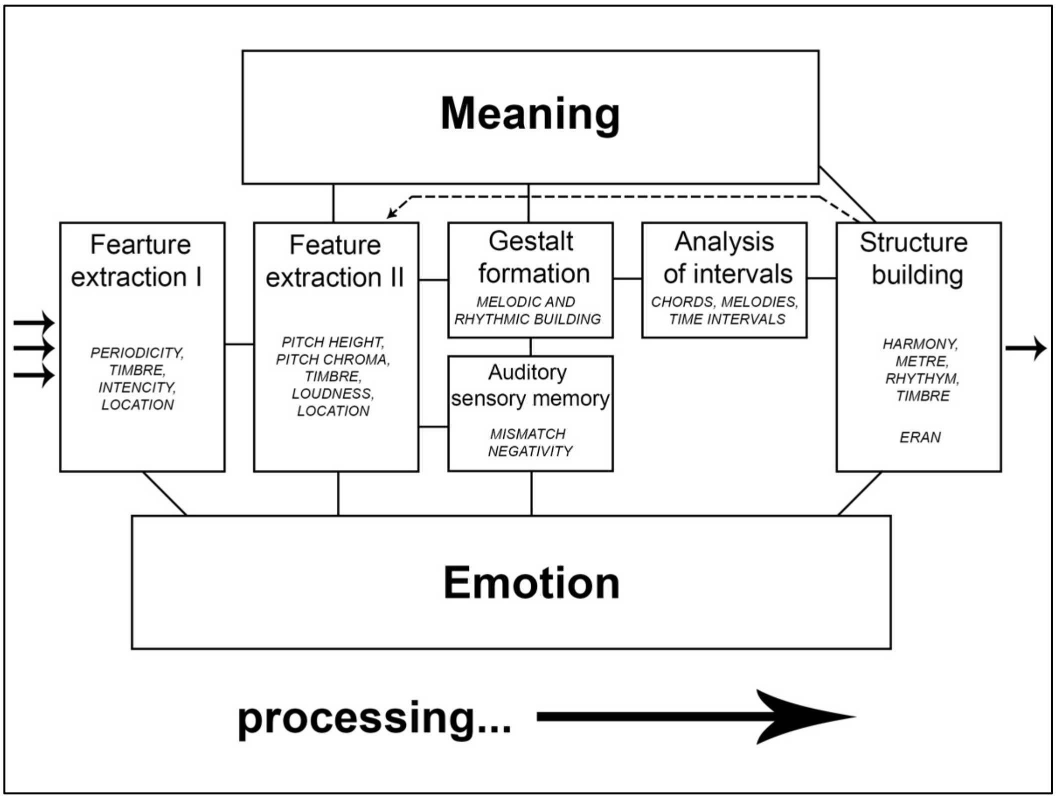


Figure 1. A simplified version of the Neurocognitive Model of Music Perception (Koelsch & Siebel, 2005). Processing occurs in stages from left to right. Meaning and emotion are continuously extracted and reevaluated. There are additional stages in the full model. (ERAN = Early right anterior negativity).

The next stage in the perceptual processing of music involves detailed processing of pitch interval and timing (Koelsch, 2012). As described in a section above, expectations of pitch intervals are based on learned statistical regularities that, in turn, affect how effectively auditory events are processed (Pearce, 2018). The timing of notes is also predicted according to learned statistical regularities, but, importantly, temporal predictions are also considerably influenced by *meter*, a hierarchical framework of periodically reoccurring and differently weighted beats that are aligned with a musical composition (Longuet-Higgins & Lee, 1984). Meter perception is also subject to cross-cultural differences, as has been demonstrated in experimental studies (Hannon & Trehub, 2005b, 2005a). In these studies, subjects consisting of North American and Bulgarian adults, as well as a group of 12-month-old infants, showed culture-specific response patterns to rhythms, indicating effects of learned regularities – also termed musical enculturation. These processes are assumed to be based on mechanisms of statistical learning and probabilistic prediction (Hannon & Trehub, 2005a, 2005b).

The next stage in the perceptual processing of music is where the elements from previous stages are combined to form a more coherent musical structure (Koelsch & Friederici, 2003). The neural component ERAN is believed to reflect expectational processes that occur at this stage (Koelsch et al., 2005).

Musical expectation

A fundamental feature of the human brain is the ability to anticipate the outcome of ensuing events since more accurate anticipations allow for faster and more appropriate adaptation to changes in the environment (Badcock et al., 2019). From an evolutionary perspective, accurate expectations and appropriate responses have important implications for an organism's survival. Therefore, expectations are assumed to be grounded in learning and memory processes since these will facilitate survival (Mitchell et al., 2009). Expectations are also closely connected to neural circuits related to the processing of emotion and reward (Zatorre, 2015). In psychological terms, action-readiness prepares individuals for some possible outcomes and not for others. A particular expectational set has implications for various psychological and cognitive systems, such as attention, emotional processing, learning, and motor sequencing (Brattico, Brigitte Bogert, & Jacobsen, 2013). For example, research has demonstrated that priming of pitch frequency and temporality allows participants to direct their attention toward an anticipated auditory event, which in turn lowers the threshold for conscious detection of the expected sound, demonstrating that higher-level cognitive processes like expectation influence detection of low-level sensory processes (G. Z. Greenberg & Larkin, 1968).

The generation of musical expectation is believed to be tuned to the particular musical environment to which a listener has been exposed (Huron, 2006, p. 204). This means that different sets of expectations pertain to different musical environments. Such musical expectancy has been demonstrated as one of the main psychological mechanisms by which emotional response is induced by music listening (Juslin & Västfjäll, 2008). In one study investigating the effects that violations of harmonic expectancy can have on emotional response, subjects were presented with several classical compositions, of which some had been manipulated to violate principals of Western music theory (Steinbeis, Koelsch, & Sloboda, 2006). Results showed correlations between increased musical unexpectedness and measures of tension, emotionality, and electrodermal activity (Steinbeis et al., 2006). Some of the physical, psychological, and neural effects that violations of musical expectancy can evoke were thereby demonstrated.

Theoretical propositions have suggested an inverted U-shaped relationship between expectancy in music and facets of both motoric and emotional response (Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014). This means that the strongest motoric and emotional response to musical events such as syncopations in rhythms are evoked when the event is deemed neither too predictable or too unpredictable (Vuust et al., 2018). Musical events that deviate too much from learned expectancy fail to elicit motoric and emotional responses, and the same is true for musical events that are perceived as too predictable. Such results substantiate the hypothesized role that statistical learning and probabilistic prediction play in linking listeners' emotional arousal with musical predictability (Pearce, 2018).

The perceptual processing of music eventually leads to generation of musical expectations that facilitate the comprehension of musical style (van der Weij, Pearce, & Honing, 2017). These processes are based on the cognitive and perceptual mechanisms that are believed to underlie the ability to categorize melodies according to style, as is required in the categorization task in this study (Pearce, 2018).

Enculturation

Some of the central aspects associated with the process of enculturation will be described in the following sections. First, similarities to the linguistic enculturation process will be looked at. Second, a look at how memory can shed light on the cognitive processing of music. The focus will then be set on how encultured cognitive models function as perceptual schemas, and how such schemas are generated, initiated, and how they affect the perception of musical styles.

A person acquires a primary musical style through the process of enculturation (van der Weij et al., 2017). This process occurs implicitly and informally through repeated social interaction with a prevailing musical culture, whereby statistical regularities of the musical environment are internalized as cognitive models (Bigand & Poulin-Charronnat, 2006; Saffran, Johnson, Aslin, & Newport, 1999). From infancy, most children experience musical engagement through songs, games, lullabies, jingles, and a multitude of other sources, and this is believed to equip even untrained listeners with a level of musical sophistication that enables them to respond to particular musical styles similarly as musically experienced listeners (Bigand & Poulin-Charronnat, 2006). Excluding instances of cognitive deficits that result in music-specific disorders, individuals will acquire knowledge of the organizational concepts representative of the particular musical style automatically (Ayotte, Peretz, & Hyde, 2002). A musical style in this sense transcends a mere musical genre (e.g., rock, jazz, or hip hop), but rather it refers to the tonal, harmonic, and rhythmic structure that underlies a broader musical tradition, for example, Western music (Bharucha & Krumhansl, 1983). This knowledge sometimes dates back centuries and has been formalized in theoretical terms (Bharucha & Krumhansl, 1983).

Similarities to language acquisition

Acquisition of musical knowledge occurs similarly to the acquisition of knowledge in other cultural practices, namely through psychological mechanisms of social learning that start in infancy and eventually establish cognitive categories that influence perception (Henrich & McElreath, 2003). Already around the age of 12 months, children start to display decreased sensitivity to violations of tonal organization and rhythmic patterns of foreign musical cultures (Hannon & Trehub, 2005b; Lynch & Eilers, 1992). This loss of sensitivity to out-of-culture auditory categories is consistent with observations in linguistic research, where infants appear to go through a similar process of decreased sensitivity for prosodic patterns of foreign languages; a process that, in turn, is believed to effectively increase sensitivity for the prosodic and phonetic patterns that constitute their native language (Kuhl, 2004).

There are also similarities between the acquisition of primary musical culture and acquisition of native language in terms of underlying cognitive mechanisms (Matsunaga, Yokosawa, & Abe, 2012; Wong, Roy, & Margulis, 2009). For example, a study by Saffran *et al.* (1999) demonstrated that the same cognitive processes are likely utilized in implicit pattern learning in both linguistic and non-linguistic material, namely statistical learning (Saffran et al., 1999).

The process of language acquisition has been described by Kuhl (2004) as a computational strategy where the detection of statistical regularities of prosodic input in time enables children to perceive larger patterns in continuous auditory stimuli (Kuhl, 2004). The process has been summarized in the *native language neural commitment* (NLNC) hypothesis, which states language learning in infants involves the production of dedicated neural networks that are sensitive to statistical regularities of prosodic cues; a process that eventually leads to word learning (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005). According to the NLNC, sensitivity to prosodic cues of the native language infers with the processing of foreign-language patterns (Kuhl, 2004), similar to what has been observed in perception of out-of-culture rhythmic patterns (Hannon & Trehub, 2005b). Therefore, a similar process to NLNC is likely involved when acquiring a native musical culture, based on the observation that the underlying cognitive machinery of statistical learning also plays a fundamental role in obtaining musical knowledge and generating perceptual expectations in music (Pearce & Wiggins, 2006).

Memory and musical processing

One approach to explore how enculturation influences musical processing is to use performance in memory tasks as a proxy for musical understanding. The underlying assumption is that if musical cognition is influenced by culture-specific content, then out-of-culture music should be more difficult to process (Demorest et al., 2016, 2009).

Research has shown that adolescents and young adults have poorer memory for out-of-culture-music, indicating more difficulty processing foreign music compared to native music (Wong et al., 2009). This difference in memory performance is not the case for bi-musical adults who have been encultured in two different musical traditions, echoing findings in linguistic research, where bilingual individuals have been shown to perform better on some memory tasks than their monolingual peers (Bialystok, 2009; Wang, Kuhl, Chen, & Dong, 2009).

In another musical-memory study, musically trained and untrained participants' ability to encode and retrieve in-culture and out-of-culture music were investigated (Demorest, Morrison, Beken, & Jungbluth, 2008). The results showed no difference in performance based on musical expertise but indicate that cultural style does appear to be a crucial variable in superior memory performance since all participants were better at remembering music from their own musical culture (Demorest et al., 2008).

These studies demonstrate how musical enculturation influences processing on a behavioral level, but differences have also been demonstrated on a neurological level. The general notion from a predictive coding perspective is that greater neural activation indicates greater difficulty in processing the information at hand (Friston, 2005). An fMRI study by Demorest *et al.* (2009) demonstrated that memory performance is better for in-culture music, and also showed greater brain activity during task performance in right frontal areas, right angular gyrus, and posterior precuneus when processing out-of-culture music in adult subjects from the US and Turkey with minimal musical training (Demorest et al., 2009). This indicates greater difficulty in processing out-of-culture musical material.

An earlier fMRI study by the same researchers along similar lines, but using Western and Chinese music instead, had shown no significant difference in activation when adult subjects listened to in-culture and out-of-culture music, although there were performance differences in memory tasks based on cultural style (Morrison, Demorest, Aylward, Cramer, & Maravilla, 2003). These findings indicate that simple listening tasks using music from different styles can activate similar neural patterns but may prompt different results in memory task performance. Interestingly, musically trained subjects did show additional activation in the superior angular gyrus, and left and right midfrontal regions for Chinese and Western music, respectively, indicating that musical expertise can influence neural response patterns when listening to in-culture and out-of-culture music (Morrison et al., 2003).

Efforts have been made to identify the specific musical feature that triggers an encultured memory response. For example, one study presented subjects with Western and Turkish compositions of three varying contextual complexity: one condition consisted of a full orchestral arrangement, one of melody alone played on a piano, and one of isochromatic pitch sequence of the melody (Demorest et al., 2016). In addition to reiterating findings mentioned above, namely that subjects perform better in memory tasks involving in-culture music compared to out-of-culture music, the results also showed that musical context in terms of arrangement and instrumentation does not appear to have a significant influence on memory performance. Neither does the subjective preference for compositions (Demorest et al., 2016). This can be taken as an indication that the encultured response to music is partly generated by the stylistic expectancies of the statistical pitch distribution that is characteristic of one's in-culture musical style.

This last point is particularly relevant for the present study, since the music used in our experiments consisted of melodic phrases that were devoid of harmony or instrumental arrangements.

Schema

A fundamental principle when obtaining appropriate knowledge in a given field or subject is to restrict the scope of inductive lessons to include only those that apply to the situation (Cosmides & Tooby, 2000). From an evolutionary perspective, context-dependent knowledge promotes evolution since all local contexts are assumed to have specific requirements (Tooby & Cosmides, 2015). Therefore, tools of information management that can shift appropriately and effectively between different contexts are assumed to be a necessity for adaptation and evolution (Tooby & Cosmides, 2015).

In cognitive psychology, the application of encapsulated perceptual or behavioral models to a particular situation has long been conceptualized through schema theory (Bartlett & Burt, 1933). A *schema* is defined as a set of mental representations that pertain to some perceived environmental regularities (Bregman & McAdams, 1994). In music, these mental representations are assumed to form sets of musical expectations that facilitate perception of the particular musical style (Krumhansl & Castellano, 1983). For example, listeners that have been enculturated in traditional Western music demonstrate considerably different expectations of tonal and harmonic structures depending on whether the music listened to is perceived to be in a minor key or a major key (Krumhansl & Kessler, 1982). Norms of harmonic progressions have also been linked to specific historical periods, meaning that the probability of certain chord-progressions changes over time (White, 2014). This suggests that music cognition and the formation of perceptual schemas are historically and culturally situated processes, based on the assumption that information processing is shaped by interactions with the particular environment in which it occurs (Byros, 2009).

The process of establishing a schema for a particular style of music involves learning to associate particular environmental markers with a generalization principle (Moelants &

Leman, 1997, p. 61). Both auditory and nonauditory markers can be associated with a musical style; for example, baroque wigs, dreadlocks, spiky colored mohawks are distinct hairstyles that are strongly associated with baroque music, reggae, and punk music, respectively (Waeber, 2007). However, extramusical cues are not of interest in this project.

One prominent musical marker for identifying a musical style is instrumentation (R. S. Wolpert, 1990). In one study, it was found that changing instrument timbre diminished recognition of compositions significantly among musicians, suggesting that timbre contributes to the process of identifying musical material (Poulin-Charronnat, 2004). Identification of musical style has also been shown to occur quite early in processing, as exemplified in a study where listeners were presented with short excerpts of commercial music in a variety of styles in order to assign genre-labels (Gjerdingen & Perrott, 2008). The results showed that listeners could make above-chance classifications of excerpts that were only 250 milliseconds long and that a near ceiling effect occurred after one second of exposure (Gjerdingen & Perrott, 2008). This is also an indication that instrumentation and timbre provide important cues for identifying musical styles, indeed maybe more so than pitch intervals associated with a style, which, by definition, must include several notes, and therefore take longer to register (R. S. Wolpert, 1990).

Musical schemas can represent both styles as well as other more specific musical elements, like major and minor modes (Krumhansl & Kessler, 1982). Many of these musical elements span over several genres and overlap different cultural styles (Bharucha & Krumhansl, 1983). Experiments suggest that musical schemas relating to structural principles in one's cultural tradition are accessed when processing music from unfamiliar genres that yet are deemed culturally familiar in a broader traditional sense (Demorest et al., 2009). These schemas are thought to facilitate better memory task performance, and also appear to correlate with the level of formal training, which corresponds with similar results among experts in other fields, for example, in chess (Gobet & Simon, 1996; Morrison et al., 2003). This is interpreted as demonstrating more sophisticated schema formation based on more advanced perceptual "chunking" mechanisms, which in turn facilitate more effective and efficient encoding and retrieval strategies (Gobet et al., 2001).

In the linguistic literature, *cohort theory* describes how both top-down and bottom-up processes contribute to word recognition (Marslen-Wilson & Tyler, 1980). According to the theory, interpretive processes facilitate accurate and effective word recognition by constraining the possibilities of words that are likely to be perceived even before any sound has been heard (Marslen-Wilson & Tyler, 1980). The point at which a word is recognized is, therefore, likely to occur earlier. A similar recognition-process may be underlying style-recognition in music. For example, there is evidence that listeners may have a "starting schema," based on musical propensities that characterize the primary style in which they are encultured. This has been demonstrated in Western-encultured listeners' tendency to assume that compositions are in the major mode (Huron, 2006, p. 207) and in a binary meter (Brochard, Abecasis, Potter, Ragot, & Drake, 2003), thereby reflecting a default "Western" schema.

Schema change

Violations of expectations in terms of pitch and harmony will signal that a more appropriate schema is needed to process the current music effectively (Moelants & Leman, 1997, p. 65). One example of such schema change in music is the concept of *key modulation*, defined as the establishment of a new tonal center in an ongoing composition, thereby introducing new chords but preserving the relative structure of chord progression. An experiment by Krumhansl and Kessler (1982) sought to investigate how quickly a new key is perceived to be established when a key modulation occurs in a song. Based on reports from subjects, the analysis suggested that a new tonal center is firmly established within three chords in the new key (Krumhansl & Kessler, 1982).

When schematic expectations are obviously and persistently violated, and therefore fail to be established in accordance with a generalization principle, it may initiate the formation of a new schema since no other schemas are adequately suitable for the musical events that are being processed (Moelants & Leman, 1997, p. 61). It has been hypothesized that inductive learning can be formed by a so-called *unusualness rule* (Holland, Holyoak, Nisbett, Thagard, & Smoliar, 2008, p. 89). This refers to situations where an unusual property acts as a primary clause for an inductively learned rule. According to this view, the unusual property may be the decisive factor that is used to distinguish two similar schemas. If a listener does not gain enough exposure to a novel musical style, it is expected that the person fails to attribute an unusual property to the style that can be used to distinguish it from other similar styles. In such cases, the listener will not create a new schema to represent it cognitively and may, therefore, either assume it is not a different style, or it may be attributed to an ambiguous category of "otherness" (Huron, 2006, p. 215).

The musical stimuli used in this study are deemed to be stylistically ambiguous by a computational model, and it is therefore assumed that listeners will have a relatively hard time noticing unusual properties that can be used as cues to shift between Danish and Chinese listening-schemas.

Music and language

Some of the most prominent similarities and dissimilarities between music and language will be examined in the following section. This examination will be done on an abstract level of analysis, thereby highlighting the organizational resemblance between the two systems. Two figures illustrating tree diagrams of a language sentence and a musical phrase will be shown.

In the second section, the perceptual assimilation model will be introduced. This model makes predictions about the perception of nonnative sound-categories, which will be used in the study's fourth hypothesis. In short, these predictions state that prototypical native sound-categories function as perceptual magnets for nonnative phonemes.

Syntactic similarities between music and language

Two notable communicative systems of the auditory modality, spoken language and music, share several fundamental syntactic features (Patel, 2003). Nevertheless, these two rich systems should not be thought of as mere variants of one another since many elimental differences make a direct comparison between the two systems' elements impossible (Patel, 2012, p. 263)

There are principles in language that govern how words form subunits, which are further combined to produce coherent sentences, as demonstrated in the syntactic tree diagram in Figure 2. Such tree structures illustrate the relationship between the constituents of a sentence that are intuitively comprehended by native speakers (Patel, 2012, p. 253). For

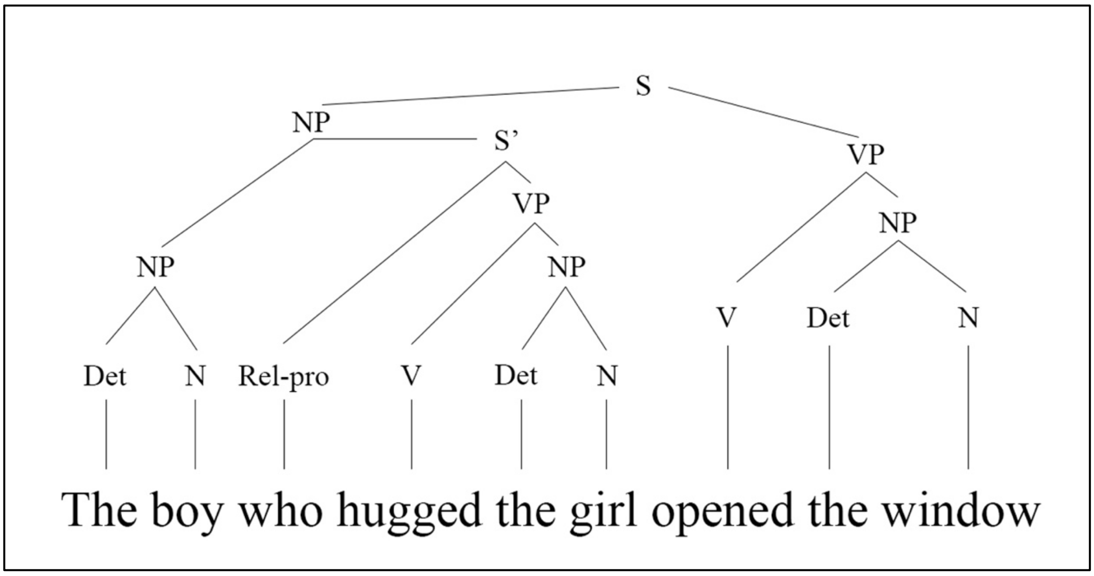


Figure 2. Syntactic tree diagram of an English sentence. S = sentence; NP = noun phrase; VP = verb phrase; S' = sentence modifier; V = verb, N = noun; Det = determiner; Rel-pro = relative pronoun (Patel, 2012, p. 253).

example, the long-distance dependency between the words "boy" and "window" in Figure 2 is apparent when following the peripheral branches that connect the first and last constituents in the sentence.

Music can be organized in similar hierarchical tree structures that also consist of subunits, which in combination form melodic phrases. However, long-distance dependencies between units are not intuitively perceived in music, as they are in a language phrase (Patel, 2003). Rather, branching patterns in music indicate the importance of events and reflect which events are relaxing or tensing (Lerdahl & Jackendoff, 1983). Musical key and rhythm signify where the most appropriate branch divisions. The example in Figure 3 shows four degrees of branches in a tree diagram. The higher the degree of branches, the more tensing is the musical event (Lerdahl & Jackendoff, 1983). In the example in Figure 3, the relaxing notes are the ones with the lowest branch-degrees, namely the 1st, 5th, 6th, 9th, and 13th. These notes also are the keynotes of the underlying chords.

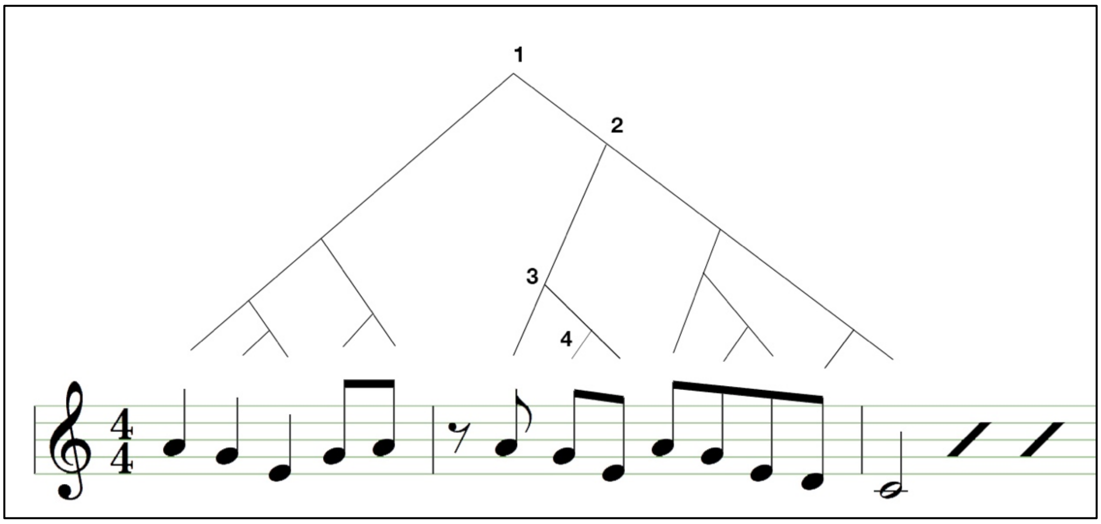


Figure 3. Syntactic tree diagram of an original musical phrase. When starting at the apex, the branch-degrees are defined by counting the number of joints it takes to reach the musical note at the bottom. For example, the first and last notes are on first-degree branches, the second note is on a third-degree branch, and the third note is on a second-degree branch (Lerdahl & Jackendoff, 1983).

In this way, the two systems are notably similar on an abstract level of analysis concerning their multilayered organization. In both systems, grammatical categories are easily replaced by other members of the same category, thereby changing the semantic information but preserving the syntactic structure (Patel, 2012, p. 265).

Nonhierarchical features of the logical syntactic structure are also comparable in both systems (Patel, 2003). For example, the concept of *verb agreement* in a sentence requires a logical agreement between a subject and its verb in a sentence (Jackendoff, 2002, p. 149). This means that words have grammatical functions that do not concern their intrinsic properties, but rather pertain to the structural role they play in the particular sentence. In music, such logical structures also exist in the sense that chords have harmonic functions derived from their relation to other chords, rather than from its intrinsic qualities (Patel, 2012, p. 266). The primary harmonic functions of chords in Western music theory are the *tonic*, the *subdominant*, and the *dominant*, referring to chords based on the first, fourth, and fifth note in a scale, respectively (Krumhansl & Kessler, 1982). This demonstrates that grammatical functions in both music and language are dependent on the logical structure of the context.

It is important not to equate grammatical categories in language, like verbs, nouns, and adjectives, with grammatical categories in music, like the tonic, the subdominant, and the dominant. Words have much more intrinsic semantic properties, while chords have more extrinsic properties (Patel, 2012, p. 260). The rich semantic properties in language are also assumed to minimize the tolerance for ambiguity, which is evident in how the cognitive system processes language compared to music (Patel, 2012, p. 264). For example, an ambiguous sentence will usually be interpreted only in one way, whereas there is much more tolerance for ambiguity in a musical phrase.

The Perceptual Assimilation Model

The Perceptual Assimilation Model (PAM; Best, 1995) describes how the perception of sound elements in a second language (L2) is dependent on the perceived similarity to phonetic elements in one's first language. Within the PAM framework, a differentiation is made between two groups of nonnative language listeners (Best & Tyler, 2007). *Naïve nonnative listeners* are defined as functional monolinguals, meaning they are not actively learning or using an L2 and are therefore naïve to the target language's stimuli. *L2-learning listeners*, on the other hand, are those who are actively learning an L2 for functional and communicative purposes and are therefore not naïve to the stimuli of the target language (Best & Tyler, 2007).

The participants in the present study, who are all Danish, are assumed to be the musical equivalent of the former group, namely functional mono-musicals that are naïve to the musical material of the out-of-culture Chinese style.

One of the PAM's underlying assumptions is that speech perception principles govern the learnability of an L2 (Best, 1995). According to the model, monolinguals have difficulty discriminating between phonetic categories of foreign languages that do not appear in their native language (Best & Tyler, 2007). These phonetic categories can include both consonants, vowels, and other tonal contrasts. The variations in perception of nonnative speech are assumed to reveal perceived similarities and dissimilarities between the phonetic structure of native stimuli and the nonnative stimuli (Tyler, Best, Faber, & Levitt, 2014). This echoes similar findings in L2 research, where discrimination of sound categories appears to be asymmetrical in such a way that discriminability is enhanced at phonetic boundaries (Kuhl, 1993). In other words, this means that discrimination is worse for good exemplars of a sound category, resulting in a "perceptual magnet"-effect around known phonetic categories, according to the closely related Native Language Magnet theory (Kuhl, 1993).

According to the PAM, four general predictions can be made about discrimination of nonnative sound categories by monolinguals. If two nonnative sounds are perceived as good exemplars of two distinct native sound categories, discrimination between the nonnative sounds will be easy (*two-category assimilation*). If two nonnative sounds are perceived to be equally good exemplars of a single native sound category, it will be difficult to discriminate (*single-category assimilation*). When two nonnative sounds are perceived as the same native sound category but differ in the goodness of fit, discrimination will be intermediate (*category-goodness assimilation*). Lastly, when one nonnative sound category is perceived to fit more than one phoneme, and the other nonnative sound category is perceived to fit one native phoneme, discrimination will be easy (*uncategorized-categorized assimilation)* (Tyler et al., 2014).

Most of the melodies in the present study are deemed to be ambiguous by the computational model IDyOM. This means that ambiguous out-of-culture melodies consist of sound elements that will be difficult to differentiate from ambiguous in-culture melodies. Therefore, the primary focus of this study will be on single-category assimilation.

These theoretical predictions do have experimental backing. In one study where the PAM framework was applied, groups were formed consisting of university students studying the English language and a control group consisting of naïve listeners (Grimaldi et al., 2014). Categorization and discrimination tests were administered, as well as a listening test where participants' neurological response was measured. Researchers, therefore, had the opportunity to investigate if the PAM's predictions could be confirmed from both a behavioral and a neurophysiological perspective. The study demonstrates that both language students and naive listeners assimilate foreign sound categories into near-fitting native sound-categories, as predicted (Grimaldi et al., 2014). Furthermore, the authors suggest that the perceived similarity of foreign and native sounds governs learnability for the naïve listeners (Grimaldi et al., 2014).

If the PAM's predictions also apply to musical material, it is expected that sound elements from ambiguous out-of-culture melodies will be assimilated into in-culture categories in the present study's categorization task.

IDyOM

This section will detail the central aspects of the computational model IDyOM, which has been used to analyze the melodies used in this study. Some fundamental concepts from information-theory will be presented, along with a graph that summarizes IDyOM's analysis of the melodies.

The *Information Dynamics of Music* (IDyOM; Pearce, 2005) is a computational model designed to simulate musical cognition through the use of statistical learning and probabilistic prediction (Pearce, 2005). From a theoretical perspective, the model derives from the hypothesis that cognitive processes generate expectations about weighted probabilities of possible continuations of notes, based on the frequency of previous experiences in similar musical contexts (Pearce, 2018; Rohrmeier & Koelsch, 2012).

When exposed to a corpus of music, the IDyOM model is capable of learning the syntactical structure of the corpus in terms of sequential regularities of note interval, and the model then calculates the probability of a musical event occurring at a particular time in the given context based on said regularities (Pearce, 2018). The model provides a quantitative measure for this expectancy in the form of information content (IC), a concept adopted from information theory (MacKay, 2004, p. 32).

The measure IC is based on the information-theoretic quantity *entropy* (MacKay, 2004, p. 32). In a setting where all events are equally probable to occur, and therefore also highly unpredictable, the entropy of a given event will be at a maximum. When predictability is introduced to the setting, thereby making some events more likely to occur than others, the entropy of the predictable events will be lower (MacKay, 2004, p. 22). The IC value of a note reflects how unexpected IDyOM finds the note in the particular context or musical style, meaning the higher the IC value of a note, the more unexpected it is in the context. Variations in IC values of notes have been shown to correspond with variations in neural responses to musical events (Omigie et al., 2013).

IC is measured in the binary unit *bits*, which is considered a more numerically stable unit of information than raw probability (MacKay, 2004, p. 32). IC is defined mathematically in the following formula, where h represents information content in bits, and p represents the raw probability: $h = -log_{10}\, p$ (Pearce, 2018). The IC of a note in a melodic sequence from a particular musical style can, therefore, be calculated to provide a quantitative measure of its sequential expectancy. When the IC content of a sequence of notes are averaged, a mean IC will be obtained that represents the melody's unpredictability in the context of the training data (Pearce, 2018).

When comparing results from two versions of IDyOM that have been trained in two different musical styles, it is possible for the model to compute a measure that functionally represents the distance between the two styles (Pearce, 2018). This measure has been termed *cultural distance* (CD) (Morrison & Demorest, 2016). In a binary comparison of musical styles, positive values represent one style, and negative values represent the other style. Zero represents an equal probability of both styles (Pearce, 2018).

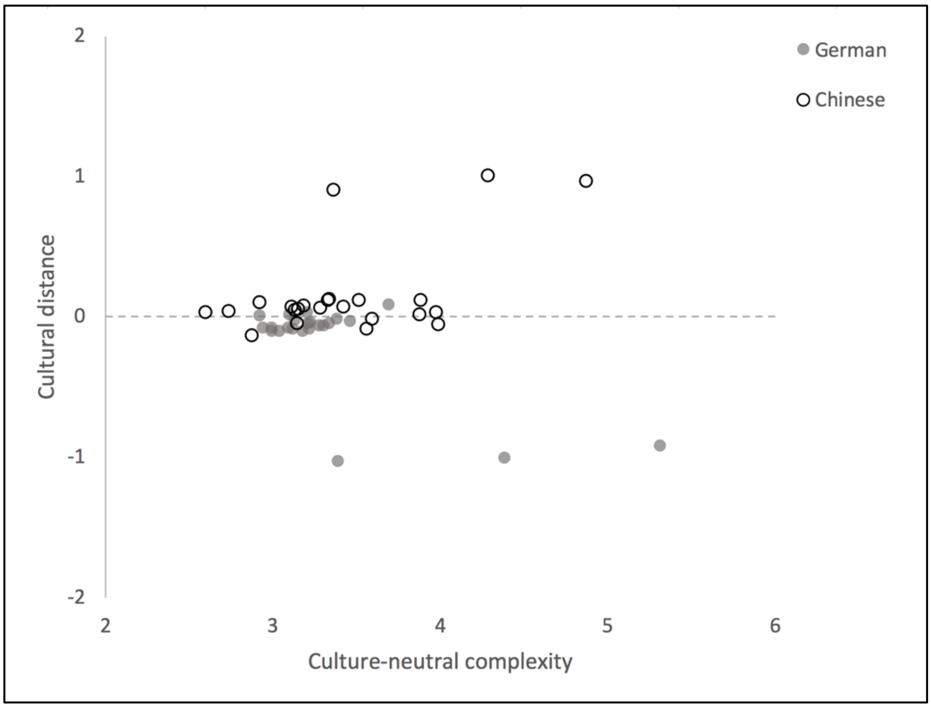


Figure 4. A plot depicting the cultural distance (CD) plotted against the culture-neutral complexity, and a line of equality. The closer a melody is to the dotted line of equality, the more stylistically ambiguous it is considered by the model.

For the melodies used in this project, two versions of the IDyOM model were trained on two separate corpora of music: a German model was trained on 769 traditional German folk songs, and a Chinese model was trained on 858 traditional Chinese folk songs (Pearce, 2018). In this way, the two models are assumed to simulate a Western listener and a Chinese listener, respectively.

In order to provide robust probabilities, each version of the model made both within-culture and between-culture predictions of the melodies in the corpora, meaning that the model trained on German melodies processed the German melodies in one condition (within-culture) and the Chinese melodies in another condition (between-culture), and the Chinese model did the same (Pearce, 2018). The within-culture predictions were made based on the same melodies that were used to train the model. In each condition, IDyOM estimated the IC for all events in the compositions. The note IC was then averaged for each melody, resulting in two mean IC values for each melody: one based on the estimates of the within-culture condition and one on the between-culture condition. These two values were then used to compute a CD value.

Figure 4 depicts a visual comparison of melodies from German and Chinese cultures. CD is plotted on the abscissa, and cultural-neutral complexity is plotted on the ordinate. Cultural-neutral complexity of a melody is defined as the IC ascribed by the melody's primary model (IC of German model for German melodies, and IC of Chinese model for Chinese melodies) (Pearce, 2018). A line of equality (x = y) marks equivalence between the two models, meaning that melodies that land on the line are considered equally predictable by both models (Pearce, 2018).

If IDyOM successfully simulates cultured listeners, the further a melody is from the line of equality, the more culturally distinct it should be perceived by participants, and the closer a melody is to the line of equality, the more culturally ambiguous it should be perceived. Distance from the line of equality can, therefore, be considered a quantitative measure of cultural distance between the two musical cultures involved (Morrison & Demorest, 2016).

IDyOM is a multidimensional model, meaning that it is configured with two functionally identical models: a long-term model (LTM) and a short-term model (STM) (Pearce, 2005). When the LTM of IDyOM is exposed to a corpus of music, for example from a specific geographic area or cultural tradition, the model decodes the syntactic structures and sequential regularities of the musical corpus through unsupervised statistical learning and thereby estimates the conditional likelihood of a musical event occurring at a particular time in the given musical style (Pearce, Ruiz, et al., 2010). The STM, on the other hand, only takes the current composition into account when estimating the probability of note continuations and discards the constructed parameters when the composition has been processed (Pearce, 2005). Predictions from the LTM and the STM can be used separately or combined in order to provide estimations for the probability of a note occurring at a particular time in the given context (Pearce, 2005). The predictions from the LTM and STM are, respectively, assumed to represent the listener's robust knowledge-based expectations based on stylistic enculturation, and their dynamic pattern-sensitive observations constructed during an active listening session (Pearce, 2018). Measures provided by a combination of predictions from the LTM and the STM is therefore considered to most accurately simulate encultured listeners. An anecdotal comparison can be made to the neural responses described in an earlier section: the mismatch negativity is elicited when short-term irregularities occur, while the early right anterior negativity is elicited when long-term irregularities occur (Koelsch et al., 2001). In this way, IDyOM is designed to represents several stages of perceptual processing in humans.

The parameters analyzed by IDyOM are melody, harmony, and rhythm since these can be transformed into symbolic representations relatively efficiently (Pearce, 2005). Research has shown that listeners use various psychological representations of pitch and time when processing music, such as pitch hight, chroma, and interval (Deutsch, 2013), and onset time and duration (Teki, Grube, Kumar, & Griffiths, 2011), either as independent or interacting events (Palmer & Krumhansl, 1987). When analyzing a musical composition, the IDyOM model can process these features of pitch and temporality separately or in combination (Pearce, 2005). The model does not currently process other musical features that rely on textual and timbral changes.

Implicit attitudes

In this section, implicit and explicit attitudes will be examined from the perspective of the APE model (Gawronski & Bodenhausen, 2006), which emphasizes differentiation of the attitudes based on how they are processes cognitively. The implications that this theoretical perspective has on the IAT test are also discussed.

Implicit attitudes are automatically activated actions and evaluations that often take place without the person's conscious awareness of causation (Greenwald, McGhee, & Schwartz, 1998). Implicit attitudes are differentiated from deliberate explicit attitudes in the scientific literature (Gawronski & Bodenhausen, 2006). The *Associative-propositional evaluation* (APE; Gawronski & Bodenhausen, 2006) model has been proposed as a way to conceptualize implicit and explicit attitudes in terms of their underlying processes. The model assumes two differentiated systems that follow different operating principles but interact to integrate reflective and impulsive processes in order to produce behavioral decisions (Strack & Deutsch, 2004).

According to the model, implicit attitudes are characterized by automatic affective activations based on learned associations between stimuli (Gawronski & Bodenhausen, 2006). These activations are termed associative processes. The associative processes do not require much cognitive capacity and are considered unaffected by the assignment of truth values since there is no intention to evaluate the imminent object or stimulus (Cunningham, Raye, & Johnson, 2004). Instead, associative activations rely mainly on spatio-temporal contiguity and feature similarity of the article at hand (Bassili & Brown, 2005; DeCoster, 2002).

Pattern activation is an important aspect of associative processes. The term refers to the idea that particular associations in memory are activated when there is a relative fit between a set of external stimuli and the pre-existing structures of association in memory (E. R. Smith, 1996). This relative fit may vary between different contexts, and the pre-existing structures of association in memory may also vary between individuals (Barsalou, 1982).

On the other hand, explicit attitudes are characterized by so-called propositional processes, according to the APE model (Strack & Deutsch, 2004). The model describes propositional processes as evaluations that are based on syllogistic conclusions derived from the propositional information available at a given time (Gawronski & Bodenhausen, 2006). The propositional processes are assumed to occur in a reflective system that assesses the validity of information transmitted from the associative system by transforming it into a propositional format and comparing it to other propositional evaluations retrieved from memory (Gawronski & Bodenhausen, 2006). As a straightforward example, a positive impulsive reaction to a stimulus will be transformed into the proposition "I like this stimulus," and the proposition will then be subject to a syllogistic analysis to determine its validity based on the person's broader held beliefs.

The most crucial distinction between propositional processes and associative processes is the contrasting dependency on truth values (Strack & Deutsch, 2004). The activation of associative processes can occur without the person considering them to be true on a propositional level of analysis (Strack & Deutsch, 2004). However, the default mode of propositional processing is to affirm the affective reactions arising from the associative processing (Gilbert, 1991). Only if the individual becomes consciously aware of the incongruity between the two processes will the affirmation be invalidated (Gawronski & Bodenhausen, 2006). In other words, if automatic affective reactions are inconsistent with the propositions held at a given time, the latter will most likely invalidate the former. For example, an individual's adverse affective reaction towards a member of a foreign culture ("I do not like this Chinese person") may be inconsistent with the propositional evaluation of another prevailing attitude ("I should not dislike people solely based on ethnicity"). In this way, inconsistency between associative and propositional processes can lead to the rejection of an automatic affective response to a stimulus (Gawronski & Bodenhausen, 2006).

Hypothetical propositions are an important aspect of propositional evaluations. They can be described as propositions of a tentative character, meaning they have been assigned a preliminary truth value (Gregg, Seibt, & Banaji, 2006). Since the default mode of propositional processing is to affirm momentarily activated associations, and because entertaining a proposition can increase activation of comparable associations in memory (Gregg et al., 2006), it is assumed that hypothetical propositions can increase the perceived validity of a supposed evaluation (Koehler, 1991). The same line of argument can be assigned to mere knowledge of a proposition, even if this knowledge is not explicitly endorsed by the individual in question (Devine, 1989). This means that knowledge of an unfavorable cultural stereotype can lead to a negative affective association, even if the individual in question does not consider the stereotype an accurate representation of members of the culture in question.

According to the APE model, consistency assessment of evaluations is solely attributed to processes of propositional reasoning since the assessment of consistency must involve the attribution and alignment of truth values to the considered propositions (Strack & Deutsch, 2004). As mentioned above, the attribution of truth value only happens through propositional reasoning. When there is an inconsistency to be resolved between two propositions, the inconsistency can be resolved through various propositional processes like hierarchical inhibition (Bodenhausen & Macrae, 1998), rejection (Gilbert, 1991), rationalization (Festinger, 1957), or justification (Crandall & Eshleman, 2003). As an example relevant to this study, when the face image of a person from a foreign culture activates negative associations but the subject taking the implicit association test has no animosity towards people from this culture on a propositional level of analysis, the automatic reaction may be inhibited, negated, rationalized, or justified, but only if there is an inconsistency with other reflective evaluations during processing (Strack & Deutsch, 2004). This underscores the point that propositional processes are superordinate to associate processes, and in an experimental setting where propositional processes are deliberately excluded through a strict response timeframe, like in the implicit association test, results should be interpreted cautiously.

The amount of active and elaborative thoughts put into an attitude or position impacts which propositional evaluations will be considered in addition to the automatic affective response (Kruglanski & Thompson, 1999). More comprehensive elaboration is generally assumed to increase the number of propositions deemed relevant for the momentary evaluation, and this increases the likelihood of inconsistencies, which in turn can reduce the correlation between automatic affective reactions and propositional judgments (Judd & Lusk, 1984). This means that more extensive cognitive elaboration on a subject does not increase or decrease the correlation between implicit and explicit attitudes per se, but only if the propositions question the validity of the affective reaction.

The implicit association test and implications of the APE model

The implicit association test (IAT; Greenwald et al., 1998) was developed and introduced to the scientific literature by researchers Greenwald, McGhee, and Schwartz (1998). The test is designed to evaluate the user's implicit attitudes towards various matters by estimating the strength of association between two given concepts and positive/negative attributes. When results from the test have been analyzed, the test produces a so-called *D-score* that can be either positive or negative, depending on which concept has a stronger association with the positive attribute (Greenwald, Nosek, & Banaji, 2003).

A principal assumption underlying the test is that by alternately grouping two discrimination tasks onto a single pair of possible responses, the strength of association between the given concepts and attributes can be revealed by analyzing response time

(Greenwald et al., 1998). It is expected that superior performance in the test reflects more effective cognitive processing of certain groups of stimuli, which results from more automated associations between constellations of the particular concept and attribute (Greenwald et al., 1998).

The IAT is designed to resist some self-presentation strategies by confining the response timeframe. This is assumed to assist in uncovering attitudes that participants do not want to express explicitly or are not fully conscious of (Greenwald et al., 1998). Only responses given between 0,2 to 3 seconds will be included in the test's analysis. The rationale behind this restriction is that responses that have not been processed cognitively (< 0.2 seconds) are excluded since the average time from stimulus onset to motor-response initiation is 200 milliseconds (Welford, 1980). Responses that are likely to have been consciously processed and adjusted (> 3 seconds) for a more favorable self-presentation are also excluded since these, in terms of the APE model, have been subject to propositional considerations (Gawronski & Bodenhausen, 2006).

The considerations in the section above have implications on the interpretation of the IAT. Results from the IAT are indirect measures of the participants' implicit attitudes that provide a proxy for activations of associative processes in memory (Gawronski & Bodenhausen, 2006). A stronger association between a concept and attribute does not necessarily reflect a generally held attitude, like a preference for a particular culture or racial group. However, studies have shown that the IAT does assess certain unobtrusive tendencies quite reliably, for example, racial attitudes (Fazio, Jackson, Dunton, & Williams, 1995).

Taking the above into account, it is clear that results from the IAT should be interpreted cautiously, as these are not highly stable measurements of a person's attitude towards a given concept. In the IAT used for this study, a negative D-score indicates shorter reaction time when Danish faces are grouped with positive words, and a positive D-score indicates shorter reaction time when Chinese faces are grouped with positive words. The larger the D-score, the more significant the difference in reaction time between the two conditions. Therefore, a negative D-score in this study does not necessarily reflect a negative cultural bias towards Chinese people or culture, but rather it may merely reveal stronger associations in memory between faces of Danish people and positive attributes (Gawronski & Bodenhausen, 2006). These associative processes sometimes correlate with propositional evaluations that are explicitly endorsed, but it cannot and should not be assumed that they always do.

The present study

The present study aims to investigate how musical enculturation and implicit cultural attitudes affect musical perception. More specifically, the aim is to investigate how accurately human listeners will classify music from two different cultural styles in a behavioral categorization task and to see if the results are comparable to results from a computational model that is designed to simulate the mechanisms underlying musical enculturation (Pearce, 2018). Furthermore, we want to investigate if categorization performance can be predicted by implicit attitudes towards the ethnic groups directly related to the musical styles used in the categorization task. To measure implicit attitudes, we use results from the implicit association test (IAT) as a proxy since it has been shown to assess ethnic attitudes somewhat reliably (Fazio et al., 1995). The objectives mentioned above will be investigated in compliance with the following research question: "*Do human listeners classify culture-specific melodies similarly to a computational model in a musical categorization task, and can their performance be predicted by implicit attitudes?*"

In order to do this, four hypotheses have been formulated. The first and second hypotheses will investigate how facets of enculturation can affect musical perception, while the third and fourth hypotheses will investigate whether implicit cultural attitudes influence the perception of musical style.

In predictive coding theory, statistical learning and probabilistic prediction are hypothesized to be fundamental mechanisms of musical cognition (Friston, 2010; Pearce, 2018). IDyOM is a computational model that has been designed to simulate musical cognition in humans (Pearce, 2005). It does this by learning the syntactic regularities of pitch intervals that underlie a particular musical style. The model then quantifies predictions of note-occurrences based on said regularities, thereby simulating the enculturation process in humans (van der Weij et al., 2017). The musical compositions used in this study have been analyzed by two versions of IDyOM that have been trained in Chinese and German musical styles. Based on the information content (IC) derived from the two versions, IDyOM provides each melody with a combined quantitative value of cultural distance (CD) that reflects the probability of a melody belonging to one of the two styles. In other words, the CD value represents the melody's stylistic specificity (Morrison & Demorest, 2016). In this study, a positive CD value indicates that the melody is most likely Chinese, and a negative CD value indicates that the melody is most likely German. The first hypothesis has been formulated in order to investigate whether human listeners perform comparatively to the computational model in a categorization task:

(H1) *"There will be no significant difference in categorization accuracy between participants and the computational model."*

Following this vein, the closer a CD value is to zero, the more ambiguous it is considered by IDyOM. The musical compositions selected for this study can be classified into two groups based on high or low CD value, namely distinct and ambiguous melodies. Therefore, a difference in the categorization accuracy of melodies with a high CD value (distinct) and low CD (ambiguous) is expected. The second hypothesis has been formulated to investigate whether human listeners are better at categorizing distinct melodies than ambiguous melodies: (H2) *"There will be a significant difference between participants' categorization accuracy of distinct and ambiguous melodies."*

The IAT was introduced to the scientific literature as a tool to measure associations between target concepts and attributes (Greenwald et al., 1998). Results from experimental studies have suggested that musical engagement can facilitate cultural understanding (Clarke, DeNora, & Vuoskoski, 2015) and influence implicit cultural attitudes (Neto, Pinto, & Mullet, 2018). In this study, we want to take a different approach to this relationship and investigate if implicit cultural attitudes can predict the perceived differentiation of musical styles. The following hypothesis has, therefore, been formulated: (H3) *"There will be a significant difference in categorization-task performance between participants who score high and low in the Implicit Association Test."*

Theoretical propositions have suggested an underlying parallel between the organization of linguistic and musical cognition in terms of hierarchical and logical structuring (Patel, 2012, p. 262). The perceptual assimilation model is a linguistic model that makes specific predictions about the perception of non-native sound categories, namely that ambiguous non-native sound categories will be assimilated into already existing native language sound categories (Best, 1995). These predictions are based on the hypothesis that native sound categories function as "perceptual magnets" that attract similar nonnative phonetic sounds. In order to inquire if these predictions remain valid for non-native musical material, the fourth hypothesis in this study will investigate if there is a relationship between implicit cultural attitudes and the categorization accuracy of ambiguous out-of-culture melodies. The assumption underlying this hypothesis is that the former variable functions as a proxy for assumed foreignness towards the non-native culture, and the latter represents the perception of non-native sound categories. Therefore, the following hypothesis has been formulated: (H4) *"There is a relationship between participants' score on the implicit* *association test and their probability to categorize ambiguous out-of-culture melodies as in-culture melodies."*

Method

The present study comprises two administered tests: a custom made musical categorization task and an implicit association test (IAT; Greenwald, McGhee, & Schwartz, 1998). In the musical categorization task, participants are asked to categorize 46 short melodies according to the cultural origin of the melodies, with the available categories being either "Kina" (China) or "Danmark" (Denmark). The second category, "Danmark," was used in the test even though the melodies in this category belonged to a German song corpus. This will be discussed later.

In the IAT, participants are asked to sort four types of stimuli into two groups as fast as possible. The types of stimuli in the IAT are positive and negative adjectives (in Danish) and face images of ethnic Chinese and Danish people.

Participants

Ten young participants took part in the experiment, age range 21 to 31 years old (M = 25.5, SD = 3.31, 5 female). They were recruited through public adverts on social media. At the time of testing, all participants were active students at Aalborg University. All participants were ethnically Danish and had spent the majority of their childhood and adolescence living in Denmark. Five of the participants (50%) reported that they had formally received musical training for three years or longer in their lifetime, and two of these participants (20%) still received training or actively played instruments at the time of testing. There was no inclusion criterion regarding musicality. None of the participants reported having hearing impairments when asked, although this factor was not measured.

Materials

Categorization task

A custom-developed computer program was used for the administration of the musical categorization task. The program is based on the programming language Python. The test was administered on a MacBook Pro computer (running on MacOS X version 10.13.6), and participants were provided with high-quality headphones (Beyerdynamics DT 770 pro) for listening to the melodies.

The 46 melodies used in the categorization task were selected from the Essen Folksong Collection (Schaffrath, 1995). The melodies were converted from their original "MIDI" format into "WAV" in order to be compatible with the computer program. *MIDI* (Musical Instrument

Digital Interface) is a standardized digital format that uses numeric information to represent musical events (Huber, 2012). A MIDI file is, therefore, in principle, not an audio file, but through the use of an appropriate musical computer program, it can be played back using any selected instrument and eventually converted into an audio file. *WAV* (Waveform Audio File Format) is a standard audio file format used to store an audio bitstream digitally (Bhatnagar, Mehta, & Mitra, 2004). The computer application Cubase (version 6) was used to convert the MIDI files into WAV files. The timbre used was a standard featured instrument called *orchestral harp*. All melodies were devoid of supplementing instrumentation or harmonization. Melodies were between 6 – 19 seconds long (M = 12.6, SD = 3.5).

The categorization task started with a training trial featuring four melodies played on either piano (50%) or harp (50%), in order to ensure that participants were acquainted with the application's technical procedure. The layout of the training trial was identical to the main trail, apart from the two options being "piano" or "harp" instead of "Kina" or "Danmark," as in the main trial. Besides the training trial, the categorization task consisted of only one trial, which included all 46 melodies. Participants were presented with a melody and asked if the melody was of Chinese or Danish origin. They were also asked to provide a confidence rating for their answer on a five-point Likert scale (*1 = Ikke sikker; 5 = Meget sikker)* and to note if they knew the melody beforehand.

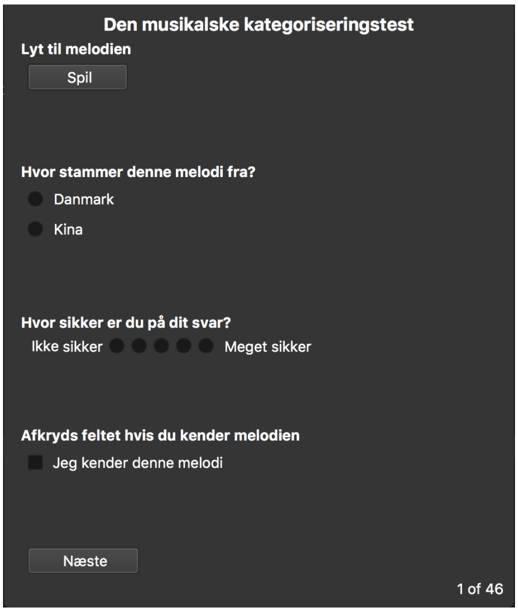


Figure 5. An example of the main screen of the categorization task. The first three items were obligatory and could not be skipped. The last item "Afkryds feltet hvis du kender melodien" was optional.

The selected melodies had been analyzed by two different versions of the computational model IDyOM. Each version provides a measure of information content (IC) for each melody. The IC reflects the probability of a melody belonging to a particular cultural style. From the two measures, a single value is extracted, which represents the melodies' cultural distance (CD) in terms of musical style. A low CD number reflects stylistic ambiguity, and a high CD number reflects stylistic distinctness. Of the selected melodies, 40 (87%) are ambiguous, and 6 (13%) are distinct.

Melodies

The German melodies used in the study are folk songs composed in the 18th and 19th centuries (Schaffrath, 1995). They belong to the Bohemian music tradition (Chesterton, 1993, pp. 42-49). The Chinese melodies used in the study are from two separate Chinese song corpora called *Han* and *Shanxi* (Schaffrath, 1995). It is unknown when these melodies were composed.

The decision to portray melodies from a German song corpus as Danish in the categorization task was made on the assumption that the minimal differences between German and Danish music traditions would disappear when compared to melodies from the distinct Chinese musical tradition. Since the concepts in the implicit association test were Chinese and Danish, it was assumed to be more appropriate to portray the German melodies as Danish as well, as this would not confuse participants unnecessarily. The categories used in the categorization task should, therefore, be understood as cultural approximations, rather than specific geographical categories.

Implicit Association Test

In order to measure the participants' implicit cultural attitudes, an open-source implicit association test (IAT) was used (Greenwald et al., 2003; Nosek, Bar-Anan, Sriram, Axt, & Greenwald, 2014). The test was administered via the computer software PsychoPy 3 (Peirce et al., 2019) on the same computer that was used for the previous task.

The IAT measures latency in response time as participants are asked to sort groupings of attributes and concepts into two categories (Greenwald et al., 1998). The attributes used in the present test were positive words and negative words, and the concepts used were face-images of Chinese people and face-images of Danish people. In this particular version of the

Block	No. of trials	Function	Item assigned to left-key response	Item assigned to right-key response
1	20	Practice	Chinese images	Danish images
2	20	Practice	Negative words	Positive words
3	20	Practice	Negative words + Chinese images	Positive words + Danish images
4	40	Test	Negative words + Chinese images	Positive words + Danish images
5	40	Practice	Danish images	Chinese images
6	20	Practice	Negative words + Danish images	Positive words + Chinese images
7	40	Test	Negative words + Danish images	Positive words + Chinese images

Figure 6. An illustration showing the blocks in the implicit association test. The position of blocks was counterbalanced between participants. Blocks 1, 3, and 4 were switched with blocks 5, 6, and 7 for half of the participants.

test, the response keys "*e*" and "*i*" are designated to represent the categories to the left and the right, respectively. In the first half of the test, the software randomly selects one attribute and one concept to be represented to the left and the remaining attribute and concept to the right. For example, positive words and Danish faces are represented to the left, and negative words and Chinese faces to the right. In the second half of the test, the concepts switch sides, but the attributes do not so that the left now represents positive words and Chinese faces, and to the right represents negative words and Danish faces.

The underlying assumption is that participants must make mental categories that represent the groupings of stimuli, and differences in response time reveal differences in the strength of association between particular attributes and concepts. Shorter response time is expected for groupings, where the association between attributes and concepts is stronger (Greenwald et al., 1998). For example, Danish participants are expected to have a shorter response time when positive words and Danish faces are grouped.

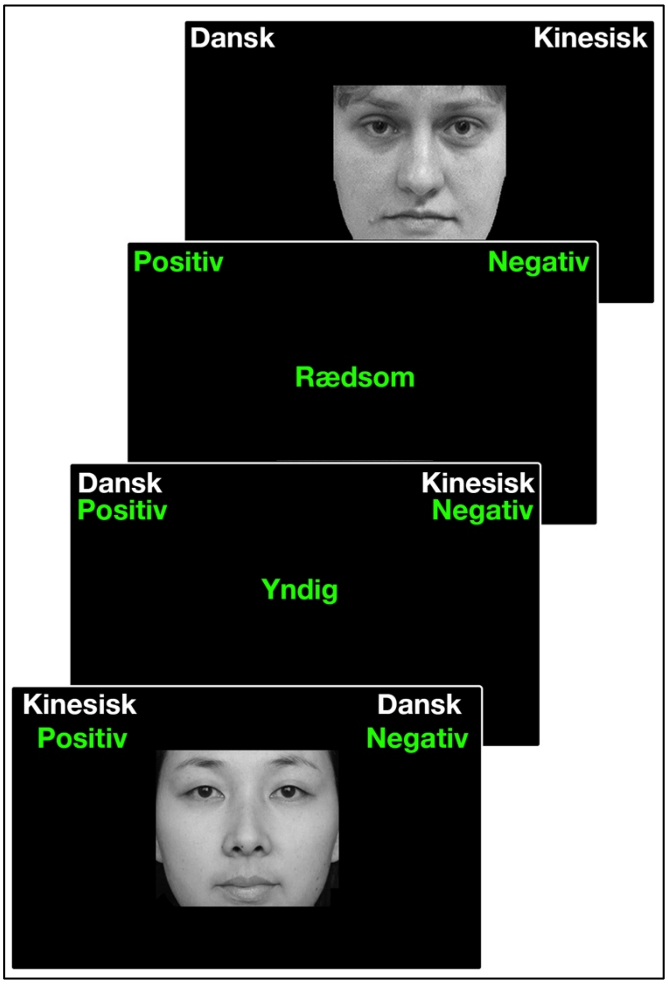


Figure 7. Illustration showing the different conditions in the IAT test. Notice that in the bottom condition the concepts have changed sides, but the attributes have not.

The facial images used are acquired from the NimStim Set of Facial Expressions database (Tottenham et al., 2009). Ten images in all, five of each ethnicity, six males and four females. All expressions are deemed neutral (Tottenham et al., 2009). The images have been edited into grayscale and cropped from the eyebrows to the chin. The image size is 350 x 225 pixels.

The present IAT consists of seven blocks, of which five are practice blocks, and two are critical blocks. Block order is balanced between conditions and participants in order to minimize the influence of primacy bias (Nosek et al., 2014).

All text in the test was presented visually, and all text was in the Danish language, including the stimuli words. The five positive words are *tiltalende, dejlig,* *yndig, pragtfult,* and *vidunderlig*. The five negative words are *frygtelig, beskidt, rædsom, afskyelig,* and *ondskabsfuld*.

The method of scoring in the present version of the IAT is the same as described in Greenwald et al. (2003). In short, the within-person difference in response latency is divided by the standard deviation of the combined critical and practice blocks in the test to yield a so-called *D-score*. In the present test, a positive D-score represents a positive bias towards Chinese culture, and a negative bias represents a positive bias towards Danish culture.

Procedure

In order to minimize potential confounding factors that might affect implicit or explicit attitudes towards the experiment as a whole, the aim of the study was not fully revealed for the participants beforehand. The tests were introduced as auditory and visual sorting tasks. Also, the participants were informed that the second task (IAT) was timed. The explicit task of categorizing melodies was administered before the implicit task. This order was used for all participants rather than, for example, randomization between tests, in an effort to access the top-down cognitive processes participants used in musical categorization without potential interference of the attributes and concepts used in the IAT (Martin, 2008, p. 31). Before the experiment started, participants received a short oral introduction and instructions regarding the tests. The participants entered a small room that was prepared for the experiment. The only objects in the room was a chair, a table, and a computer with headphones. Each test had detailed technical instructions on screen before it began.

Only the musical categorization task required headphones. The sound volume on the computer was set to 70% of full volume. The categorization task started with a short training trial. After the training trial was over, the participants were asked if they were comfortable with the sound volume and the technical procedure of the application. Adjustments were made to the volume if necessary, and the administrator left the room. The 46 melodies in the categorization task were presented in random order. Each melody could only be played once, and it was not possible to proceed to the next melody without listening and providing answers related to the present melody. The categorization task took between 15–20 minutes to complete.

When the categorization task was complete, the participant notified the administrator, who reentered the room in order the start the IAT on the computer. Participants were asked to sit at arm's length from the computer screen and instructed to answer as quickly as possible.

The categories were visible in the top right and left corners during the test. Halfway through the test, the category groupings changed, as described in the section above. The IAT took about 6 minutes to complete.

The entire experiment took about 25–30 minutes to complete. After completion, participants were debriefed about the study and had the opportunity to ask questions.

Results

In the following four sections, results from the study's hypotheses will be presented separately. Ten subjects participated in this experiment. The data used in these analyses were found to be approximately normally distributed. Therefore, parametric statistical tests were appropriate in order to test the hypotheses. Alpha level of .05 was used for all statistical tests.

Hypothesis 1

The first hypothesis in this study states the following: *"There will be no significant difference in categorization accuracy between participants and the computational model."* In order to investigate this hypothesis, the participants' and the computational model's mean categorization accuracy must be calculated and compared in statistical analysis.

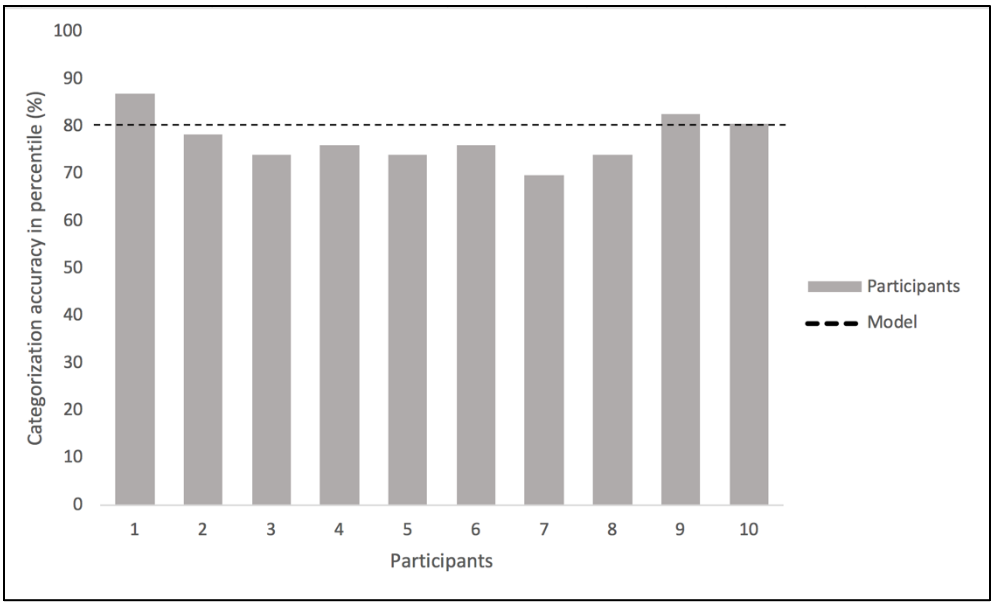


Figure 8. Columns showing participants' categorization accuracy and a dashed line showing IDyOM's categorization accuracy (80.43%). No statistically significant difference was found between participants' average categorization accuracy (77.17%) and that of the computational model.

Two versions of the IDyOM model were used to analyze the melodies selected for this study, namely one trained in Chinese music and one trained in German music. A combined measure of cultural distance (CD) is then provided for each musical composition in the dataset, reflecting the probability of a particular melody belonging to a particular style. In the present paradigm, Chinese melodies are represented by CD values greater than zero (> 0), and German melodies are represented by CD values less than zero (< 0). A CD value of absolute zero (= 0) would indicate an equal probability of being Chinese or German, but no melody in the dataset met this criterion. Of the selected compositions used in this study (N = 46), nine were given an "incorrect" CD value (positive CD value for German melodies and negative CD value for Chinese melodies), meaning that the model found it more probable that the melody belonged to the opposing cultural style than it actually did. Therefore, the remaining 37 of the total 46 melodies (80.43%) were considered accurately categorized by the computational model. The participants' combined mean categorization accuracy was calculated and found to be 77.17% (*SD* = 5.05).

A One-sample t-test was conducted to determine if a statistically significant difference in means existed between the categorization accuracy of human listeners and that of the computational model. The results from the One-Sample t-test ($\mu \neq \mu_0$) indicate that there is no statistically significant difference between the categorization accuracy of human listeners (*M* = 77, *SD* = 5) and the computational model, $t(9) = -2.04$, $p = 0.072$, $d = -0.65$.

Hypothesis 2

The second hypothesis in this study states the following: *"There will be a significant difference between participants' categorization accuracy of distinct and ambiguous melodies."* In order to investigate this hypothesis, the participants' mean categorization accuracy of distinct and ambiguous melodies must be calculated and compared in statistical analysis.

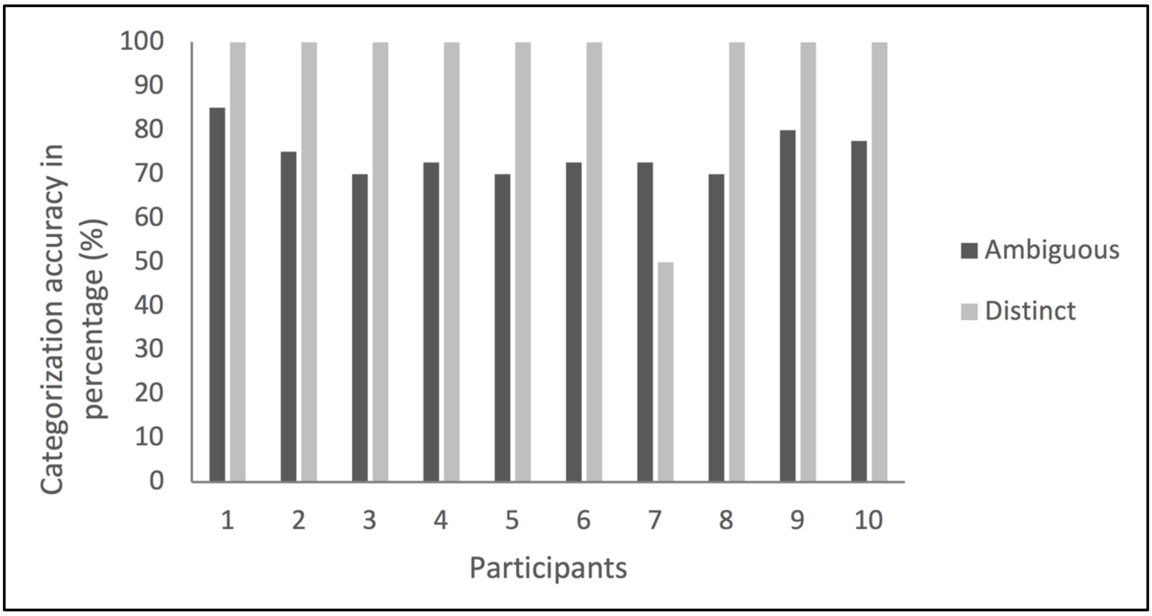


Figure 9. Columns showing participants' categorization accuracy of ambiguous and distinct melodies. A statistically significant difference between categorization accuracy of ambiguous and distinct melodies was found, *p = .003.

The CD value of the 46 melodies selected for this study ranges from 1.01 to -1.02. Of these, 40 (86.96%) had CD values ranging from 0.13 to -0.13 and were considered stylistically ambiguous. The remaining 6 melodies (13.04%) were considered stylistically distinct and had CD values that ranged from 1.01 to 0.91 (three Chinese melodies) and -0.92 to -1.02 (three German melodies). Therefore, a group of ambiguous melodies and a group of distinct melodies were formed. Participants' categorization accuracy for each group of melodies was calculated and found to be 74.5% (*SD* = 4.97) for the ambiguous group, and 94.9% (*SD* = 15.81) for the distinct group. The ambiguous group consisted of 40 melodies, and the distinct group consisted of 6 melodies.

A Paired-Samples t-test was conducted to determine if a statistically significant difference in means exists between the categorization accuracy of ambiguous melodies and the categorization accuracy of distinct melodies for human listeners. The results from the Paired-Samples t-test indicate that there is a statistically significant difference between the categorization accuracy of ambiguous melodies and the categorization accuracy of distinct melodies for human listeners, $t(9) = 4.08$, $p = .003$, $d = 1.97$.

Hypothesis 3

The third hypothesis in this study states the following: *"There will be a significant difference in categorization-task performance between participants who score high and low in the Implicit Association Test."* In order to investigate this hypothesis, participants must be divided into two groups according to their IAT performance, so that the mean categorization accuracy for each group can be compared in statistical analysis.

One group is comprised of participants who scored low in the IAT. This "low-group" was calculated to have a mean categorization accuracy of 75.22% (*SD* = 4.24) in the categorization task. The "high-group" consists of the remaining participants who got higher scores on the IAT and was calculated to have a mean categorization accuracy of 79.13% (*SD* = 5.46) in the categorization task. The formation of the two groups was based on finding the median value of the participants' IAT scores and partitioning two equally sized groups. In this sense, the terms "high-group" and "low-group" in reference to IAT scores are local and do not reflect global scoring tendencies.

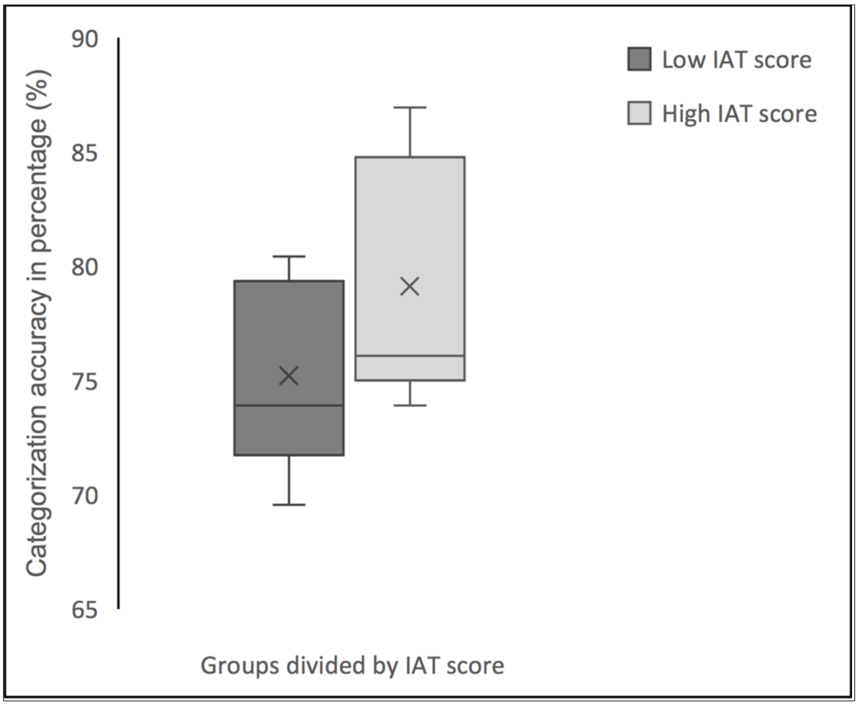


Figure 10. Boxplot of participants' categorization accuracy in percentage, grouped according to IAT scores (D-scores).

An Independent Samples t-test was conducted to compare the categorization accuracy of human listeners who scored low (D-scores between 0 and ±.54) and who scored high (D-scores between ±.55 and ±1.2) on the IAT. The results from the Independent Samples t-test ($\mu_1 \neq \mu_2$) indicate that there was no statistically significant difference in categorization accuracy between low-scoring (M = 75.22, SD = 5.5) and high-scoring (M = 79.13, SD = 4.2) listeners, $t(8) = 1.27$, $p = .24$, $d = 0.8$.

Hypothesis 4

The fourth hypothesis in this study states the following: *"There is a relationship between participants' score on the implicit association test and their probability to categorize ambiguous out-of-culture melodies as in-culture melodies."* In order to investigate this hypothesis, the participant's categorization accuracy of ambiguous melodies and the participants' IAT scores must be used as variables in a correlational analysis. Each participant's categorization accuracy of ambiguous melodies was found by isolating responses to the 40 ambiguous melodies in the categorization task and converting the categorization accuracy of this group of melodies into a percentile.

A Pearson's correlation was conducted to test if there is a relationship between the categorization accuracy of ambiguous melodies and participants' IAT scores. The results from Pearson's correlation indicate that there is no statistically significant correlation between the categorization accuracy of ambiguous melodies and participants' IAT score, $r(10) = .352$, $p = .159$.

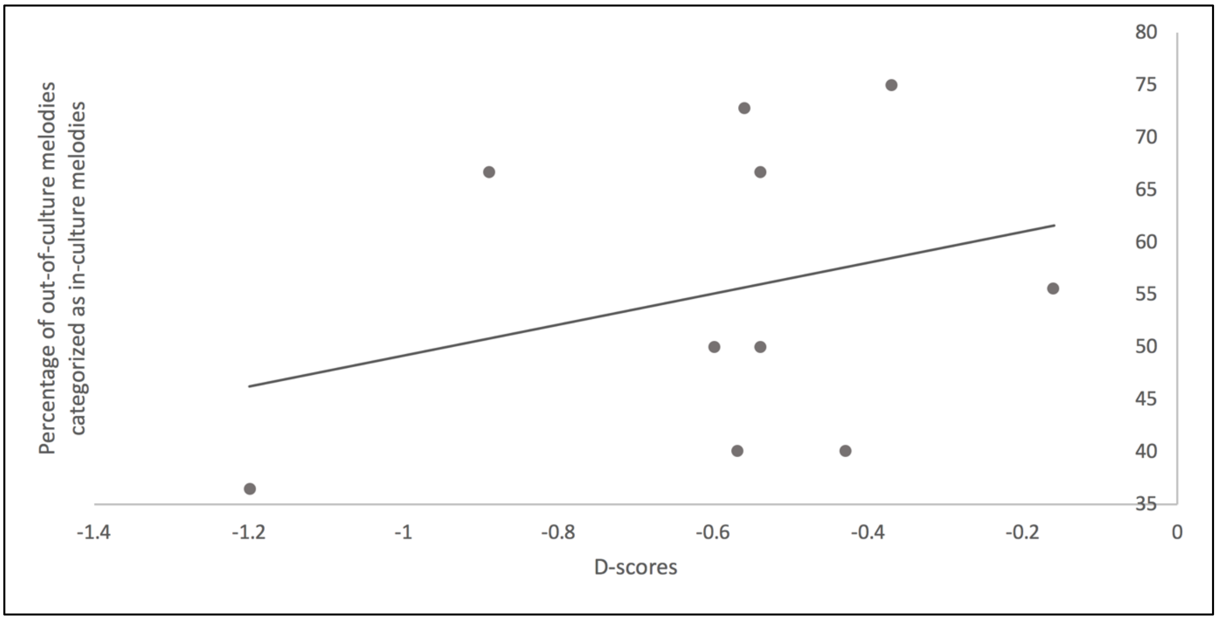


Figure 11. A graph depicting the relationship between the IAT's D-scores and percentage of out-of-culture melodies categorized as in culture melodies in the categorization task. The correlation test was not significant, *p* = .159.

Discussion

This discussion is organized into two main sections. In *Implications of results,* the results from each hypothesis are discussed separately. The main point in the discussion of the first and second hypotheses is that the computational model appears to accurately simulate musical processing in humans, as it is designed to do. Also, the potential issue of comparing groups of different quantities is discussed. In the discussion of the third and fourth hypotheses, which both remain unconfirmed, the focus is on the possibly that the small group of participants may be too uniform to yield a significant result, and whether predictions from a linguistic model can be directly transferred to musical material.

In the section *General discussion*, the overall results are discussed relative to the study's research question. It is also discussed whether the music that the IDyOM model has been trained on is representative of the music in which participants are encultured. Lastly, other perspectives that could be relevant for future research in the subject are discussed, such as musicality, the variance of empathy, and personality structure.

Implications of results

This study aimed to investigate the effect that musical enculturation and implicit cultural attitudes have on the categorization of short musical compositions. The results confirm the first and second hypotheses, namely that (H1) there will be no significant difference in categorization accuracy between participants and a computational model, and that (H2) there will be a significant difference between participants' categorization accuracy of distinct and ambiguous melodies. The results did not confirm the third hypothesis (H3), which states that there will be a significant difference in categorization accuracy between participants who score high and participants who score low on the implicit association test (IAT). The results did not confirm the fourth hypothesis (H4) either, which states that there will be a relationship between participants' IAT scores and their probability to categorize ambiguous out-of-culture melodies as in-culture melodies.

The first two hypotheses are based on the assumption that the computational model IDyOM is capable of providing quantitative measures of stylistic specificity by simulating cognitive processes of enculturation (Pearce, 2018; van der Weij et al., 2017). The findings in this study demonstrate that statistical learning and probabilistic prediction can plausibly account for the effect that musical enculturation has on the perception of distinct musical styles.

Hypothesis 1

The 46 melodies used in the study were selected from two corpora consisting of 858 Chinese and 769 German compositions (Pearce, 2018; Schaffrath, 1995). The selected melodies were disproportionally ambiguous, in reference to the cultural distance (CD) value provided by the IDyOM model. In terms of CD value, IDyOM's categorization accuracy of the selected dataset was 80.43%. In a study using the full corpora previously mentioned, Pearce (2018) found that the IDyOM model correctly classified 98.52% of the melodies. It is strongly assumed that the categorization accuracy of the participants in the present study would be higher if the dataset included proportionally more distinct melodies. This assumption can be corroborated by the second hypothesis's results, in which distinct melodies were shown to be significantly more accurately classified than ambiguous melodies.

For a computational model that simulates auditory cognition in humans, it is important not to overgeneralize nor undergeneralize the learning principles used in the analysis of the data at hand (Pham & Triantaphyllou, 2008). *Overgeneralization* occurs when a model overfits existing but inappropriate datapoints inferred from training trials onto new data, thereby producing false-positive classifications (Pham & Triantaphyllou, 2008). *Undergeneralization*, on the other hand, occurs when true data points are not adequately inferred from training trials, resulting in the poor predictive performance of a model. When considering the results from the first hypothesis, this study provides preliminary evidence that when analyzing new data, the computational model IDyOM neither overfits nor underfits the generalization principals learned in the training trials. Instead, the model performs comparatively to human listeners when classifying the melodies in the present dataset. It remains unconfirmed whether the computational model's categorization accuracy and that of human listeners would be statistically similar if the data used were more representative of the full song corpora from which the melodies have been selected, but it is strongly assumed.

Hypothesis 2

The second hypothesis demonstrates that cultural distance (CD) is a valid measure of cultural specificity since there is a significant difference between participants' categorization accuracy of distinct and ambiguous melodies. The CD values of the entire corpora, from which the melodies were selected, ranged from -1.09 to 1.42. In the present paradigm, negative CD values indicate German origin, and positive values indicate Chinese origin. As mentioned earlier, the selected melodies were disproportionally ambiguous. 86.96% ($N = 40$) belonged to an "ambiguous group" consisting of melodies with CD values on the interval -0.13 to 0.13, and the remaining 13.04% (N = 6) belonged to a "distinct group" consisting of melodies with CD values on the intervals -0.92 to -1.02 (German) and 0.91 to 1.01 (Chinese).

All participants except one categorized all distinct melodies correctly, as can be seen in Figure 10 in the section *Results*. Participant no. 7 miscategorized three distinct melodies, resulting in a lower categorization accuracy for distinct melodies than the categorization accuracy for ambiguous melodies. All other participants had higher categorization accuracy for distinct melodies. This highlights the drawback of having a large difference in the number of melodies representing each group, namely that a small number of mistakes can have a disproportionally large influence on the overall score. The result should not necessarily be taken as an indication that the individual found it easier to categorize ambiguous melodies than to categorize distinct melodies, but rather, the result points to an underlying problem in the experimental design.

Notably, all three categorization mistakes of distinct melodies by participant no. 7 were German melodies that were perceived and categorized as Chinese melodies. This is interesting in at least two ways. First, this is an inversion of the sentiment of the study's fourth hypothesis, which predicts that ambiguous out-of-culture melodies will be perceived as in-culture melodies. In the case of participant no. 7, distinct in-culture melodies were perceived to be out-of-culture melodies. Second, this may suggest that the timbre used in the melodies in the categorization task, a so-called orchestral harp, sounded somewhat like a traditional Chinese harp instrument. This was also noted by some participants directly after the experiment. Therefore, the participant may have taken the timbre used in the categorization task as evidence for a particular classification that proved to be incorrect. This corresponds with results from a study reviewed in the section *Schema*, where genre identification was suggested to be strongly influenced by style-typical timbres (Gjerdingen & Perrott, 2008).

Hypothesis 3

The third and fourth hypotheses in this study are based on the assumption that implicit attitudes, as measured by the IAT, can influence participants' perception of musical material. The findings here cannot corroborate either of the hypotheses, and they remain unconfirmed. The large effect size in the third hypothesis (d = 0.8) suggests that a larger sample size might yield a statistically significant result, given the participants perform comparably to the participants in this pilot study. When running the numbers from the third hypothesis in *G*Power*, a program that computes statistical power analysis, it predicted that 35 participants in each group, 70 altogether, might yield a significant result if performance was similar to the small sample used here. The output from the power analysis is included in the appendix.

The scores from the implicit association test (IAT) are shown in Figure 12 in the section *Results*. IAT scores are based on differences in participants' response time in opposing conditions. For example, when a concept is grouped with a positive attribute in one condition, as opposed to when the concept is grouped with a negative attribute in another condition. Half of the IAT scores appear to cluster in an interval between -0.54 to -0.6, indicating that these participants have similar implicit cultural attitudes. One explanation for this cluster may be that the participants belong to a homogeneous demographic segment: ethnically Danish university students in their 20's, living in Aalborg. It is plausible that their exposure, affiliation, and attitude to Chinese culture, as portrayed in the IAT test, is similar.

Studies have shown that listening to music from foreign cultures can affect implicit cultural attitudes, as measured by the IAT (Vuoskoski et al., 2017). The third hypothesis in this study sought to investigate other possible aspects of the relationship between implicit affiliation and musical engagement, namely whether implicit cultural attitudes can influence the perception of musical style. It could be that people who scored low on the IAT have had more exposure to the foreign culture in question since a low score indicates stronger associations between the foreign culture and positive attributes. They are, therefore, more likely to have listened to Chinese music, which, in turn, would provide them with perceptual representations that can facilitate better performance in the categorization task. However, the results from the third hypothesis do not indicate that implicit cultural attitudes influence perception of musical style.

Hypothesis 4

The fourth hypothesis states that ambiguous foreign melodies are more likely to be classified as native melodies, based on the assumption that the cognitive model used to process pitch elements in the melodies tends to assimilate out-of-culture sound categories into the most appropriate in-culture sound category. In other words, ambiguous perceptual events are most likely attributed to a familiar source. The results could not confirm a correlational relationship between IAT scores, and the number of out-of-culture ambiguous melodies perceived as in-culture melodies, so the hypothesis remains disconfirmed.

Through years of informal exposure, participants are assumed to have formed cognitive models of the native musical style that consists of mental representations of the prominent characteristics of said style, like pitch interval, rhythm, structure, and timbre. When listening to an out-of-culture musical style, expectations of ensuing musical events are regularly violated. When there is a large discrepancy between the expected event and the actual event, as is expected in distinct melodies, the stylistic classification will be easy. On the other hand, when there is a small discrepancy between the expected event and the actual event, as is expected in ambiguous melodies, the stylistic classification will be difficult. If predictions from the perceptual assimilation model (PAM) can be generalized to include musical sound categories, perception of such ambiguous sound elements will presumably be assimilated into already established native sound categories, resulting in the perception of out-of-culture melodies as being in-culture melodies.

It is doubtful whether specific predictions from a linguistic model like the PAM can be directly transferred onto musical material. As discussed earlier, there are similarities between organizational principles underlying the language system and the musical system in terms of hierarchical and logical structuring. However, direct comparison between the systems should be avoided, and it is questionable to what extent such abstract structural similarities in two different systems will be reflected in concrete predictions pertaining to perceptual processing (Patel, 2003).

Another potential issue regarding the fourth hypothesis is whether results from the IAT test serve as a valid proxy for foreignness towards a non-native culture, as assumed. Automatic activations in memory are hypothesized to underlie implicit attitudes, as measured by IAT (Gawronski & Bodenhausen, 2006). Such activations have been shown to correspond to measures of racial attitude (Fazio et al., 1995). However, equating these indirectly associated variables to reveal a potential relationship with participants' perception of ambiguous out-of-culture melodies has not been successful in this study.

Overall, the perceptual magnet affect, as predicted by the PAM, could not be reproduced in this study.

General discussion

Overall, these results provide confirmation for the first part of the research question that this study set out to investigate, namely that human listeners do classify culture-specific melodies similarly to the computational model IDyOM. Based on the cognitive mechanisms statistical learning and probabilistic prediction, which are hypothesized to underlie the process of musical enculturation, it is likely that encultured listeners have mental representations that pertain to the statistical intervals of the in-culture musical style. This facilitates cognitive processing of the musical style and influences perceptual decision-making, as was required in the present study.

In a binary categorization task as the one used here, it is likely adequate to only be encultured in one of the musical styles represented. This is because there is no practical difference when a participant is sure that a melody foreign (bi-musical) and being sure that a melody is not native (mono-musical). However, it is likely advantageous for participants to be encultured it both musical styles when the melodies are ambiguous, as the majority of the melodies in this experiment were, since this would presumably mean that they have more comprehensive cognitive representations of the styles at hand, and this would facilitate a more accurate perceptual decision-making.

The results in this study could not confirm that implicit attitudes, as measured by the IAT, had any effect on performance in the categorization task – not as a direct predictor as in the third hypothesis, and neither when used as a proxy for foreignness in the fourth hypothesis.

The associative activations that are hypothesized to underlie implicit attitudes do not appear to influence the perceptual processing of musical material. It is likely that perceptual decision-making regarding musical style, as in this experiment's categorization task, requires reflective evaluations of the stimuli at hand, thereby activating propositional processes that invalidate the influence of implicit attitudes.

Regarding the computational model

If the IDyOM model is to be successful at simulating musical cognition in humans, one fundamental requirement is that the musical corpus by which the model has been trained represents the music by which listeners have been encultured. Here lies a potential limitation that must continuously be considered. Namely, it is open to inquiry how representative the training material is of the music that participants actively and passively listen to. The compositions used in this study were German and Chinese folksongs from the 19th and 20th centuries (Schaffrath, 1995). An underlying assumption of IDyOM is that pitch intervals in musical cultures are anchored in endurable musical traditions that remain more or less unaffected by shifts in popular music genres, which often are accompanied and driven by compelling extramusical elements. As an example, technological advancements in the mid 20th century paved the way for the introduction of electric instruments into popular music, like electric guitars and electric organs. This shift, in what was perceived as the vanguard of popular music in the 1960s, was arguably much more a general cultural shift than it was a purely musical one (Sweers, 2010). Another anecdote of preserved musical tradition despite cultural changes is Johan Sebastian Bach's (1685–1750) famous hymn "O Sacred Head, Now Wounded," whose main melody was reused in 1973 when Paul Simon (1941–) released the song "American Tune." Interestingly, Bach also borrowed part of the melody from a previous composition, namely Hans Leo Hassler's (1564–1612) "Passion Chorale" (Benninghof, 2007, p. 65).

Still, differences in statistical structures in terms of pitch interval and harmonic organization do exist between different musical genres belonging to the same broad musical tradition, as well as between music from different eras (Huron, 2006, p. 203). This potential issue of representation is assumed to be steadily increasing since out-of-culture music has become increasingly accessible through streaming services on the internet, like Spotify, Apple Music, and YouTube, resulting in increased exposure to out-of-culture music and possibly also decreased exposure to traditional in-culture music. Therefore, considerations regarding which compositions represent a musical style in a given computational model should be continuous, in order to achieve adequate correspondence with music that people listen to in everyday situations. When considering the results from this study, in particular from the first two hypotheses, it is evident that the participants' internal cognitive models do correspond with the computational model in terms of categorization accuracy (H1) and the perceived specificity of melodies (H2).

General musicality

As illustrated in Figure 9 in the section *Results*, participant no. 1 performed best in the categorization task and attained a categorization accuracy of 86.96%. This was above the computational model's accuracy, which is averaged to 80.43% between both musical styles. When conversing after the experiment, the participant declared to have had several years of musical training and self-identified as a musician. This may be an indication that musicality influences categorization performance. An extended paradigm, as the one initially planned, might be able to uncover potential relationships between the ability to categorize melodies according to style and different aspects of musicality. For example, results from the Musical Ear Test strongly correlate with musical expertise and can distinguish between musicians and non-musicians in experiments (Wallentin et al., 2010). It is plausible that highly musical individuals have more sophisticated cognitive representations of musical structures present in the styles in this study, and that this would prove to be an advantage when asked to classify ambiguous melodies. It is also plausible that there is a relationship between categorization ability and general musical engagement, rather than musical expertise. The Goldsmith's MSI questionnaire assesses self-reported musical engagement that has been shown to correlate well with results from listening tasks (Müllensiefen et al., 2014). The idea here is that more sophisticated representations of musical structures do not necessarily depend on musical expertise, but instead on previous exposure to said structures. In turn, the musical structures have become an integrated part of the individual's musical repertoire via the process of enculturation. In this way, an individual who has a broad musical knowledge but does not have notable musical expertise performance-wise may still perform well in the categorization task.

Considering this, different facets of musicality and musical engagement may be a better predictor of perceptual variability than implicit attitudes, as was the assumption in this study.

Facets of personality structure

Regarding the third hypothesis, there may be other factors mediating the relationship between implicit attitudes and musical engagement, such as facets of personality structure like *trait empathy* and *openness to experience* (Clarke et al., 2015; D. M. Greenberg, Müllensiefen, Lamb, & Rentfrow, 2015). In the case of empathy, there is some evidence to suggest that listeners who score high on trait empathy are more susceptible to show increased implicit attitudes toward a foreign culture after listening to music from said culture (Vuoskoski et al., 2017). It has been proposed that empathy as a psychological trait consists of two separate but interacting systems: *cognitive empathy*, referring to the ability to understand other's thoughts and feelings, and *affective empathy*, referring to the ability to respond with appropriate behavior and emotion to other's mental and emotional state (Shamay-Tsoory, Aharon-Peretz, & Perry, 2009). In the context of the present study, cognitive empathy may facilitate the understanding of another musical culture's underlying characteristics so that participants who score high on cognitive empathy would perform better in the categorization task. Affective empathy, on the other hand, is likely to facilitate changes in attitude towards the culture after engaging in musical activity, as demonstrated in other studies (Neto, Da Conceiçao Pinto, & Mullet, 2016; Neto et al., 2018; Vuoskoski et al., 2017).

Openness to experience (openness) is a personality trait from the Five-Factor Model of Personality (McCrae & John, 1992). It has been demonstrated as a predictor of musical sophistication (D. M. Greenberg et al., 2015). People who score high on trait openness are considered to be imaginative, have diverse interests, and are generally open to changes in their environment (McCrae & Ingraham, 1987). Furthermore, trait openness correlates positively with trait empathy (Barrio, Aluja, & García, 2006). It is therefore likely that individuals who score high on openness tend to engage more with foreign cultures, for example, by listening to out-of-culture music and thereby establishing familiarity with the statistical structures underlying the particular foreign musical style. Such engagement could plausibly translate into better performance in the categorization task and possibly also into lower scores on the IAT, based on stronger associations between out-of-culture artifacts and positive attributes.

Including tests in the study that measure empathy or openness to experience, such as the Interpersonal Reactivity Index (Davis, 1980) or a version of the personality test NEO-PI-3 (Costa & McCrae, 1992), could prove fruitful for future research.

Conclusion

This study investigated whether human listeners classify culture-specific melodies similarly to a computational model in a musical categorization task and whether implicit attitudes can predict performance. Four hypotheses were formulated in order to guide the investigation.

The results from the study demonstrate that there is no difference in the categorization accuracy of Chinese and German melodies between Danish participants and the computational model IDyOM. The results also demonstrate that Danish participants are significantly better at categorizing distinct melodies than they are at categorizing ambiguous melodies. These findings show that human listeners do classify culture-specific melodies similarly to IDyOM, indicating that the mechanisms underlying the IDyOM model can plausibly predict the effect that musical enculturation has on human listeners' ability to categorize melodies according to the cultural origin.

The results from this study also demonstrate that implicit attitudes, as measured by the implicit association test (IAT), can not convincingly predict performance in the musical categorization task. In addition to these findings, the results indicate no statistically significant relationship between IAT scores and the probability to categorize ambiguous out-of-culture melodies as in-culture melodies.

Overall, this study demonstrates that the IDyOM model is successful in simulating human listeners perception of musical style in terms of categorizing Chinese and German melodies, but it does not appear that variations in such perception can be predicted by implicit cultural attitudes, as measured by the IAT.

References

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia. A group study of adults afflicted with a music-specific disorder. *Brain*. https://doi.org/10.1093/brain/awf028

Badcock, P. B., Friston, K. J., Ramstead, M. J. D., Ploeger, A., & Hohwy, J. (2019). The hierarchically mechanistic mind: an evolutionary systems theory of the human brain, cognition, and behavior. *Cognitive, Affective and Behavioral Neuroscience*. https://doi.org/10.3758/s13415-019-00721-3

Baillet, S. (2011). Electromagnetic brain mapping using MEG & EEG. *Handbook of Social Neuroscience*.

Barrio, V. del, Aluja, A., & García, L. F. (2006). RELATIONSHIP BETWEEN EMPATHY AND THE BIG FIVE PERSONALITY TRAITS IN A SAMPLE OF SPANISH ADOLESCENTS. *Social Behavior and Personality: An International Journal*. https://doi.org/10.2224/sbp.2004.32.7.677

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*. https://doi.org/10.3758/BF03197629

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2008.0319

Bartlett, F. C., & Burt, C. (1933). REMEMBERING: A STUDY IN EXPERIMENTAL AND SOCIAL PSYCHOLOGY. *British Journal of Educational Psychology*. https://doi.org/10.1111/j.2044-8279.1933.tb02913.x

Bassili, J. N., & Brown, R. D. (2005). Implicit and explicit attitudes: Research, challenges, and theory. *The Handbook of Attitudes*.

Benninghof, J. (2007). *The Words and Music of Paul Simon*. (J. Perone, Ed.). Westport, Connecticut: Praeger.

Best, C. T. (1995). A direct realist view of cross-language speech perception. *Speech Perception and Linguistic Experience. Issues in Cross-Language Research*. https://doi.org/10.1016/0378-4266(91)90103-S

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. https://doi.org/10.1075/lllt.17.07bes

Bharucha, J. J. (1987). Music Cognition and Perceptual Facilitation: A Connectionist Framework. *Music Perception*. https://doi.org/10.2307/40285384

Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*. https://doi.org/10.1016/0010-0277(83)90003-3

Bhatnagar, G., Mehta, S., & Mitra, S. (2004). The WAV File Format. In *Introduction to Multimedia Systems*. https://doi.org/10.1016/b978-012500453-4/50011-0

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism*. https://doi.org/10.1017/S1366728908003477

Bigand, E., & Poulin-Charronnat, B. (2006). Are we 'experienced listeners'? A review of the musical capacities that do not depend on formal musical training. *Cognition*. https://doi.org/10.1016/j.cognition.2005.11.007

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In *Stereotype activation and inhibition: Advances in social cognition, Volume 11*.

Brattico, E., Brigitte Bogert, & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2013.00206

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*.

https://doi.org/10.1016/j.brainres.2006.08.023
Bregman, A. S., & McAdams, S. (1994). Auditory Scene Analysis: The Perceptual Organization of Sound. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.408434
Brent, M. R. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Sciences*. https://doi.org/10.1016/S1364-6613(99)01350-9
Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The 'Ticktock' of our Internal Clock: Direct Brain Evidence of Subjective Accents in Isochronous Sequences. *Psychological Science*. https://doi.org/10.1111/1467-9280.24441
Byros, V. (2009). *Foundations of tonality as situated cognition, 1730-1830 : An enquiry into the culture and cognition of Eighteenth-Century Tonality with Beethoven's Eroica Symphony as a Case Study*. *Foundations*.
Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology: A Journal of Research in Music Cognition*. https://doi.org/10.1037/h0094276
Chesterton, G. K. (1993). Bohemia's Claim for Freedom. *The Chesterton Review*. https://doi.org/10.5840/chesterton199319376
Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. *Physics of Life Reviews*. https://doi.org/10.1016/j.plrev.2015.09.001
Cosmides, L., & Tooby, J. (2000). Consider the Source: The Evolution of Adaptations for Decoupling and Metarepresentation. In D. Sperber (Ed.), *Metarepresentations: A Multidisciplinary Perspective* (p. 448). New York, New York: Oxford University Press.
Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI)*. *Odessa FL Psychological Assessment Resources*.
Crandall, C. S., & Eshleman, A. (2003). A Justification-suppression Model of the Expression and Experience of Prejudice. *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.129.3.414
Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/0898929042947919
Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*.
DeCoster, E. R. S. and J. (2002). Dual-Process Models in Social and Cognitive Psychology. *Personality and Social Psychology Review 2000, Vol. 4, No. 2, 108-131*. https://doi.org/10.1207/S15327957PSPR0402
DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*. https://doi.org/10.1038/nn1504
Demorest, S. M., Morrison, S. J., Beken, M. N., & Jungbluth, D. (2008). Lost in translation: An enculturation effect in music memory performance. *Music Perception*. 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*. https://doi.org/10.1525/MP.2016.33.5.590
Demorest, S. M., Morrison, S. J., Stambaugh, L. A., Beken, M., Richards, T. L., & Johnson, C. (2009). An fMRI investigation of the cultural specificity of music memory. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsp048
Deutsch, D. (2013). The Processing of Pitch Combinations. In *The Psychology of Music*. https://doi.org/10.1016/B978-0-12-381460-9.00007-9
Devine, P. G. (1989). Stereotypes and Prejudice: Their Automatic and Controlled Components. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.56.1.5

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*. https://doi.org/10.1016/0364-0213(90)90002-E
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*. https://doi.org/10.1037/0022-3514.69.6.1013
Festinger, L. (1957). A theory of cognitive dissonance / Leon Festinger. *A Theory of Cognitive Dissonance*.
Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2005.1622
Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn2787
Furl, N., Kumar, S., Alter, K., Durrant, S., Shawe-Taylor, J., & Griffiths, T. D. (2011). Neural prediction of higher-order auditory sequence statistics. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2010.10.038
Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.132.5.692
Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*. https://doi.org/10.1037/0003-066X.46.2.107
Gjerdingen, R. O., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research*. https://doi.org/10.1080/09298210802479268
Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*. https://doi.org/10.1016/S1364-6613(00)01662-4
Gobet, F., & Simon, H. A. (1996). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory and Cognition*. https://doi.org/10.3758/BF03200937
Greenberg, D. M., Müllensiefen, D., Lamb, M. E., & Rentfrow, P. J. (2015). Personality predicts musical sophistication. *Journal of Research in Personality*. https://doi.org/10.1016/j.jrp.2015.06.002
Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-Response Characteristic of Auditory Observers Detecting Signals of a Single Frequency in Noise: The Probe-Signal Method. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.1911290
Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.74.6.1464
Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). 'Understanding and using the Implicit Association Test: I. An improved scoring algorithm': Correction to Greenwald et al. (2003). *Journal of Personality and Social Psychology*. https://doi.org/10.1037/h0087889
Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.90.1.1
Grimaldi, M., Sisinni, B., Fivela, B. G., Invitto, S., Resta, D., Alku, P., & Brattico, E. (2014). Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: A behavioral and ERP study. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2014.00279
Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*. 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 of the United States of America*. https://doi.org/10.1073/pnas.0504254102

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*. https://doi.org/10.3389/fnins.2018.00218

Henrich, J., & McElreath, R. (2003). The Evolution of Cultural Evolution. *Evolutionary Anthropology*. https://doi.org/10.1002/evan.10110

Holland, J. H., Holyoak, K. J., Nisbett, R. E., Thagard, P. R., & Smoliar, S. W. (2008). Induction: Processes of Inference, Learning, and Discovery. *IEEE Expert*. https://doi.org/10.1109/mex.1987.4307100

Huber, D. (2012). Midi. In *Audio Engineering Explained*. https://doi.org/10.4324/9780240812748-26

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

Huron, D. (2008). Science & Music: Lost in music. *Nature*. https://doi.org/10.1038/453456a

Jackendoff, R. (2002). *Foundations of Language*. New York, New York: Oxford University Press.

Judd, C. M., & Lusk, C. M. (1984). Knowledge structures and evaluative judgments: Effects of structural variables on judgmental extremity. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.46.6.1193

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X08005293

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2004.10.007

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.110.3.499

Koelsch, S. (2012). Brain and Music. *Brain and Music*.

Koelsch, S., & Friederici, A. D. (2003). Toward the Neural Basis of Processing Structure in Music. *Annals of the New York Academy of Sciences*. https://doi.org/10.1196/annals.1284.002

Koelsch, S., Gunter, T. C., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: An ERP study. *NeuroReport*. https://doi.org/10.1097/00001756-200105250-00019

Koelsch, S., Gunter, T., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/089892905774597290

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.2007.00517.x

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*. https://doi.org/10.1038/nn1197

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2005.10.001

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*. https://doi.org/10.1207/S15327965PL100201

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition*. https://doi.org/10.3758/BF03202445

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*. https://doi.org/10.1037/0033-295X.89.4.334

Kuhl, P. K. (1993). Innate Predispositions and the Effects of Experience in Speech Perception: The Native Language Magnet Theory. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. https://doi.org/10.1007/978-94-015-8234-6_22
Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn1533
Kuhl, P. K., Conboy, B., Padden, D., Nelson, T., & Pruitt, J. (2005). Early Speech Perception and Later Language Development: Implications for the 'Critical Period'. *Language Learning and Development*. https://doi.org/10.1207/s15473341lld0103&4_2
Kumar, S., Joseph, S., Gander, P. E., Barascud, N., Halpern, A. R., & Griffiths, T. D. (2016). A brain system for auditory working memory. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.4341-14.2016
Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2007.11.004
Large, E. W. (2017). Music from the air to the brain and body. In R. Ashley & R. Timmers (Eds.), *The Routledge Companion to Music Cognition*. New York, New York: Taylor & Francis.
Lerdahl, F., & Jackendoff, R. (1983). An Overview of Hierarchical Structure in Music. *Music Perception*, *1*(2), 229–252. https://doi.org/10.2307/40285257
Longuet-Higgins, H. C., & Lee, C. S. (1984). The Rhythmic Interpretation of Monophonic Music. *Music Perception*. https://doi.org/10.2307/40285271
Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception and Psychophysics*. https://doi.org/10.3758/BF03193946
Loui, P., & Wessel, D. (2008). Learning and liking an artificial musical system: Effects of set size and repeated exposure. *Musicae Scientiae*. https://doi.org/10.1177/102986490801200202
Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*. https://doi.org/10.3758/BF03211696
MacKay, D. (2004). Information Theory, Inference, and Learning Algorithms. *IEEE Transactions on Information Theory*. https://doi.org/10.1109/TIT.2004.834752
Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*. https://doi.org/10.1016/0010-0277(80)90015-3
Martin, D. W. (2008). *Doing psychology experiments*. *PsycCRITIQUES*. https://doi.org/10.1037/023365
Matsunaga, R., Yokosawa, K., & Abe, J. ichi. (2012). Magnetoencephalography evidence for different brain subregions serving two musical cultures. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2012.10.002
McCrae, R. R., & Ingraham, L. J. (1987). Creativity, Divergent Thinking, and Openness to Experience. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.52.6.1258
McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x
Medalla, M., & Barbas, H. (2014). Specialized prefrontal 'auditory fields': Organization of primate prefrontal-temporal pathways. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2014.00077
Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2007.07.034
Mitchell, A., Romano, G. H., Groisman, B., Yona, A., Dekel, E., Kupiec, M., … Pilpel, Y.

(2009). Adaptive prediction of environmental changes by microorganisms. *Nature*. https://doi.org/10.1038/nature08112

Moelants, D., & Leman, M. (1997). Music and Schema Theory: Cognitive Foundations of Systematic Musicology. *Revue Belge de Musicologie / Belgisch Tijdschrift Voor Muziekwetenschap*. https://doi.org/10.2307/3687199

Morrison, S. J., & Demorest, S. M. (2016). Quantifying Culture: The Cultural Distance Hypothesis of Melodic Expectancy. In *The Oxford Handbook of Cultural Neuroscience*.

Morrison, S. J., Demorest, S. M., Aylward, E. H., Cramer, S. C., & Maravilla, K. R. (2003). fMRI investigation of cross-cultural music comprehension. *NeuroImage*. https://doi.org/10.1016/S1053-8119(03)00300-8

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). Measuring the facets of musicality: The Goldsmiths Musical Sophistication Index (Gold-MSI). *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2013.07.081

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x

Narmour, E. (1991). The Top-Down and Bottom-Up Systems of Musical Implication: Building on Meyer's Theory of Emotional Syntax. *Music Perception*. https://doi.org/10.2307/40286156

Neto, F., Da Conceiçao Pinto, M., & Mullet, E. (2016). Can music reduce anti-dark-skin prejudice? A test of a cross-cultural musical education programme. *Psychology of Music*. https://doi.org/10.1177/0305735614568882

Neto, F., Pinto, M. da C., & Mullet, E. (2018). Can music reduce national prejudice? A test of a cross-cultural musical education programme. *Psychology of Music*. https://doi.org/10.1177/0305735618774867

Nosek, B. A., Bar-Anan, Y., Sriram, N., Axt, J., & Greenwald, A. G. (2014). Understanding and using the brief implicit association test: Recommended scoring procedures. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0110938

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*. https://doi.org/10.1038/381607a0

Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2013.05.010

Palmer, C., & Krumhansl, C. L. (1987). Independent Temporal and Pitch Structures in Determination of Musical Phrases. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/0096-1523.13.1.116

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*. https://doi.org/10.1038/nn1082

Patel, A. D. (2012). *Music, Language, and the Brain*. *Music, Language, and the Brain*. https://doi.org/10.1093/acprof:oso/9780195123753.001.0001

Pearce, M. T. (2005). *THE CONSTRUCTION AND EVALUATION OF STATISTICAL MODELS OF MELODIC STRUCTURE IN MUSIC PERCEPTION AND COMPOSITION*. *December*. City University, London. Retrieved from https://openaccess.city.ac.uk/id/eprint/8459/

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

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010). The role of expectation and probabilistic learning in auditory boundary perception: A model comparison. *Perception*. https://doi.org/10.1068/p6507

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2009.12.019

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*. https://doi.org/10.1525/mp.2006.23.5.377

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2006.03.006

Pham, H. N. A., & Triantaphyllou, E. (2008). The impact of overfitting and overgeneralization on the classification accuracy in data mining. In *Soft Computing for Knowledge Discovery and Data Mining*. https://doi.org/10.1007/978-0-387-69935-6_16

Pickles, J. O. (2012). *Introduction to the Physiology of Hearing. Introduction to the Physiology of Hearing.*

Picton, T. W., Durieux-Smith, A., & Moran, L. M. (1994). Recording auditory brainstem responses from infants. *International Journal of Pediatric Otorhinolaryngology*. https://doi.org/10.1016/0165-5876(94)90001-9

Platel, H., Baron, J. C., Desgranges, B., Bernard, F., & Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *NeuroImage*. https://doi.org/10.1016/S1053-8119(03)00287-8

Poulin-Charronnat, B. (2004). Effects of a Change in Instrumentation on the Recognition of Musical Materials. *Music Perception: An Interdisciplinary Journal*. https://doi.org/10.1525/mp.2004.22.2.239

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*. https://doi.org/10.1038/nn.2331

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2011.12.010

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*. https://doi.org/10.1016/S0010-0277(98)00075-4

Schaffrath, H. (1995). The Essen Folksong Collection in Kern Format. *Assisted Research in the Humanities, Menlo Park, CA: Center for Computer*. Retrieved from https://kern.humdrum.org/

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/j.1749-6632.2012.06447.x

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*. https://doi.org/10.1093/brain/awn279

Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature*. https://doi.org/10.1038/nature04485

Smith, E. R. (1996). What Do Connectionism and Social Psychology Offer Each Other? *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.70.5.893

Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*. https://doi.org/10.1016/0013-4694(75)90047-4

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.2006.18.8.1380

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*. https://doi.org/10.1207/s15327957pspr0803_1

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2009.06.003

Sweers, B. (2010). *Electric Folk: The Changing Face of English Traditional Music*. *Electric Folk: The Changing Face of English Traditional Music*. https://doi.org/10.1093/acprof:oso/9780195174786.001.0001

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

Tooby, J., & Cosmides, L. (2015). The Theoretical Foundations of Evolutionary Psychology THE EMERGENCE OF EVOLUTIONARY PSYCHOLOGY: WHAT IS AT STAKE? *Foundations*.

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*. https://doi.org/10.1016/j.psychres.2008.05.006

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*. https://doi.org/10.1159/000356237

van der Weij, B., Pearce, M. T., & Honing, H. (2017). A probabilistic model of meter perception: Simulating enculturation. *Frontiers in Psychology*, *8*(MAY), 1–18. https://doi.org/10.3389/fpsyg.2017.00824

Volgsten, U. (2014). Music, Culture, Politics: Communicating Identity, Authenticity and Quality in the 21st Century. *Nordisk Kulturpolitisk Tidskrift*.

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

Vuust, P., Dietz, M. J., Witek, M., & Kringelbach, M. L. (2018). Now you hear it: A predictive coding model for understanding rhythmic incongruity. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.13622

Waeber, J. (2007). Searching for the plot: Charles-Valentin Alkan's Souvenirs: Trois morceaux dans le genre pathétique, Op. 15. *Journal of the Royal Musical Association*. https://doi.org/10.1093/jrma/fkm005

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*. https://doi.org/10.1016/j.lindif.2010.02.004

Wang, Y., Kuhl, P. K., Chen, C., & Dong, Q. (2009). Sustained and transient language control in the bilingual brain. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2008.12.055

Welford, A. T. (1980). Choice reaction time: basic concepts. In *Reaction Times*.

White, C. W. M. (2014). Changing styles, changing corpora, changing tonal models. *Music Perception*. https://doi.org/10.1525/MP.2014.31.3.244

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0094446

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology : CB*. https://doi.org/10.1016/s0960-9822(01)00432-8

Wolpert, R. S. (1990). Recognition of Melody, Harmonic Accompaniment, and Instrumentation: Musicians vs. Nonmusicians. *Music Perception*. https://doi.org/10.2307/40285487

Wong, P. C. M., Roy, A. K., & Margulis, E. H. (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

Zatorre, R. J. (2015). Musical pleasure and reward: Mechanisms and dysfunction. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.12677

Zatorrea, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1301228110

Test for normal distribution

The data are a little skewed and kurtotic, but it doesn't differ significantly from normality.

Shapiro-Wilk test: Null hypothesis is that the data are normally distributed. P-value is 0.689, meaning the null hypothesis will be kept.

It can be concluded from this that the data from the categorization test (categorization accuracy) are approximately normally distributed, and therefore parametric test can be applied.

Descriptives

			Statistic	Std. Error
Accuracy	Mean		77.1740	1.59593
	95% Confidence Interval for Mean	Lower Bound	73.5638	
		Upper Bound	80.7842	
	5% Trimmed Mean		77.0528	
	Median		76.0900	
	Variance		25.470	
	Std. Deviation		5.04677	
	Minimum		69.57	
	Maximum		86.96	
	Range		17.39	
	Interquartile Range		7.07	
	Skewness		.634	.687
	Kurtosis		.319	1.334

Z-values:

Skewness: (0.634/0.687) = 0.923

Kurtosis: (0.319/1.334) = 0.239

Tests of Normality

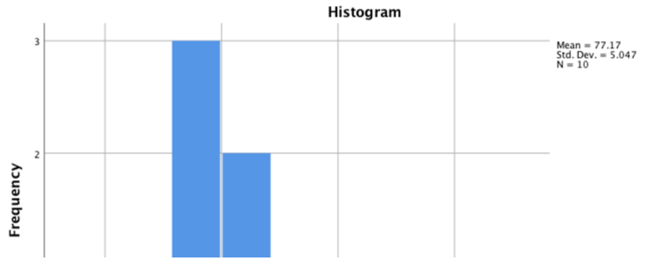
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Accuracy	.185	10	.200*	.952	10	.689

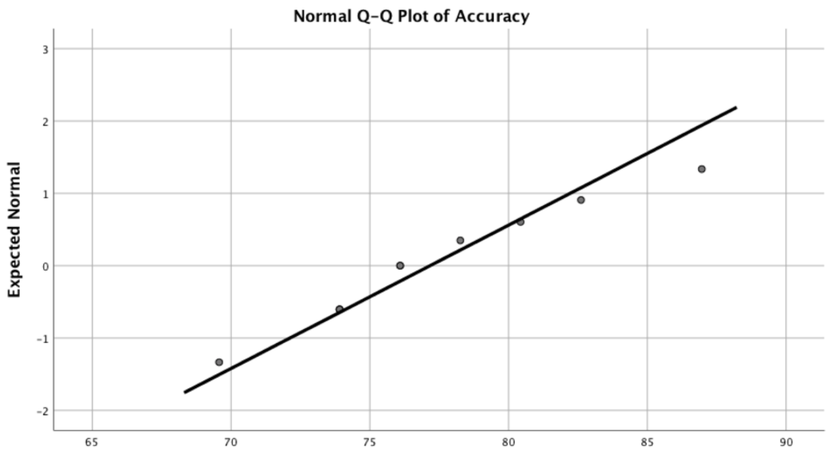
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

P-value: 0.689

Accuracy





Participant data and analyses output

Participant data

Subject	Gender	Age	Categor. accuracy	Distinct accuracy	Ambiguous accuracy	IAT score	IAT group
1	M	28	86.96	100	85	-0.57	High
2	M	21	78.26	100	75	-0.43	Low
3	F	22	73.91	100	70	-0.89	High
4	F	24	76.09	100	72.5	-0.56	High
5	M	31	73.91	100	70	-0.37	Low
6	M	25	76.09	100	72.5	-1.2	High
7	F	23	69.57	50	72.5	-0.54	Low
8	M	28	73.91	100	70	-0.54	Low
9	F	29	82.61	100	80	-0.6	High
10	F	24	80.43	100	77.5	-0.16	Low
Model			80.43	100	77.5		
			Mean	**Mean**	**Mean**		
		Low group	75.22	90	73		
		High group	79.13	100	76		
		Combined	77.174	95	74.5		
		St. deviation	4.23	22.36515191	3.259601203		
			5.45	0	6.274950199		
		STDEV combin	5.046768383	15.81455058	4.97214463		

Hypothesis 1

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	10	77.1740	5.04677	1.59593

One-Sample Test

	Test Value = 80.43					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Accuracy	-2.040	9	.072	-3.25600	-6.8662	.3542

Hypothesis 2

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Distinct accuracy	95.0000	10	15.81139	5.00000
	Ambiguous accuracy	74.5000	10	4.97214	1.57233

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Distinct accuracy & Ambiguous accuracy	10	.141	.697

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Distinct accuracy - Ambiguous accuracy	20.50000	15.89025	5.02494	9.13280	31.86720	4.080	9	.003

Hypothesis 3

T-Test

Group Statistics

	High/low	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	0	5	75.2160	4.23498	1.89394
	1	5	79.1320	5.45781	2.44081

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Accuracy	Equal variances assumed	.872	.378	-1.268	8	.241	-3.91600	3.08943	-11.04023	3.20823
	Equal variances not assumed			-1.268	7.535	.243	-3.91600	3.08943	-11.11745	3.28545

Hypothesis 4

Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
Foreign as native	55.9180	14.88114	10
IAT d-score	-.5860	.28426	10

Correlations

		Foreign as native	IAT d-score
Foreign as native	Pearson Correlation	1	.352
	Sig. (1-tailed)		.159
	N	10	10
IAT d-score	Pearson Correlation	.352	1
	Sig. (1-tailed)	.159	
	N	10	10

Power analysis (hypothesis 2)

t tests - Means: Difference between two independent means (two groups)

```
Analysis:   A priori: Compute required sample size
Input:      Tail(s)                          =   One
            Effect size d                    =   0.8
            α err prob                       =   0.05
            Power (1-β err prob)             =   0.95
            Allocation ratio N2/N1           =   1
Output:     Noncentrality parameter δ        =   3.3466401
            Critical t                       =   1.6675723
            Df                               =   68
            Sample size group 1              =   35
            Sample size group 2              =   35
            Total sample size                =   70
            Actual power                     =   0.9523628
```

Flyer

Undersøgelsesdeltagere søges

Er du universitetsstuderende og har du lyst til at deltage i et pilotstudie der omhandler musik og psykologi? Så vil jeg meget gerne høre fra dig!
Du vil blive bedt om at udføre to computerbaserede tests, som sammenlagt tager ca. 30 min at færdiggøre. Efter det vil du blive tilbudt kaffe og kage, som tak for hjælpen.

Undersøgelsen foregår på XXXX i Aalborg centrum.

Kunne du tænke dig at deltage er du velkommen til at kontakte mig på
mail: XXXX@student.aau.dk, eller telefon: XXXX.