Title
Children infer the behavioral contexts of unfamiliar foreign songs

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Hilton, Courtney B
Crowley-de Thierry, Liam
Yan, Rachel
et al.

Publication Date
2021

Peer reviewed

Children infer the behavioral contexts of unfamiliar foreign songs

Liam Crowley-de Thierry*
(crowleliam@myvuw.ac.nz)
School of Psychology, Victoria University Wellington
Wellington, New Zealand

Courtney B. Hilton*
(courtneyhilton@g.harvard.edu)
Department of Psychology, Harvard University,
Cambridge, MA 02138 USA

Rachel Yan (ryan49@smith.edu)
Department of Psychology, Smith College,
Northampton, MA 01063 USA

Alia Martin (alia.martin@vuw.ac.nz)
School of Psychology, Victoria University Wellington
Wellington, New Zealand

Samuel Mehr (sam@wjh.harvard.edu)
Department of Psychology, Harvard University,
Cambridge, MA 02138 USA

* These authors contributed equally

Abstract

Humans readily make inferences about the behavioral context of the music they hear. These inferences tend to be accurate, even if the songs are in foreign languages or unfamiliar musical idioms: upon hearing a Blackfoot lullaby, a Korean listener with no experience of Blackfoot music, language, or broader culture is far more likely to judge the music's function as "to soothe a baby" than as "for dancing". Are such inferences shaped by musical exposure or does the human mind naturally detect certain links between musical form and function? Children's developing experiences with music provide a clear test of this question. We studied musical inferences in a large sample of children (N = 2,418), who heard dance, lullaby, and healing songs from 70 world cultures and were tasked with guessing the original behavioral context in which each was performed. We found little evidence for the effect of experience on musical inferences: children reliably inferred the original behavioral contexts of unfamiliar foreign songs, with only minimal improvement in performance from the youngest (age 3 or younger) to the oldest (age 12) participants. Children's inferences tightly correlated with those of adults for the same songs, as collected from a similar massive online experiment (N = 85,068). Moreover, the same acoustic features explained variability in both children's and adults' inferences. These findings imply that accurate inferences about the behavioral contexts of music, driven by links between form and function in music across cultures, do not require extensive musical experience.

Keywords: Music; Development; Form-function links; Social cognition

Introduction

Music is a ubiquitous part of human culture whose presence plays various functional roles in our day-to-day lives (Mehr et al., 2019; Trehub et al., 2015): We sing songs to entertain each other, to find spiritual connection, to tell stories, to regulate our emotions, or just to pass the time. Whereas some musical contexts are idiosyncratic, like how complex geographical knowledge is culturally encoded in song by some groups of Indigenous Australians (Norris & Harney, 2014), others are recurrent across societies. These latter contexts tend to also be accompanied by a recurrent set of features. For example, lullabies tend to have slow tempos and a narrow dynamic and pitch range (Bergeson & Trehub, 1999; Trehub et al., 1993b; Trehub & Trainor, 1998), and this is found not only for western lullabies but also those in a globally representative sample of small-scale societies (Mehr et al., 2019). That music appears universally in conjunction with a given behavioral context (e.g., infant-care), and that examples of music in that context share particular musical features (e.g., slow tempos), implies a link between form and function in music production.

Adults are sensitive to this link, in that they reliably distinguish song forms from one another even when examples are unfamiliar and foreign. This has been demonstrated in experiments where naïve listeners are asked to identify the context in which a song was originally used, purely on the basis of the sounds it contains. Indeed, they successfully identify lullabies when paired with love songs as foils (Trehub et al., 1993a) and distinguish lullabies, dance songs, healing songs, and love songs from one another (with weakest performance at identifying love songs) (Mehr et al., 2018, 2019).

How do adults come to be sensitive to these musical form-function links? Given the ubiquity and richness of musical experience over the lifespan, and especially in the younger years (Mendoza & Fausey, 2021), it is plausible that this sensitivity is shaped by direct musical experience. This seems all the more plausible given the detailed nature of people's long-term memories for music (Krumhansl, 2010) and known importance of implicit associative learning in musical knowledge more generally (Rohrmeier & Rebuschat, 2012).

Two forms of evidence point to an alternative explanation, however. First, across species, systematic form-function pairings are found in vocalizations throughout the animal kingdom (Morton, 1977), possibly reflecting innate aspects of vocal signaling (Darwin, 1871; Mehr et al., 2020). Indeed, emotional vocalizations are not only cross-culturally intelligible in humans (Scherer et al., 2001; including in music: Balkwill & Thompson, 1999), but also between species (Filippi et al., 2017). Form-function inferences in song might thus be grounded in these instinctive and evolutionarily ancient signaling mechanisms.

Secondly, infants, who have far less exposure to music than do adults, relax more in response to lullabies than non-lullabies, even when the particular examples are culturally unfamiliar (Bainbridge et al., 2020). While this does not demonstrate an ability to identify a lullaby as such, it does suggest an innate sensitivity to its physiological effects, which may in turn come to act as an internal cue for higher-level cognitive inferences. Together, these findings suggest an innate or early-appearing sensitivity to form and function in music, which could subsequently interact with musical experience to produce adult-level musical inferences.

Here, we test between these explanations by studying musical inferences across early and middle childhood, so as to measure whether and how these inferences change as children's musical experience grows. Children participated in a citizen-science experiment (readers can try it with their children at https://themusiclab.org) where they listened to examples of lullabies, dance songs, and healing songs from 70 mostly small-scale societies, drawn from the Natural History of Song Discography (Mehr et al., 2019). For each example, they guessed the original behavioral context of the song.

To examine the degree to which musical experience affected their intuitions, we conducted a series of preregistered and exploratory analyses, asking (1) whether children are sensitive to form and function in music; (2) how this sensitivity changes over the course of childhood; (3) how children's inferences compare to those of adults; and (4) whether and how children's inferences are predictable from the acoustical features of the music.

Methods

Participants

Children visited the citizen science website https://themusiclab.org to participate. We posted links to the experiment on social media, which were spread via organic sharing, and we also advertised the experiment on https://childrenhelpingscience.com, a website disseminating information on web-based research for children. Recruitment was/is open-ended; at the time of analysis, we had complete data from 5,774 children with reported ages from "3 or under" to "17 or older." Given our interest in early development, we chose to analyze only data from the youngest 10 age bands (ages "3 or under" to "12"). We excluded older children (n = 3,502); children whose parents indicated that they had assisted their child during the experiment (n = 658); children who indicated they had played the game previously (n = 499); children with known hearing impairments (n = 350); and children who were missing data from the test phase (n = 37).

This left 2,418 participants for analysis (1,031 male, 1,292 female, 95 other) representing 107 native languages and 110 countries.

No compensation was given, reducing the likelihood that adults were posing as children, but as an additional check, we examined participants' response times during the study. Consistent with prior work showing that response time in perceptual and cognitive tasks decreases over the course of childhood (Kiselev et al., 2009), response times during the task decreased reliably as a function of age ($F(1, 2{,}416) = 205.54$, $p < .001$, $R^2 = 0.078$; SI Figure 1). More generally, there is growing evidence that online and lab-based data collection produce similar results, with limited evidence for reduced reliability in online cohorts (Hartshorne et al., 2019).

Stimuli

Song excerpts were drawn from the Natural History of Song Discography (Mehr et al., 2019), a corpus of vocal music collected from 86 predominantly small-scale societies across the world, including hunter-gatherers, pastoralists, and subsistence farmers. Each recording in this corpus is performed in one of four behavioral contexts: dance, healing, lullaby, and love. The song recordings in this corpus are selected and categorized based on the ethnographic records pertaining to each song, collated within the NHS Ethnography (see Mehr et al., 2018, 2019). Each song is described by a text excerpt which summarizes the song performance, its use within a given society or both. For example, "Urai turuk bilou", a dance song from the Mentawai people located off the western coast of Sumatra in Indonesia is described by dancers remaining still during singing, or stamping in rhythm while making animal noises.

Because love songs were ambiguously detected in some prior work (Mehr et al., 2018), and given the relative difficulty of explaining this category to children, we omitted love songs from the experiment and studied only the remaining three categories (dance, lullaby, and healing). This left 88 songs from 70 mostly small-scale societies, sung in 30 languages and originating from locations corresponding to 46 modern countries (see SI Figure 2). The 88 songs consisted of 30 dance, 28 healing, and 30 lullaby.

Procedure

Parents were instructed to begin the game without the child present, so that they could provide us with demographic information and become oriented to the interface. This information included whether the child had previously participated in the experiment; the child's age, gender, country of residence, native language, any known hearing impairments; and whether or not the child was wearing headphones. We also asked about the frequency with which the child heard singing and recorded music in the home.

Parents then turned the experiment over to their children, who interacted with an animated character ("Susie the Star"). First, in a training phase, Susie played children a likely-to-be-familiar song (the "Happy Birthday" song) and asked them to identify what they thought it was used for, by selecting one of three choices: "singing for bath time", "singing for school assembly" or "singing for celebrating a birthday"). Once children correctly identified the song's behavioral context, Susie then presented the three song categories that children would be tested on in the experiment (lullaby, "for putting a baby to sleep"; healing, "to make a sick person feel better"; dance, "singing for dancing"). A static illustration accompanied each song category's verbal description, so as to facilitate responses for children too young to read.

Susie then played the children a counterbalanced set of six songs (two per song type, 14s each) drawn randomly from the Natural History of Song Discography and presented in a randomized order. Children were unable to advance to the next trial until the song excerpt played in its entirety. After each song, Susie asked "What do you think that song was for?" and children made a guess by clicking or tapping on an illustration corresponding with the song type. After each song, we also asked children to rate how much they liked the song, for use in a separate experiment. Encouraging feedback was always given after each trial (e.g., "Good job!"), with an arbitrary number of "points" awarded (+20 points when correct and +15 points when incorrect).

Last, parents answered debriefing questions. The questions confirmed whether the child had worn headphones during the experiment (if, at the beginning of the study, the parent had stated the child would wear headphones), if the parent had assisted the child during the training portion of the experiment, and if the parent had assisted the child during the test questions.

Pilot study and preregistration

We ran a pilot version of the experiment with 500 children (50 children in each age group, range 3-12; 221 male, 267 female, 12 other) to provide a dataset for exploratory analyses. Based on these data, and our broader theoretical interests motivated in the introduction section of this paper, we preregistered three confirmatory hypotheses to be tested in a larger sample (see the preregistration at https://osf.io/56zne): in the full cohort of children, accurate identification of behavioral contexts overall and in all three song types; a statistically significant but practically non-significant effect of age on accuracy ($R^2 < .05$); small or null associations of musical exposure in the home on accuracy.

We also pre-specified four other hypotheses that were not investigated in the exploratory sample. Two of these are studied in this paper: predicted correlations between children's intuitions about the songs and adult's intuitions (with a sampled expanded from our previous study in Mehr et al., 2019); and a prediction that the musical features of the songs that are predictive of children's intuitions about them, within a given song type, will correspond with musical features previously identified as universally associated with that song type (from Mehr et al., 2019). We leave the remaining two hypotheses for future research.

Results

Children's musical inferences are accurate

Children accurately inferred the behavioral context of the songs they listened to at a rate significantly above chance level of 33.3% (Figure 1a; M = 52.3%, SD = 20.8%, d' = 0.82), as they did in exploratory results, and confirming the first preregistered hypothesis.

Notably, children did not use the three response options evenly: they guessed "dance" most frequently (42.6%), followed by "healing" (35.3%), and least often guessed "lullaby" (22.2%). As such, we computed *d*-prime scores for each of the three song types for each of the participants, to assess discrimination independently from response bias. Dance songs were the most reliably discriminated (d' = 0.88, 95% *CI* [0.84, 0.91], $t(2417) = 50.57$, $p < .001$, one-sample two-tailed t-test), followed closely by lullabies (d' = 0.85, 95% *CI* [0.82, 0.88], $t(2417) = 49.51$, $p < .001$). Healing songs were the least reliably discriminated but still robustly above chance (d' = 0.28, 95% *CI* [0.24, 0.32], $t(2417)$ = 15.04, $p < .001$).

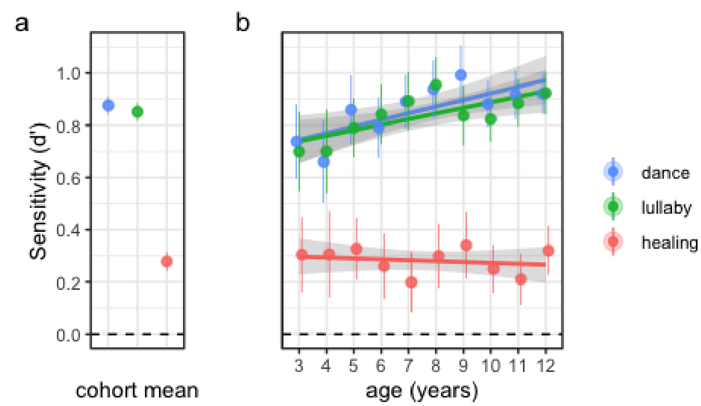

Figure 1: Accurate identification of all three song types, cohort-wide and as a function of age. **a**, Mean *d*-prime scores across all children show above-chance identification of each song type, independent of their response bias. **b**, Accuracy increases only modestly from the youngest to the oldest children, and only for lullabies and dance songs. In both panels, the circles indicate mean *d*-prime scores and the error bars indicate the 95% confidence intervals. In panel **b**, the three thick lines depict a linear regression for each song type and the shaded regions represent the 95% confidence intervals from each regression.

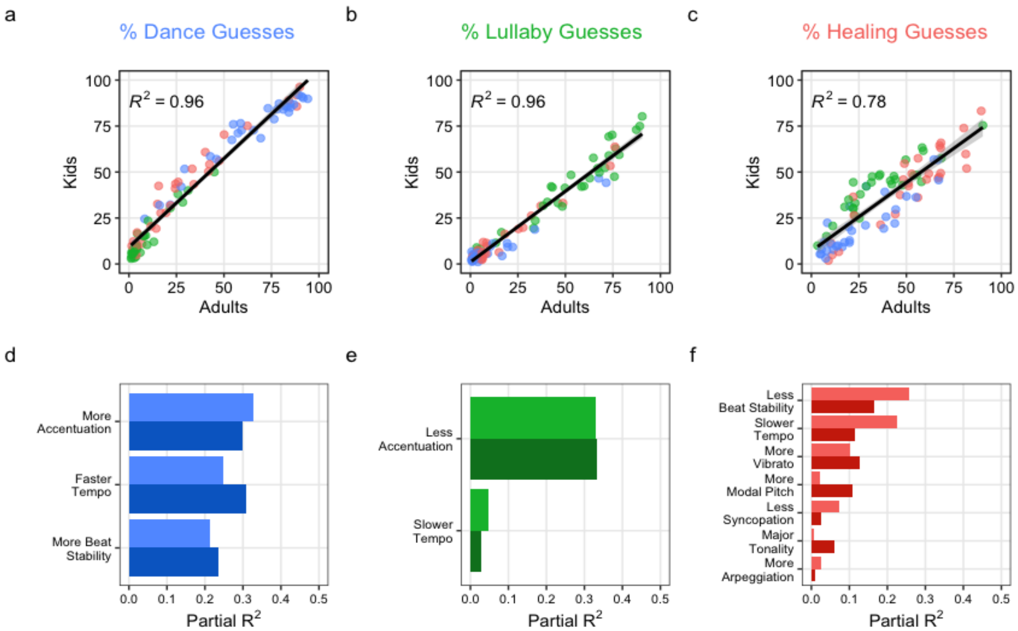

Figure 2: Children and adults make similar musical inferences that are driven by the same acoustical features. The scatterplots (**a-c**) show the tight correlations between children's and adults' musical inferences. Each point represents average percent guesses that was (**a**) "for dancing"; (**b**) "for putting a baby to sleep"; and (**c**) "to make a sick person feel better"; the songs' *actual* behavioral contexts are color-coded, with dance songs in blue, lullabies in green, and healing songs in red. The lines depict simple linear regressions and the gray shaded areas show the 95% confidence intervals from each regression. The bar plots (**d-f**) show the similar amounts of variance (partial-$R2$) in childrens' (lighter bars) and adults' (darker bars) guesses that is explained by musical features selected via LASSO regularization, for each of the three song types, computed from multiple regressions.

Musical inferences improve only slightly with age

To measure the degree to which these inferences change through development, we fit a simple linear regression predicting children's average accuracy from age, finding a small but significant effect ($F(1, 2{,}416) = 7.38$, $p = 0.007$, $R^2 = 0.003$), consistent with our exploratory analyses and preregistered hypothesis (i.e., a small, or practically non-significant increase in accuracy). As a more rigorous test (which was not preregistered), we ran a mixed-effects model on participant-wise *d*-prime scores, with random intercepts for participants and random intercepts and slopes for song types. In this model, the effect of age did not reach statistical significance ($F(1, 235.63) = 3.73$, $p = 0.055$).

Musical inferences are unrelated to children's home musical environment

We tested the relation between parent-child musical interactions and children's musical inferences. Consistent with the preregistered hypothesis, we found no evidence for relations between *d*-prime scores and the frequency of parental singing (SI Figure 4a; $F(1, 2{,}414) = 0$, $p > 0.99$), or the frequency of recorded music (SI Figure 4b; $F(1, 2{,}414) = 0.06$, $p = 0.801$).

Musical inferences are highly similar between children and adults

With practically no change in accuracy from early childhood to early adolescence, might children's musical inferences be similar to those of adults? We used all available data from the "World Music Quiz" on https://themusiclab.org, a similar experiment for adults. Data were available from 85,068 participants who were older than 18 years old (13,157 male, 6,643 female, 205 other, for 65,063 gender information was not available; mean age = 31.3, SD = 12.9, range: 18-99; n.b., a portion of these data were previously analyzed in Mehr et al. (2019), but not in relation to children's performance on the task). For both adults and children, we computed the song-wise proportion of guesses for each song type (i.e., a measure of how strongly each song cued a given behavioral context) and then regressed these scores on each other.

In all cases, children's inferences were highly predictive of adults', with R^2 values approaching 1 (Figure 3a-c; dance: $F(1, 86) = 1{,}967.28$, $R^2 = 0.96$; lullaby: $F(1, 86) = 1{,}940.38$, $R^2 = 0.96$; healing: $F(1, 86) = 297.49$, $R^2 = 0.78$; all $p < .001$). This relationship was also robust when inferences were reduced from a continuous score to a discrete ranking of the songs by guessing percentage. We quantified the strength of this relationship using Kendall's rank correlation coefficient $\boldsymbol{\tau}$: a more conservative, non-parametric measure of correlation. Again, there was a strong positive correlation between children's and adults' inferences for all song types (dance: τ = 0.874; lullaby: τ = 0.787; healing: τ = 0.706; all p <.001).

Child and adult musical inferences are driven by the same musical features

To investigate what might drive children's musical inferences, we analyzed which musical features were predictive of their guesses, within each song; and whether these musical features corresponded with those that were predictive of adults' guesses. We began with the set of 36 musical features (expert annotations and transcription variables) previously studied in Mehr et al. (2019). To reduce overfitting, we selected variables via LASSO regularization.

The model was trained to predict the child and adult inferences for each song on the basis of the 36 musical features. The key hyperparameter of the model (lambda), which determines how many features are selected, was chosen on the basis of a 10-fold cross-validation procedure across random subsets of the data, repeated 100 times for each song function to maximize robustness. The model mean squared error was averaged over these iterations, and we selected the highest lambda value whose model error was within one standard error of the minimum. This procedure yielded a small number of musical features predictive of each song-type guess; we then discarded those with coefficients below an arbitrary threshold of 0.01, as a form of secondary regularization (to err on the side of analyzing a smaller set of features).

We then built multiple regression models predicting the percentage of song-wise guesses in each category from the selected musical features. For all three song types, these features explained substantial variability in listener inferences (dance: $F(1, 84) = 52.66$, $p < .001$, $R^2 = 0.65$; lullaby: $F(1, 85) = 36.05$, $p < .001$, $R^2 = 0.46$; healing: $F(1, 80) = 17.89$, $p < .001$, $R^2 = 0.61$). An analysis of the partial R^2s from the models showed that the musical features explained comparable degrees of variability in guessing behavior between children and adults (Figure 3d-f; full regression reporting is in SI Tables 1-3). To estimate the degree of similarity, we pooled song-wise guessing percentages and tested whether a model that could differentiate child/adult was statistically equivalent to one that could not, using a chi-square test. Only the model for dance inferences showed a statistical difference between children and adults (dance: $\chi^2 = 5.95$, $p = 0.005$; lullaby: $\chi^2 = 6.45$, $p = 0.073$; healing: $\chi^2 = 3.35$, $p = 0.217$).

Discussion

We showed that 3- to 12-year-old children can discriminate whether songs are intended for soothing an infant, for dancing, or for healing the sick. The songs were unfamiliar to the children, taken from a representative sample of vocal music from 70 human societies, yet their guessing patterns were well above chance and strikingly similar to those of adults. While older children guessed more accurately than younger children, on average, this difference was small, with three-year-olds already reaching near-adult levels of performance. Additionally, musical exposure in the child's home, from either recorded music or from parental singing, had no discernible effect. Together, these results imply that for the particular song types sampled in this study (dance, lullaby, and healing) human intuitions about their function develop early and seemingly independently from direct musical experience.

The similarities in children's and adults' inferences were detectable even at the level of individual musical features of each song. For example, faster tempo, more stable beat structure, and more rhythmic accentuation led listeners to guess that a song was used for dancing; less rhythmic accentuation and slower tempo cued listeners that a song was used as a lullaby; and songs with a slower tempo, less beat stability, more melodic arpeggiation, and a major tonality suggested that a song was used for healing. But across the board, these acoustical features were comparably predictive of both children's and adults' inferences, suggesting similar mechanisms in both cohorts.

What explains the associations between particular musical features and listener inferences? We speculate that these associations may tie into the prototypical emotional and physiological content of songs' behavioral contexts. Across cultures, dance songs typically aim to increase arousal (e.g., energizing groups of people to dance) whereas lullabies aim to decrease it (e.g., soothing an infant). This is mirrored in the arousal-mediating effects of both rhythmic accentuation (Ilie & Thompson, 2006; Schubert, 2004; Weninger et al., 2013) and tempo (Balch & Lewis, 1996; Holbrook & Anand, 1990; Husain et al., 2002; Yamamoto et al., 2007)—consistent with the inferences made by both children and adults.

Dance songs also aim to not just excite but to have people move their bodies. This may be supported by the psychological effects of 'groove': an impulsive drive for rhythmic body movement in response to music with high beat stability, faster tempos, more rhythmic accentuation, and moderate amount of rhythmic complexity (Janata et al., 2012)—also consistent with the musical features correlated with 'dance' inferences.

The acoustical correlates of healing songs are more difficult to interpret. Their traditional behavioral context is characterized by formal and religious activity across cultures (Mehr et al., 2019), such as manifest in shamanic healing rituals (Singh, 2018). But the difficulty in their interpretation here is that healing rituals sometimes involve prominent dance elements, blurring the distinction between healing songs and dance songs. Indeed, looking at the songs which received the highest percentage of 'dance' guesses: two of the top five were healing songs (as can be seen in Figure 2a), with the highest of these receiving 96% 'dance' guesses, and clearly exemplify the characteristic dance features discussed in the previous paragraph. More generally, despite better than chance identification of healing songs, participants were more likely to guess healing songs as having a dance function

(43.7%) than a healing function (42.6%). But if we instead look at the proportions of song types inferred to be healing songs, 40.20% were actually healing songs and 39.33% were lullabies, and indeed, the songs with the highest percentage of 'healing' guesses were characterized by unaccompanied vocal chanting, differentiated from lullabies by not having reduced accentuation (and having other intensifying features like vibrato), and differentiated from dance songs by having less precise and prominent rhythmic structures.

Taken together, our results suggest that the ability to distinguish dance, lullaby, and healing songs from one another—to infer their behavioral function—comes naturally to humans and relies minimally on experience. Indeed, many music perception abilities appear to be precocial, including beat-based processing (Kirschner & Tomasello, 2009; Phillips-Silver & Trainor, 2005; Winkler et al., 2009; Zentner & Eerola, 2010); sensitivity to tonal structure in music (Lynch & Eilers, 1992); rich memories of musical stimuli (Granier-Deferre et al., 2011) that persist after long delays (Mehr et al., 2016); and early sensitivity to metrical (Hannon & Trehub, 2005) and tonal structures (Lynch et al., 1990) typical of infants' native musical environment. The nature of human music perception — including inferences about particular form-function relationships — might therefore be similar to the perception of arousal signals in vocalizations across species (Filippi et al., 2017; Owren & Rendall, 2001). This contrasts with abilities that require extensive direct experience, such as those involved in using a specific language, or indeed, other aspects of music psychology, such as specific stylistic competence.

This is not to say that experience does not shape listeners' understanding of music: it obviously does, as evidenced by common sense and a variety of studies showing modest differences in the interpretation of musical structure across cultures (e.g., Drake & Heni, 2003). But the results of the present study provide support for the view that some aspects of our musical understanding stem from innate signaling mechanisms not unlike those found in nonhuman vocalizations (e.g., those found in the contexts of infant care and territorial signaling; Mehr et al. 2020). Musical understanding might thus result from a unique combination of acoustic predispositions, built on those shared with other species; social experience, which plays a key role in early learning (Baldwin et al., 1996; Liberman et al., 2017; Tomasello et al., 2005), including in music (Cirelli et al., 2018; Mehr et al., 2016, 2017; Xiao et al., 2018); and their interaction with related modalities, such as movement and emotion (Sievers et al., 2013, 2019).

The findings reported here raise the possibility that despite the vast variety of ways that music entangles with human lives, many of which shaped in large ways by culture and our idiosyncratic experiences, some of these may be rooted in our biology, developing independently of experience as a natural part of the human mind.

Supplementary figures

All supplementary figures can be found at: https://osf.io/29cs4/

References

Bainbridge, C. M., Bertolo, M., Youngers, J., Atwood, S., Yurdum, L., Simson, J., Lopez, K., Xing, F., Martin, A., & Mehr, S. A. (2020). Infants relax in response to unfamiliar foreign lullabies. *Nature Human Behaviour*.

Balch, W., & Lewis, B. (1996). Music-dependent memory: The roles of tempo change and mood mediation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R., Irwin, J. M., & Tidball, G. (1996). Infants' Reliance on a Social Criterion for Establishing Word-Object Relations. *Child Development*, *67*(6), 3135–3153.

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, *17*(1), 43–64.

Bergeson, T., & Trehub, S. (1999). Mothers' singing to infants and preschool children. *Infant Behavior and Development*, *22*(1).

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and Early Language Acquisition. *Frontiers in Psychology*, *3*.

Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants: Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences*, *1423*(1), 66–72.

Darwin, C. (1871). *The descent of man*. Watts & Co.

Drake, C., & El Heni, J. B. (2003). Synchronizing with Music: Intercultural Differences. *Annals of the New York Academy of Sciences*, *999*(1), 429–437.

Filippi, P., Congdon, J. V., Hoang, J., Bowling, D. L., Reber, S. A., Pašukonis, A., Hoeschele, M., Ocklenburg, S., Boer, B. de, Sturdy, C. B., Newen, A., & Güntürkün, O. (2017). Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: Evidence for acoustic universals. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1859).

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847–857.

Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y., & Bassereau, S. (2011). Near-term fetuses process temporal features of speech: Fetuses process temporal features. *Developmental Science*, *14*(2), 336–352.

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

Holbrook, M., & Anand, P. (1990). Effects of tempo and situational arousal on the listener's perceptual and affective responses to music. *Psychology of Music*, *18*(2), 150–162.

Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus monkeys (Macaca mulatta) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *12*.

Husain, G., Thompson, W., & Schellenberg, E. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, *20*(2), 151–171.

Ilie, G., & Thompson, W. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*(4), 319–330.

Kiselev, S., Espy, K., & Sheffield, T. (2009). Age-related differences in reaction time task performance in young children. *Journal of Experimental Child Psychology*, *102*(2), 150–166.

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The Origins of Social Categorization. *Trends in Cognitive Sciences*, 13.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, Experience, and Music Perception. *Psychological Science*, *1*(4), 272–276.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*, *52*(6), 599–608.

Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagen, E. H. (2020). Origins of music in credible signaling. *Behavioral and Brain Sciences*, 1–41.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468), 957–970.

Mehr, S. A., Singh, M., York, H., Glowacki, L., & Krasnow, M. M. (2018). Form and function in human song. *Current Biology*, *28*(3), 356–368.

Mehr, S. A., Song, L. A., & Spelke, E. S. (2016). For 5-month-old infants, melodies are social. *Psychological Science*, *27*(4), 486–501.

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, *111*(981), 855–869.

Norris, R. P., & Harney, B. Y. (2014). Songlines and Navigation in Wardaman and other Australian Aboriginal Cultures. *Journal of Astronomical History and Heritage*, *17*(2), 15.

Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology: Issues, News, and Reviews*, *10*(2), 58–71.

Phillips-Silver, J., & Trainor, L. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430–1430.

Rohrmeier, M., & Rebuschat, P. (2012). Implicit Learning and Acquisition of Music. *Topics in Cognitive Science*, *4*(4), 525–553.

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, *4*.

Scherer, K., Banse, R., & Wallbott, H. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*(1), 76–92.

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, *21*(4), 561–585.

Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences*, *110*(1), 70–75.

Sievers, B., Lee, C., Haslett, W., & Wheatley, T. (2019). A multi-sensory code for emotional arousal. *Proceedings of the Royal Society B*, *286*(1906).

Trehub, S. E., Ghazban, N., & Corbeil, M. (2015). Musical affect regulation in infancy. *Annals of the New York Academy of Sciences*, *1337*(1), 186–192.

Trehub, S. E., & Trainor, L. (1998). Singing to infants: Lullabies and play songs. *Advances in Infancy Research*, *12*, 43–78.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, *16*(2), 193–211.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, *16*(3), 285–295.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691.

Weninger, F., Eyben, F., Schuller, B., Mortillaro, M., & Scherer, K. (2013). On the acoustics of emotion in audio: What speech, music, and sound have in common. *Frontiers in Psychology*, *4*.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, *106*(7), 2468–2471.

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2018). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science*, *21*(2), e12537.

Yamamoto, M., Naga, S., & Shimizu, J. (2007). Positive musical effects on two types of negative stressful conditions. *Psychology of Music*, *35*(2), 249–275.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, *107*(13), 5768–5773.