

Children infer the behavioral contexts of unfamiliar foreign songs

Courtney B. Hilton^{1,2,^,*}, Liam Crowley-de Thierry^{3,^,*}, Ran Yan^{2,4}, Alia Martin³, & Samuel A. Mehr^{1,2,5,*}

¹Haskins Laboratories, New Haven, CT 06511, USA.

²Department of Psychology, Harvard University, Cambridge, MA 02138, USA.

³School of Psychology, Victoria University of Wellington, Wellington 6012, New Zealand.

⁴Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA.

⁵Data Science Initiative, Harvard University, Cambridge, MA 02138, USA.

[^]These authors contributed equally.

*Corresponding author. Emails: courtneyhilton@g.harvard.edu, liam.crowleydethierry@vuw.ac.nz, sam@wjh.harvard.edu

Author Note. All data and materials, including a reproducible version of this manuscript with all analysis and visualization code, are available at <https://github.com/themusiclab/form-function-kids>. Audio excerpts from the *Natural History of Song Discography* are available at <https://osf.io/vcybz> and can be explored interactively at <https://themusiclab.org/nhsplots>. The study was preregistered at <https://osf.io/56zne>. For assistance, please contact C.B.H., L.C-T., and S.A.M.

Music commonly appears in behavioral contexts in which it can be seen as playing a functional role, as when a parent sings a lullaby with the goal of soothing a baby. Humans readily make inferences, on the basis of the sounds they hear, regarding the behavioral contexts associated with music. 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 “used to soothe a baby” than “used for dancing”. Are such inferences shaped by musical exposure or does the human mind naturally detect links between these kinds of musical form and function? Children’s developing experience of music provides a clear test of this question. We studied musical inferences in a large sample of children recruited online ($N = 4,825$), who heard dance, lullaby, and healing songs from 70 world cultures and who were tasked with guessing the original behavioral context in which each was performed. Children reliably inferred the original behavioral contexts with only minimal improvement in performance from the youngest (age 4) to the oldest (age 16), providing little evidence for an effect of experience. Children’s inferences tightly correlated with those of adults for the same songs, as collected from a similar online experiment ($N = 94,207$). Moreover, similar acoustical features were predictive of the inferences of both samples. These findings suggest that accurate inferences about the behavioral contexts of music, driven by universal links between form and function in music across cultures, do not always require extensive musical experience.

Music is a ubiquitous part of human culture that plays a variety of functional roles in our day-to-day lives (Mehr et al., 2019; Trehub et al., 2015; Nettl, 2005; Lomax, 1968): 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, as in the encoding of complex geographical knowledge in song by some groups of Indigenous Australians (Norris & Harney, 2014), other contexts are recurrent across societies and tend to co-occur with the same acoustic features. For example, lullabies tend to have slow tempos; smooth, minimally accented contours; and a narrow dynamic and pitch range, as has been long-established in western societies (Bergeson & Trehub, 1999; Trehub & Trainor, 1998; Trehub et al., 1993a), and recently in more globally representative samples of societies (Hilton et al., 2021; Mehr et al., 2019). The findings that music appears universally in conjunction with particular behavioral contexts (e.g., infant care), and that examples of music in a given context share particular acoustic features (e.g., minimal accents), imply a link between form and function in music production.

Adults are sensitive to this link, in that they reliably distinguish song functions 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. Adult listeners indeed identify lullabies when paired with love songs as foils (Trehub et al., 1993b) and distinguish lullabies, dance songs, healing songs, and love songs from one another (Mehr et al., 2019, 2018).

How do adults come to be sensitive to these musical form-function links? Across the lifespan, musical experience is ubiquitous and rich, including through infancy (Mendoza & Fausey, 2021; Yan et al., 2021) and childhood (Bonneville-Roussy et al., 2013; Mehr, 2014). It is therefore plausible that a sensitivity to form and function in music is the result of acquired associations through direct musical experience; especially so given humans’ detailed long-term memories for music (Krumhansl, 2010; Levitin & Cook, 1996), including in infancy (Mehr et al., 2016; Saffran et al., 2000), and the importance of implicit associative learning in musical knowledge (Rohrmeier & Rebuschat, 2012).

Two forms of evidence point to an alternative explanation, however. First, systematic form-function pairings are found in vocalizations throughout the animal kingdom (Fitch et al., 2002; Morton, 1977), and are thought to reflect innate aspects of vocal signaling (Mehr et al., 2021; Darwin, 1871). Indeed, emotional vocalizations are not only cross-culturally intelligible within humans (in speech: Cowen et al., 2019; Chronaki et al., 2018; Scherer et al., 2001; in music: Balkwill & Thompson, 1999), but also between species (Kamiloğlu et al., 2020; Filippi et al., 2017). Humans even extend these emotional inferences to non-animate environmental sounds (Ma & Thompson, 2015) and abstract shapes (Sievers et al., 2019). And a broader class of innate psychoacoustical relationships are thought to underlie the surprisingly robust cross-linguistic associations between certain speech sounds and meaning in the world’s languages (Yu et al., 2021; Blasi et al., 2016; Imai et al., 2008; Nuckolls, 1999). Form-function inferences in music may rely upon similar, largely-innate mechanisms.

Second, despite their lesser degree of exposure to music, relative to adults, infants are affected by music in ways that are predictable from behavioral function. For example, infants relax more in response to lullabies than non-lullabies, even when the songs are in unfamiliar languages and from unfamiliar cultures (Bainbridge et al., 2021); and infants move rhythmically more in response to metrical music than speech (Zentner & Eerola, 2010). While neither implies that infants are *aware* of the functions of the songs they hear, these early-developing responses to music may provide a grounding for later higher-level social inferences about music, which may develop through interaction with precocial social intelligence (Bohn & Frank, 2019; Mehr et al., 2016; Mehr & Spelke, 2017; Powell & Spelke, 2013; Herrmann et al., 2007; Tomasello et al., 2005), along with functional intuitions about other social signals, such as speech (Vouloumanos et al., 2014; Martin et al., 2012).

To test between these competing explanations — invoking learned vs. innate mechanisms — we studied musical inferences across early and middle childhood, focusing on 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/quizzes/kfc>) 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), and were tasked with guessing the original behavioral context.

To examine the degree to which musical experience affected children’s inferences, we conducted a series of preregistered and exploratory analyses, asking (1) whether children are sensitive to form and function in music; (2) how their 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

(Sheskin et al., 2020). Recruitment was/is open-ended; at the time of analysis, we had complete data from 7,926 participants with reported ages from “3 or under” to “17 or older” in one-year increments.

We opted to only analyze the precisely year-long age bands, excluding the participants who responded with “3 or under” and “17 or older” ($n = 2,653$) since these are harder to interpret (but see SI Text 1 for a supplementary analysis of these participants). We also excluded children whose parents indicated that they had assisted their child during the experiment ($n = 894$); children who indicated they had played the game previously ($n = 685$); children with known hearing impairments ($n = 496$); children who were missing data from the test phase ($n = 42$); and children who had trials with responses faster than 100ms or slower than 10 seconds ($n = 2,168$). While the exclusion criteria for response times were not preregistered, we note that the main findings replicate under a variety of alternative exclusion decisions, or when not excluding any trials at all.

This left 4,825 participants for analysis (1,821 male, 2,804 female, 200 other; the cohort’s breakdown of age and gender, after exclusions, is in SI Figure 1). Participants reported speaking 127 native languages and being located in 126 countries (SI Tables 1 and 2).

Readers should note that the experiment itself was conducted in English (including both written and spoken/audible instructions), however. As such, despite the apparent diversity of countries of origin and native languages, these data should not be considered a representative sample of cultural or linguistic experiences, and is likely to be biased in certain ways (e.g., those with internet access who can speak English; see Discussion).

Stimuli

Song excerpts were drawn from the *Natural History of Song Discography* (Mehr et al., 2019), a corpus of vocal music collected from 86 mostly small-scale societies, including hunter-gatherers, pastoralists, and subsistence farmers. Each song in this corpus was originally performed in one of four behavioral contexts: dance, healing, lullaby, and love. All recordings were selected on the basis of supporting ethnographic material, as opposed to the acoustic features of the songs (see: Mehr et al., 2018, 2019).

Because love songs were ambiguously detected in some prior work (Mehr et al., 2018), and given the difficulty of explaining this category to children, we omitted love songs from the experiment and studied only the remaining three contexts (dance, lullaby, and healing). This left 88 songs from 70 societies, sung in 30 languages (SI Table 3) and originating from locations corresponding to 46 countries (SI Figure 3).

Procedure

Participants could use a computer, mobile phone, or tablet; they were encouraged to use a computer, however, for easier user interface navigation. A schematic of the experiment is in Figure 1; it included a progress bar, displayed throughout the experiment, which indicated the times when the child should be assisted by a parent and the times the child should not be assisted.

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 would complete the experiment while wearing headphones. We also asked about the frequency with which the child was exposed to parental singing or recorded music in the home.

Parents then turned the experiment over to their children, who were guided by an animated character, “Susie the Star”. Susie provided verbal encouragement to keep the children motivated; explained the task instructions, which were also presented visually (although many of the youngest children were likely unable to read them); and reminded the children of these instructions throughout the experiment.

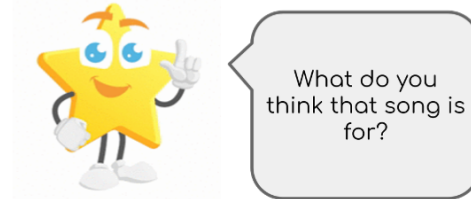
a [progress bar with parent compliance guidelines]



b [narrated instructions]



d [test trial: child's response]



c [test trial: audio]

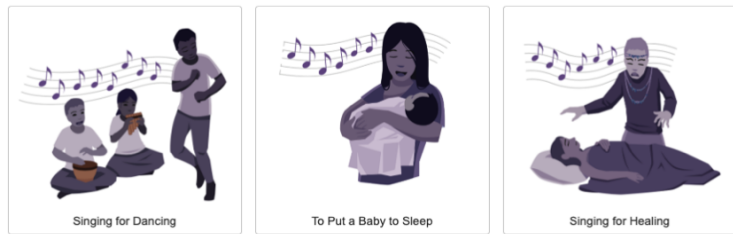


Figure 1. Schematic of the test phase of the experiment. Children listened to the songs and indicated their responses in the context of an online game. **a**, Parents guided their children's participation with the assistance of a progress bar, which indicated which portions of the experiment were intended to be for parents alone, parents and children together, or children alone. The progress bar was visible throughout the experiment. **b**, The game was narrated by *Susie the Star*, an animated character whose instructions were spoken aloud and printed in speech bubbles on the screen. **c**, During testing children listened to a series of songs. **d**, After each song, *Susie* asked them to guess its behavioral context and indicate their response by clicking or tapping an image (with accompanying text).

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”; 98% of children did so correctly on the first try.

Susie then explained 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”). An illustration accompanied each song category’s description, to facilitate responses for children too young to read (see Figure 1). Before proceeding to the test trials, we asked parents whether they felt their child understood the task; 98% responded in the affirmative and those that did not were required to repeat the training trials.

Susie then played the children a counterbalanced set of six songs (two per song type) drawn randomly from the *Natural History of Song Discography* and presented in a random order. The excerpts were 14 seconds long and each played in its entirety before the child could advance. After each song, Susie asked “What do you think that song was for?” and children made a guess by clicking or tapping on a labeled illustration corresponding with one of the three song types (see Figure 1; i.e., a three-response classification task). After they indicated their inference, we also asked children to rate how much they liked the song, for use in a different study. Positive 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 (those who had were excluded from analyses; see Participants, above).

Pilot studies and preregistration

In designing the experiment, we first piloted the experiment in-person with four children (ages 6, 7, 9, and 10), to explore whether the procedure was well-explained and intuitive for children, and to observe potential parental interference. The children easily understood the task and did not elicit help from their parents; at times they *rejected* parental assistance (e.g., as if to say “I can do this on my own, mum”).

We continued by running a pilot version of the experiment with 500 children recruited online (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: (i) in the full cohort of children, accurate classification of behavioral contexts overall and in all three song types, including after adjusting for response bias via d' analyses; (ii) a positive but modestly sized effect of age on accuracy ($R^2 < .05$); and (iii) little to no effect 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: (iv) predicted correlations between children’s intuitions about the songs and adults’ intuitions, using an expanded sample from a previous study of adults (Mehr et al., 2019); and (v) 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 preregistered hypotheses for future research.

The preregistration is available at <https://osf.io/56zne>.

Notes on the unmoderated citizen-science approach

Most psychological studies of children are conducted in-person (e.g., in a laboratory or school) or in a moderated online setting (e.g., via videoconference). Our approach instead relies on an unmoderated, *citizen-*

science approach. While this approach has the advantage of larger-scale, and often more diverse recruitment than in-person or moderated online studies (Hilton & Mehr, 2022; Li et al., submitted), it has two potential risks.

The risk of participant misrepresentation: Are the children really children?

As we did not directly observe the participants, we cannot verify that participants in the study are, in fact, children, as opposed to adults posing as children, or malicious automated participants (e.g., bots).

As no compensation was given, it is not clear what incentive there would be for either possibility, and there is growing evidence that uncompensated online and compensated lab-based data collection produce similar results across a number of domains (Huber & Gajos, 2020; Hartshorne et al., 2019; Germine et al., 2012). Nevertheless, we analyzed several forms of metadata to explore the risk of participant misrepresentation.

First, we examined the timecourse of participant recruitment across the child and adult cohorts (SI Figure 3). In particular, we inspected large spikes in recruitment, with the idea that spikes that are not explicable by known events could be attributable to bots or other malicious participants. All spikes in recruitment were attributable to known internet exposure such as viral news coverage, however, suggesting that increases in recruitment during these spikes corresponded with *bona fide* participants.

Second, we compared the times of study completion across the child and adult cohorts across several countries (SI Figure 4). Children generally keep different schedules than do adults (e.g., most adults do not attend elementary school) and thus, presumably, would have different opportunities to participate in the experiment. We found that participants in the child cohort were less likely to participate in the late night/early morning, relative to adults, a pattern consistent across 5 countries with large samples in both children and adults.

Third, we examined the patterns of response times within the child cohort in the experiment (SI Figure 5). Consistent with prior work showing that response time in perceptual and cognitive tasks decreases over the course of childhood (Hale, 1990; Kiselev et al., 2009), response times during the task decreased reliably as a function of self-reported age ($F(1, 4,823) = 808.01, p < .001, R^2 = 0.14$; SI Figure 5).

These considerations suggest that the risk of participant misrepresentation is limited.

The risk of parental non-compliance: Did parents bias children’s responses?

As we did not directly observe the participants, we cannot verify whether or not parents complied with the instructions to let their children participate uninterrupted and without providing advice or other interference. We believe this risk is mitigated by several concerns, however.

First, we designed the experiment to be easy to understand, so as to limit the need for children to ask for help. We used clear, child-friendly visual and audio elements and interactive training trials to help the child intuitively understand the task; for the parents, we included a progress bar throughout to communicate when parents should and should not intervene. These are depicted in Figure 1 and should in principle limit parent intervention. Indeed, in in-person pilot participants, children did not elicit any help from parents, and even discouraged their parents from intervening (see Pilot studies and preregistration, above).

Second, at the completion of the online experiment, we explicitly asked parents whether they complied with the request to not interfere with the child’s performance. Post-hoc prompts on task compliance have been shown to be effective in other uncompensated online experiments (Reinecke & Gajos, 2015); parents have little incentive to lie in response to such a question. The vast majority of parents stated that they did not assist their children in the experiment and we excluded those who did (see Participants, above).

Third, even if parents *did* try to help their child, adult performance in the task reported here is far below ceiling (Mehr et al., 2019): parents would unlikely have been confident that they knew the answers themselves.

Fourth, the large sample size afforded by the unmoderated citizen-science approach can absorb considerable noise arising from violations to task compliance. It seems unlikely, for example, that the majority of parents

(i.e., thousands of people) would have intentionally disregarded the clear instructions to not interfere. This is especially the case given one reward we did provide to parents: a few minutes of free time while their child was busy completing the experiment. As much of our cohort participated during the COVID-19 pandemic, when many parents had reduced childcare resources, we suspect that many parents welcomed the opportunity to *not* pay attention to their child’s responses.

These considerations suggest that the risk of parental non-compliance is limited.

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%, both overall (Figure 2a; $M = 54.0\%$, $SD = 21.6\%$, $p < .001$) and within each song type (dance = 71.5%; lullaby = 45.6%; healing = 44.8%; $ps < .001$), as they did in exploratory results, and confirming the first preregistered hypothesis.

However, children did not use the three response options evenly (SI Figure 6a): they guessed “dance” most frequently (42.2%), followed by “healing” (36.2%), and least often guessed “lullaby” (21.6%). This is suggestive of a response bias, and as such, raw accuracy (i.e., percent correct) is difficult to interpret.

Thus, we also computed d' scores to assess sensitivity to song function independently from response bias, doing so for the whole cohort, and for each age-group using a pooled estimator (Macmillan & Kaplan, 1985; see a simulation of the effect of this analytic choice in SI Figure 7). Lullabies were the most reliably classified ($d' = 1.21$, 95% CI [1.15, 1.26], $t(12) = 47.47$, $p < .001$, $d = 1.20$; one-sample two-tailed t-test), closely followed by dance songs ($d' = 1.17$, 95% CI [1.12, 1.22], $t(12) = 49.88$, $p < .001$, $d = 1.69$). Healing songs were the least reliably classified but still robustly above chance ($d' = 0.37$, 95% CI [0.31, 0.42], $t(12) = 14.73$, $p < .001$, $d = 0.48$). Estimated criterion scores are reported in SI Figure 6b.

This principal result was robust to the inclusion or exclusion of two subgroups of participants whose reported ages were ambiguous (SI Text 1) and was consistent across subgroups of participants with different native languages and locations (SI Text 2).

Musical inferences did not appreciably improve 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. While the effect of age was statistically significant, the effect size was miniscule ($F(1, 4,823) = 11.42$, $p < .001$, $R^2 = 0.002$), confirming our second pre-registered hypothesis. In an exploratory analysis, we tested the effect of age on d' scores (instead of raw accuracy); the effects were comparable, with no age effect overall or in any of the song types individually ($ps > 0.05$).

Musical inferences are unrelated to children’s home musical environment

We tested the relationship between the frequency of parent-child musical interactions (i.e., singing or playing recorded music) and children’s musical inferences. Consistent with our third preregistered hypothesis, we found no evidence for a relationship between accuracy and the frequency of parental singing (SI Figure 8a; $F(1, 4,821) = 0.57$, $p = 0.45$), or with the frequency of recorded music (SI Figure 8b; $F(1, 4,821) = 0.82$, $p = 0.37$).

Musical inferences are highly similar between children and adults

With no substantive 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

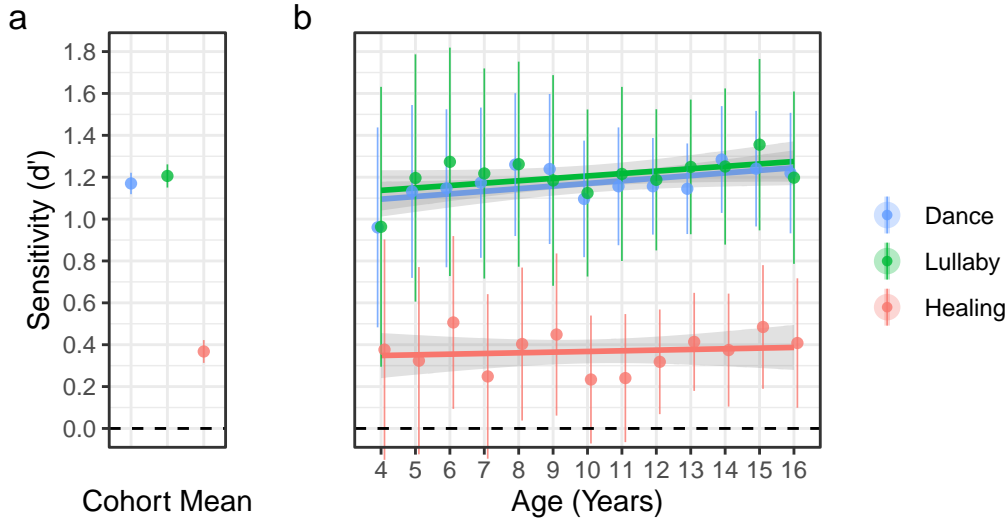


Figure 2 | Accurate classification of all three song types and cohort-wide, with no effect of age. (a) Mean d' -prime scores across all children show above-chance classification of each song type, independent of their response bias. (b) Above-chance sensitivity replicates for each age group, except for some of the healing song estimates whose 95% confidence-intervals crossed zero. Performance did not improve with age, either averaging across all song-types, or within dance, lullaby, or healing songs individually (p s > 0.05). In both panels, the circles indicate 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.

<https://themusiclab.org>, a similar experiment for adults¹ using the same general task design and stimuli (but with a fourth song type, “love songs”, and response option, “used to express love to another person”, neither of which were used in the children’s experiment). Data were available from 94,207 participants who were older than 18 years (55,132 male, 36,028 female, 1,991 other, 1,056 unknown/missing; mean age = 31.0 years, SD = 12.8, range: 18-99). 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) within each cohort (i.e., children or adults) and 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: $F(1, 86) = 2,924.38$, $R^2 = 0.97$; lullaby: $F(1, 86) = 2,677.81$, $R^2 = 0.97$; healing: $F(1, 86) = 367.22$, $R^2 = 0.81$; p s < .001), confirming our fourth pre-registered hypothesis. The strong relationship between the guesses of children and adults also held robustly across children of all ages and across the three song types, with uniformly high R^2 values and slight increases in R^2 through childhood (SI Figure 9).

In a more conservative, exploratory analysis, we transformed the measures of inferences from song-wise continuous scores (i.e., comparing the cohorts’ proportions of responses for each song type) to song-wise rank scores (e.g., integers from “most-guessed to be a lullaby” to “least-guessed to be a lullaby”). We quantified the strength of these relationships using the non-parametric Kendall’s rank correlation coefficient τ and found a strong positive correlation between children’s and adults’ inferences for all song types (dance: $\tau = 0.89$; lullaby: $\tau = 0.85$; healing: $\tau = 0.72$; p s < .001).

¹Note that a portion of these data were previously reported in Mehr and colleagues (2019) but were not analyzed in relation to children’s performance.

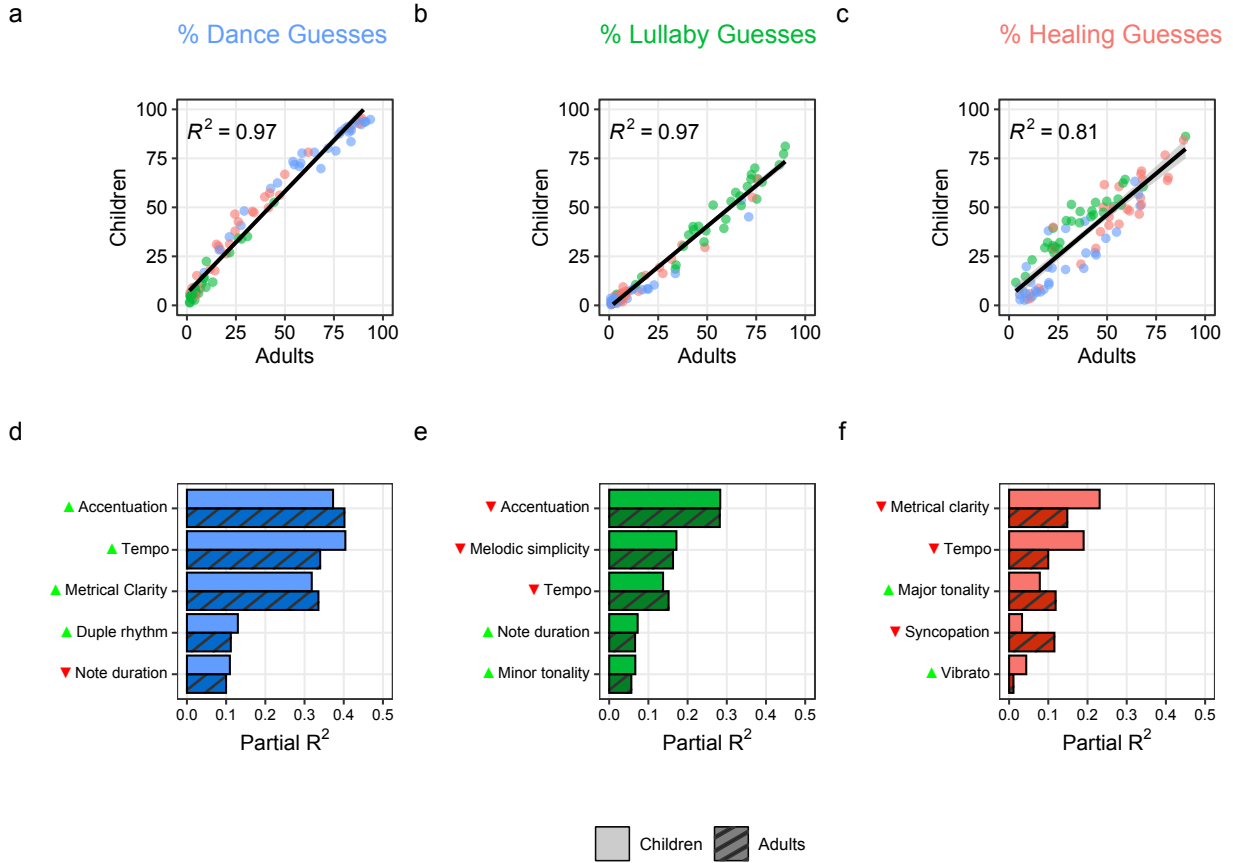


Figure 3 | Children and adults make highly similar musical inferences, which are driven by the same acoustical features. The scatterplots (a-c) show the tight correlations between children's and adults' musical inferences. Each point represents average percent guesses that it was (a) “for dancing”; (b) “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- R^2) 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. The arrows beside each musical feature indicate the direction of effect, with green upwards arrows indicating increases (e.g., faster tempo) and red downward arrows indicating decreases (e.g., slower tempo).

Child and adult musical inferences are driven by the same acoustic correlates of musical function

As a further test of the fourth preregistered hypothesis, we explored what might drive children’s musical inferences. We analyzed which musical features were predictive of children’s guesses, within each song, and asked whether these musical features corresponded with those that were predictive of adults’ guesses.

We began with an initial set of 36 musical features (annotations from expert musicians and variables derived from transcriptions of the songs) that were previously studied in Mehr and colleagues (2019), which were z-scored to facilitate model fitting and comparison across features. To avoid overfitting, we then selected a smaller subset of the full 36 features via LASSO-regularized modeling (Friedman et al., 2016), performed separately for each song type (dance, lullaby, healing) and for each population (child, adult). Each model predicted the naive listener inferences for each of the 88 songs, and was trained with 10-fold cross-validation, repeated 10 times for robustness. We then conservatively discarded resulting model coefficients below an arbitrary threshold of 0.02 as a form of secondary regularization (erring on the side of a simpler model; further minimize overfitting).

The results lay out the musical features that most highly influence children’s and adults’ inferences for each song type (SI Table 4; which also describes the features themselves). To quantify their influence, we regressed the percentage of song-wise guesses in each category on the LASSO-selected musical features. For all three song types, the features explained substantial variability in listener inferences (dance: $F(1, 82) = 45.44$, $R^2 = 0.73$; lullaby: $F(1, 76) = 13.67$, $R^2 = 0.66$; healing: $F(1, 81) = 20.45$, $R^2 = 0.60$; $ps < .001$).

Notably, the partial R^2 s from these models show that children’s and adults’ inferences, especially in the cases of lullabies and dance songs, were guided in a highly similar fashion: each musical feature explained a similar degree of variability in guessing behavior across the two cohorts (Figure 3d-f; full regression reporting is in SI Tables 5-7). Indeed, in a model predicting inferences from the LASSO-selected features for each song type, with a categorical variable for the cohort (children vs. adults), and interaction terms for each feature with cohort, children’s and adults’ inferences were not statistically distinguishable at the level of any musical feature (all interaction term $ps > 0.05$).

Children’s inferences are driven by objective acoustic correlates of musical function

Previous work demonstrated that a core set of musical features reliably distinguished dance songs, lullabies, and healing songs from one another across 30 world regions (Mehr et al., 2019). In an exploratory test of our fifth preregistered hypothesis, we tested the degree to which children’s inferences reflected these universal musical biases.

We first trained a LASSO multinomial logistic classifier to infer song type from the acoustic features, trained with leave-one-out cross-validation at the level of world region. We then derived predictions for each song in the corpus such that the predictions for songs in a given world-region were based on information from only songs in other regions. This yielded groupings of three probabilities per song (e.g., for an ambiguous song, the model might predict [0.3 dance; 0.3 lullaby; 0.4 healing], whereas for a less ambiguous song, the model might predict [0.8 dance; 0.0 lullaby; 0.2 healing]).

We asked how similar these groupings were to the songwise guessing proportions derived from the children’s responses. We found strong positive correlations for dance inferences ($r = 0.80$) and lullaby inferences ($r = 0.67$), and a moderate correlation for healing inferences ($r = 0.44$). Thus, children’s inferences, which, of course, are made without any explicit reference to acoustic features, were shaped by many of the same general acoustic features that objectively characterize song types across cultures.

This was not the case across *all* musical features, however. In contrast to the model comparing children and adults (above), where the two cohorts’ inferences could not be reliably distinguished on the basis of particular musical features (i.e., feature-by-cohort interaction terms were non-significant), we found some reliable differences in how the musical features operated worldwide and how they influenced children’s inferences.

In a model predicting song-wise probabilities from the LASSO-selected features for each song type, with a categorical variable for the data type (children’s inferences vs. objective classification), and interaction terms for each feature with data type, we found several statistically significant differences. For example, children were more likely to rate a song as “for dancing” when it had more rhythmic accentuation, faster tempo, greater metrical clarity, more duple rhythm, and shorter note durations, but in each case, children *over*-estimated their importance relative to the modeled objective differences (interaction $ps < .05$; see the full models in SI Tables 8-10). But the high degree of similarity between children’s and adults’ inferences suggests that the two groups make such errors in similar fashions, deviating from the features that reliably distinguish song types from one another worldwide.

Discussion

We found that children make accurate inferences concerning the behavioral contexts of unfamiliar foreign lullabies, dance songs, and healing songs. The songs were unfamiliar to the children and drawn from a representative sample of vocal music from 70 human societies, but their guessing patterns were well above chance and strikingly similar to those of adults. Older children (teenagers) performed no better than the youngest children in our sample (age 4), who already performed at adult-like levels. Musical exposure in the home, from either recorded music or parental singing, was also unrelated to children’s performance. These results suggest that the ability to infer musical function from acoustical forms, at least in the contexts studied here, develops early and requires minimal direct 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 and less beat stability suggested that a song was used for healing. Children’s inferences were also robustly correlated with those of a classifier trained on the objective acoustic differences in the corpus, demonstrating a link between early perceptual intuitions and universal acoustic biases that shape human music production. This link, and the few cases where children’s intuitions differed from universal biases, present an intriguing target for future research.

What explains these robust associations between musical features and functional inferences? We speculate that they 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: [Hilton et al., 2021](#); [Mehr et al., 2019](#)). This is mirrored in the arousal-mediating effects of both rhythmic accentuation ([Weninger et al., 2013](#); [Ilie & Thompson, 2006](#); [Schubert, 2004](#)) and tempo ([Yamamoto et al., 2007](#); [Husain et al., 2002](#); [Balch & Lewis, 1996](#); [Holbrook & Anand, 1990](#)). As in other universal acoustic form-function mappings ([Perlman et al., 2021](#); [Patten et al., 2018](#); [Fitch et al., 2002](#); [Fitch, 1997](#)), this relationship may be grounded in the physics of sound production: the production of sounds with high tempo and accentuation, for example, requires more energy than does the production of slower, less accentuated sounds. These features are thus associated with higher arousal. Indeed, from a young age, infant arousal is reliably modulated by these features ([Bainbridge et al., 2021](#); [Cirelli et al., 2019](#)).

Beat stability was also predictive of “used for dancing” inferences. Dancing characteristically involves temporal coordination, often between people and a perceived musical beat. As such, the more stable the beat, the more reliable it is as a cue to support temporal coordination. Children may naturally make this connection via the psychological effects of “groove”, an impulse for rhythmic body movement in response to music with high beat stability, fast tempos, more rhythmic accentuation, and a moderate amount of rhythmic complexity ([Witek et al., 2014](#); [Janata et al., 2012](#)). In both children and adults, these features were the strongest predictors of “used for dancing” inferences, and there is some evidence that infants as young as 5 months share such impulses linking rhythmic musical features to movement ([Zentner & Eerola, 2010](#)).

The acoustic correlates of healing songs are more difficult to interpret. Their traditional behavioral context tends to be characterized by formal and religious activity across cultures ([Mehr et al., 2019](#)), such as those found in shamanistic healing rituals ([Singh, 2018](#)). But these rituals sometimes include dancing, and may

thus constitute a fuzzier and less distinct category than dance or lullaby; for example, among both children and adults, the song most consistently categorized as a dance song (96% of the time) was, in fact, a healing song. The weaker discriminability of healing songs by naive listeners, and the weaker correlations between human inferences and objective distinguishing features underscores this fuzziness. This fuzziness is also germane to the behavioral context of healing songs more broadly, since a person’s health is rarely externally observable, unlike the relatively more concrete referents of babies sleeping and people dancing. Indeed, Singh (2018) argues that healing rituals often function in part to signal a healer’s otherworldliness and ability to transact with unobserved spiritual causes of health. So while music appears universally in the context of healing (Mehr et al., 2019; Singh, 2018), and has well-recognized potential in modern clinical settings (Cheever et al., 2018), the mechanisms and acoustical correlates of inferences about traditional healing songs are far less clear and require further study.

Together, our findings suggest that some form-function inferences about music may be maturational rather than experiential in origin: requiring minimal direct experience to develop. Many other facets of music perception develop precociously in this way, including beat processing (Zentner & Eerola, 2010; Winkler et al., 2009; Phillips-Silver, 2005); sensitivity to basic tonal structure (Perani et al., 2010; Lynch & Eilers, 1992); prenatally present attentional orientations to music (Granier-Deferre et al., 2011); biases for socially-relevant musical information (Mehr et al., 2016; Mehr & Spelke, 2017); and indeed, distinct physiological responses to lullabies, relative to non-lullabies (Bainbridge et al., 2021). Some of these abilities continue to develop with experience; for example, in the first year of life, the perception of metrical (Hannon & Trehub, 2005) and tonal (Lynch et al., 1990) structures becomes tuned to regularities of the infants’ native musical culture, and later stages of this can even continue past 12 years of age in ways that depend upon rich experiential input (Brandt et al., 2012). By contrast, the present findings show an adult-like competence as late as 4 years of age (and potentially earlier; see SI Text 1), irrespective of variation in musical exposure.

This is not to say that experience does *not* shape listeners’ understanding of music. Experience shapes many facets of musicality, from musical preferences (Schellenberg et al., 2008; Peretz et al., 1998), to the enculturation of extramusical associations (Margulis et al., 2022). Experience can even exert subtle influences on our perception of basic structures of rhythm (Drake & Heni, 2003) and pitch (Jacoby et al., 2019); the influences are not be limited to *musical* experience, as shown by the effect of linguistic experiences on music processing (Liu et al., 2021). And while we studied a relatively diverse sample of children and a highly diverse set of recordings (see SI Tables 1-3), we did not systematically sample diverse experiences: the study instructions were in English and the children were recruited via an English-language website. These limitations suggest a need for further studies conducted in multiple languages and recruiting from a wide variety of geographic regions, so as to better characterize the potential effects of early musical experience on listeners’ understanding of music.

Our results raise the possibility that some forms of musical meaning may not be entirely experience-dependent, building instead on instinctive signalling mechanisms, not unlike those thought to govern many nonhuman-animal vocalizations (Mehr et al., 2021). From this perspective, foundations for musical understanding come from the combination of acoustic predispositions, shared with other species (Kamiloğlu et al., 2020; Filippi et al., 2017; Owren & Rendall, 2001); human-unique social-cognitive skills and biases (Herzmann et al., 2007), which play a key role in early learning (Liberman et al., 2017; Kinzler et al., 2007; Baldwin et al., 1996), including in music (Mehr & Krasnow, 2017; Xiao et al., 2017; Mehr et al., 2016); and their interaction with related modalities, such as movement and emotion (Sievers et al., 2019, 2013).

Thus, despite the variety of roles that music plays in our lives and the different meanings it can afford, shaped by culture and experience, the results reported here suggest that aspects of our musical intuitions are rooted in our biology. The psychology of music, like other acoustic predispositions, may develop as a natural component of the human mind.

Context of the research

This research builds on previous studies from our group (Mehr et al., 2019, 2018; Hilton et al., 2021), demonstrating (a) cross-cultural regularities in the associations between acoustical forms of music and par-

ticular behavioral contexts in which music appears; and (b) the sensitivity of adults to these associations. Here, we explored how musical experience in childhood shapes that sensitivity, using the same stimuli as previous work. We also compared children’s intuitions about music to adults’ directly, and analyzed how these intuitions correlate with specific musical features (providing clues about the mechanisms underlying the main effects). The results complement those of another recent study ([Bainbridge et al., 2021](#)), where infants showed differential physiological responses to unfamiliar foreign lullabies, relative to non-lullabies; those early-appearing, implicit responses to different forms of music, may help to explain children’s high performance in classifying behavioral contexts for music. Together, these studies inform theories of the basic design features of a human psychology of music, including debates surrounding the biological foundations of music ([Mehr et al., 2020](#)) and its relation to more general aspects of cognition ([Hilton et al., 2020](#)).

Acknowledgments

We thank all the families who participated in this research; C. Bainbridge, A. Bergson, M. Dresel, and J. Simson for assistance with designing and implementing the experiment; and L. Yurdum for helpful comments on the manuscript.

Author contributions

- S.A.M. and A.M. conceived of the research and provided funding and other resources.
- L.C-T., A.M., and S.A.M. designed the experiment.
- L.C-T. wrote code to implement the experiment, with contributions from S.A.M.
- L.C-T., A.M., and S.A.M. piloted the experiment.
- L.C-T., C.B.H., and S.A.M. collected the data.
- C.B.H. led analyses and designed the figures, with contributions from L.C-T., R.Y., A.M., and S.A.M.
- C.B.H. and S.A.M. wrote the original manuscript, with contributions from L.C-T., R.Y., and A.M.
- C.B.H. wrote the revision, with contributions from L.C-T and S.A.M., and all authors approved it.

Supplementary Information

SI Text 1. Analysis including participants aged “3 or younger” and “17 or older”

In the main text, we excluded participants who reported their age in the ambiguous response categories of “3 or under” or “17 or older”, as these spanned wider ranges of possible ages than the other, one-year response categories, complicating their interpretation. As this exclusion decision was not preregistered, for transparency we repeated the main analyses without excluding these two groups, so as to ensure that the post-hoc exclusion decision did not affect our conclusions.

It did not: accurate discrimination was robustly replicated within both “3 or under” (lullaby $d' = 1.15$, 95% CI [0.98 1.31]; dance $d' = 1.21$, 95% CI [1.08 1.33]; healing $d' = 0.56$, 95% CI [0.43 0.68]) and “17 or older” (lullaby $d' = 1.36$, 95% CI [1.29 1.43]; dance $d' = 1.30$, 95% CI [1.25 1.34]; healing $d' = 0.46$, 95% CI [0.41 0.51]). These findings raise the possibility that the ability to discriminate form and function in music may well extend to children younger than four years of age.

Other key findings reported in the main text, such as those concerning age effects, did not change substantively when including these two groups. For brevity, we did not repeat them here, but readers interested in exploring this question are welcome to use our open data and code to do so.

SI Text 2. Variation across countries, languages, and songs

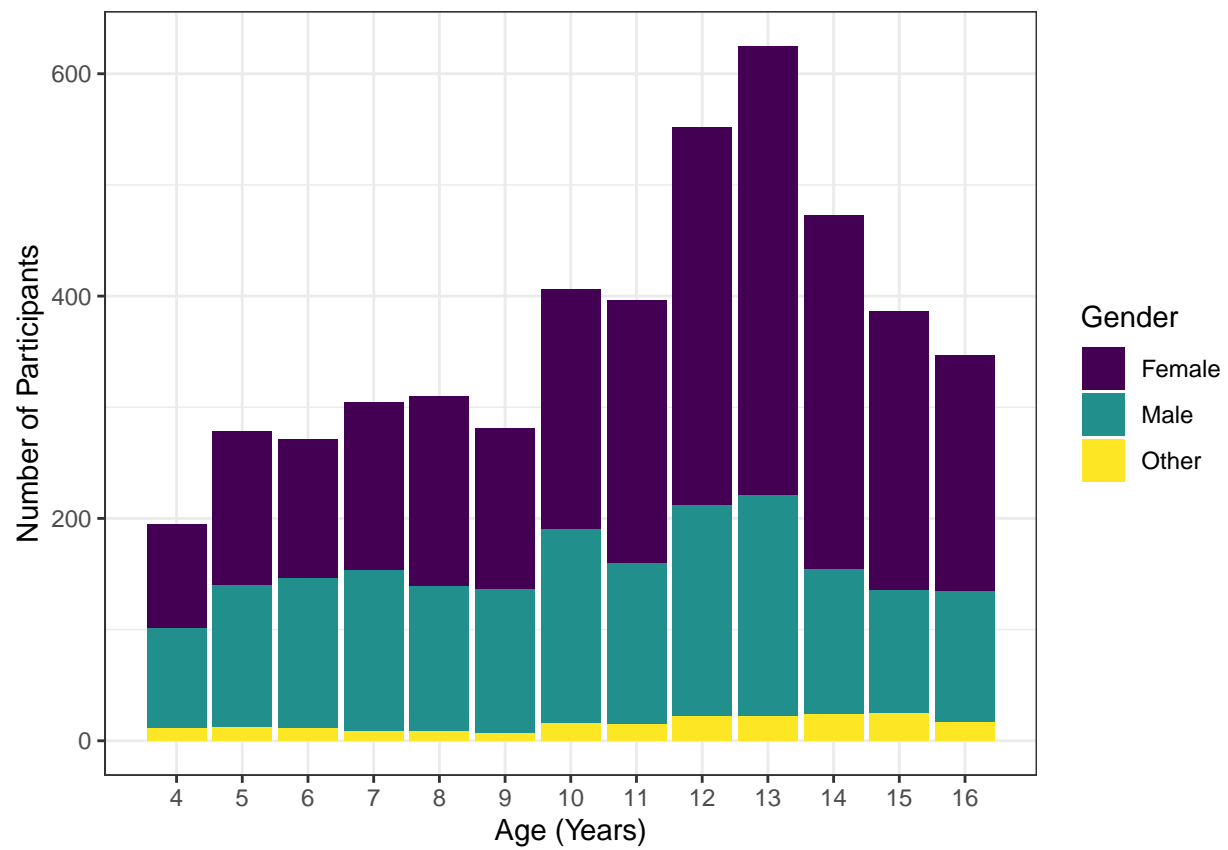
Because the instructions of the experiment were only presented in English, there could be differences in performance between children whose native language is English and children whose native language is not English.

We tested this question in two ways. First, we replicated the main d' analyses using only data from English-speaking participants. The findings robustly replicated, with slightly lower scores (dance $d' = 1.05$; lullaby $d' = 0.94$; healing $d' = 0.26$; $ps < 0.05$). The effect of age was not statistically significant for any of the song types ($ps > 0.05$).

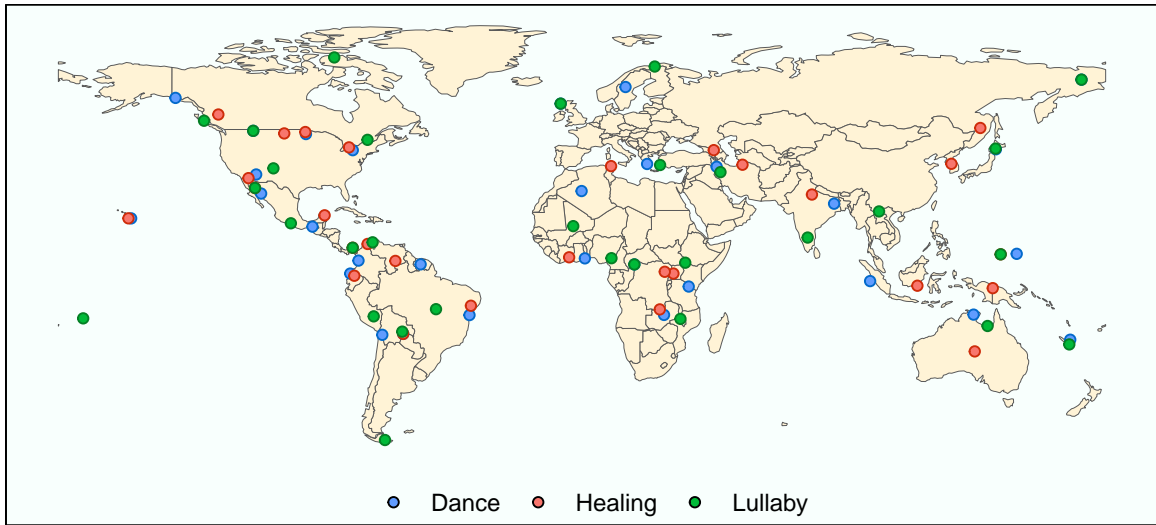
Second, as a more general test of this issue, we measured the degree of variability across the 126 countries of origin and the 127 languages using coefficient of variation scores derived from a mixed-effects model predicting accuracy. In each case we found very little variation (0.036 and 0.026, respectively). To put these scores in context, the coefficient of variation for the 88 songs used in the experiment was 0.463 (i.e., ~10/20 times larger).

Together, this suggests that the heterogeneity of our participants, and the varying degrees to which the children may have comprehended English, likely had a negligible influence on our results.

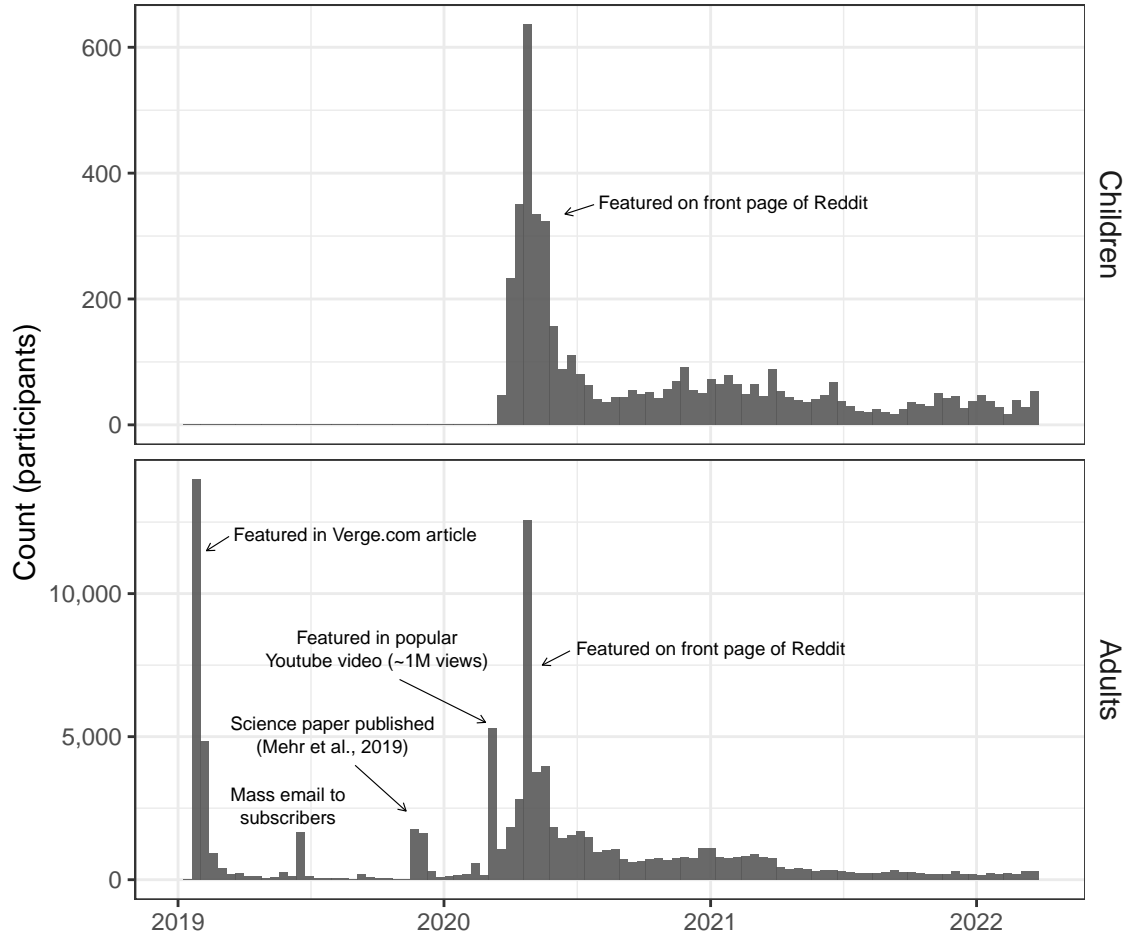
Supplementary Figures



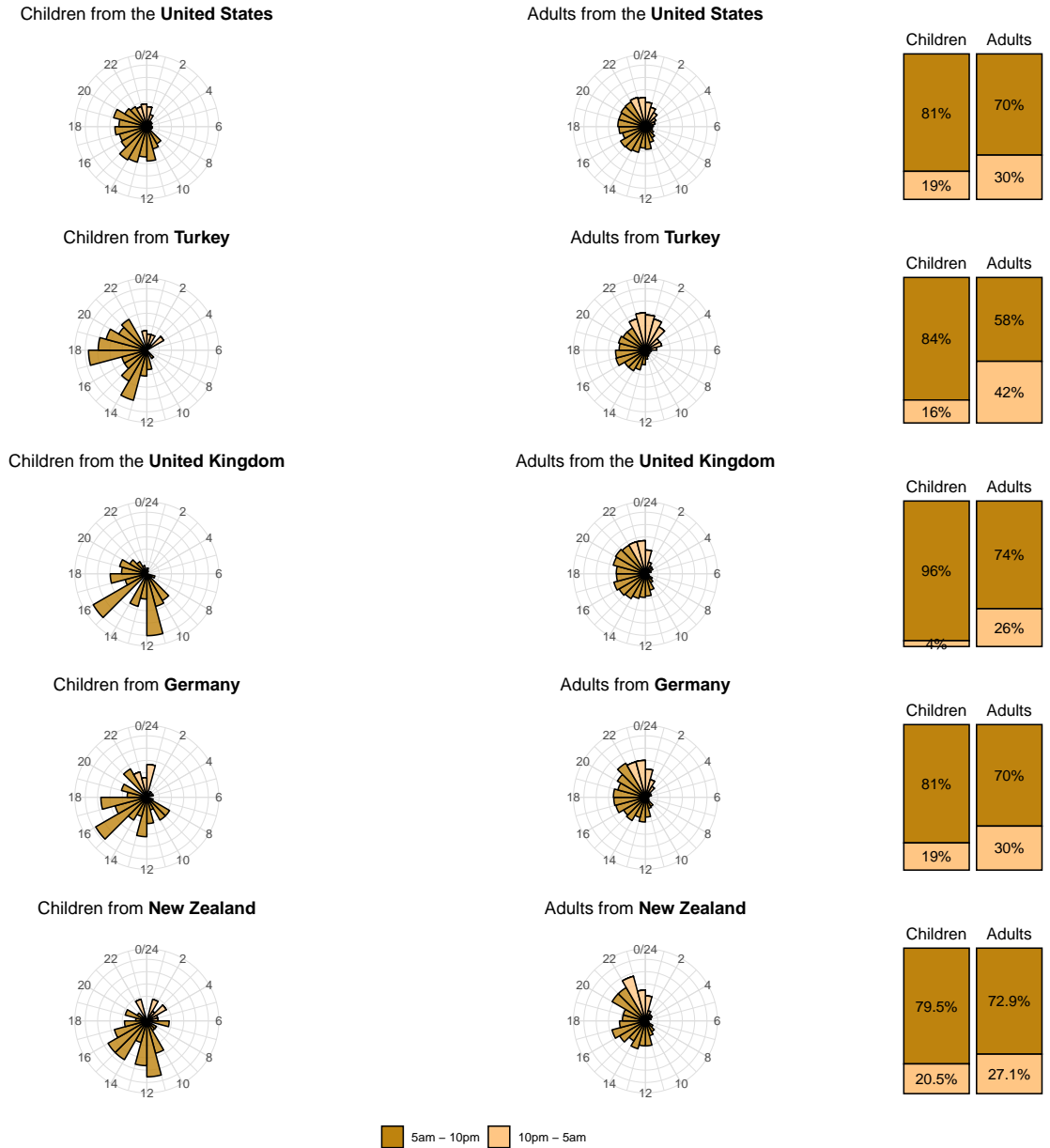
SI Figure 1. The distributions of ages and genders of the participants.



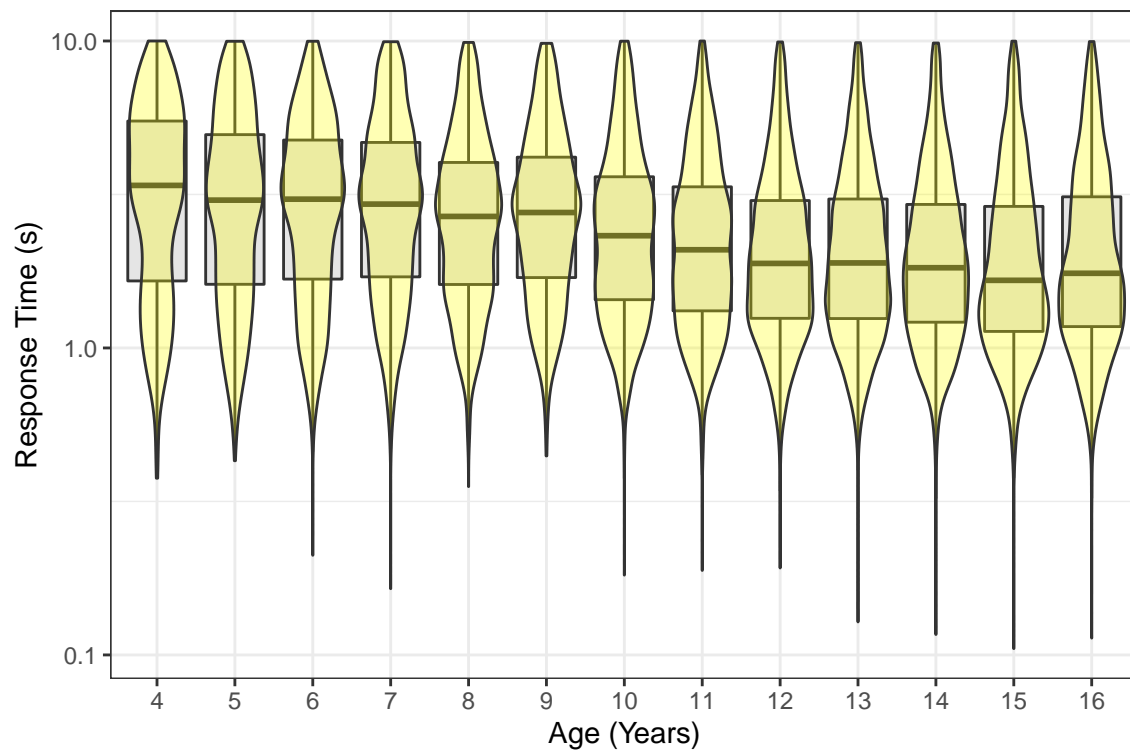
SI Figure 2. The approximate locations of societies in which the *Natural History of Song Discography* recordings used in the experiment were originally recorded.



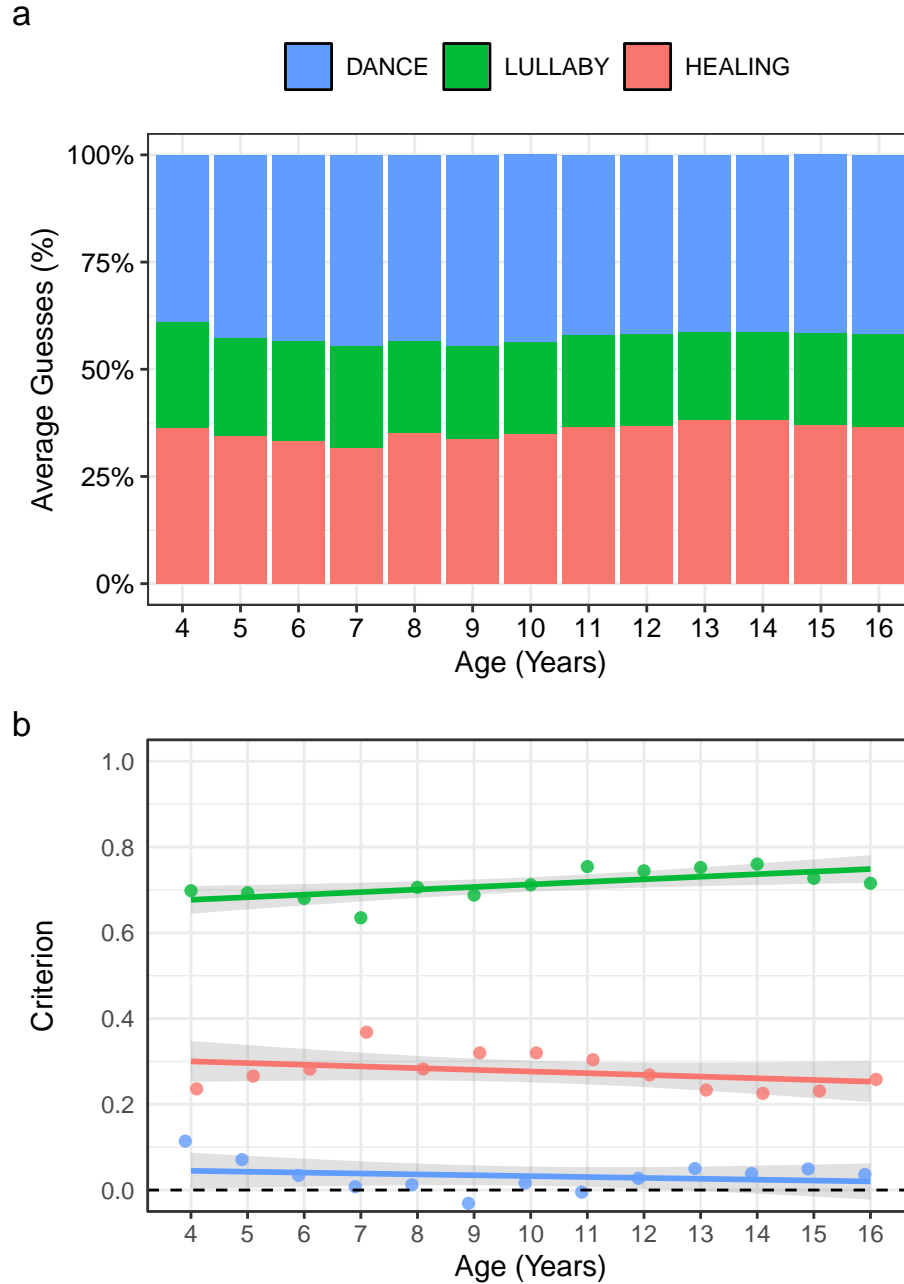
SI Figure 3. The time-series of recruitment in children and adults. The histograms show the number of participants (after exclusion) that completed the experiments (the child version, top; or adult version, bottom) between 2019 and 2022. The date of participation is indicated on the x -axis. The annotations on each plot indicate specific events that coincided with spikes in participation, such as viral posts on social media, news coverage, etc. Other than these explicable spikes in participation, we see evidence of steady recruitment commensurate with all of our prior and current online experiments, suggesting that the participants were genuine, and not malicious entities (i.e., bots).



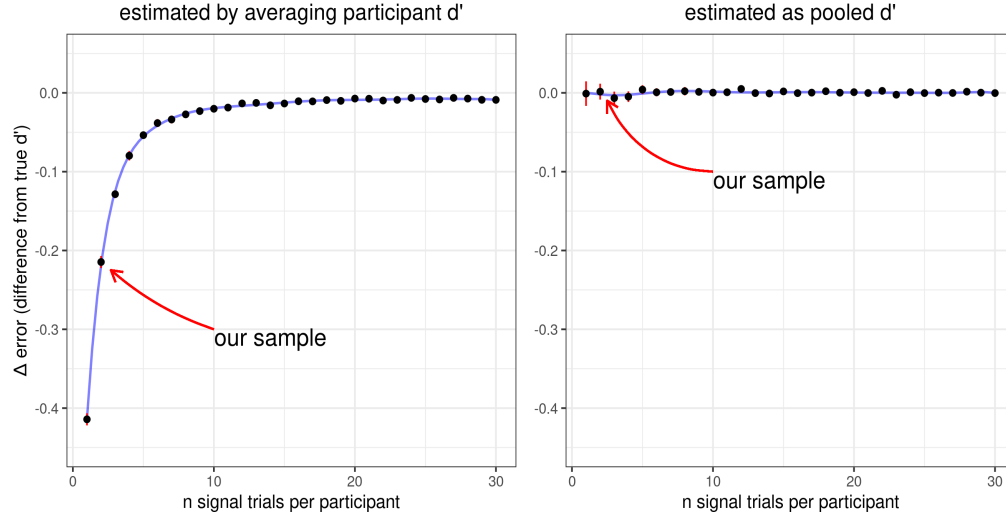
SI Figure 4. The timing of participation differs across children and adults across five high-recruitment countries. The polar histograms show the relative proportion of the cohort that participated at a particular time of day (on a 24-hour clock); the two shades differentiate an arbitrary threshold of ‘day’ versus ‘late night/early morning’ participation (see legend). The bar plots show the proportions between these two cutoffs in the data. The distributions of participation times differ markedly across the cohorts, with general consistency across countries within each cohort. This suggests that the participants who self-report being children are, in fact, children, as they are less likely to participate late at night. Note that four of the plotted countries have a single time zone; the United States data is aggregated across three time zones, but is included here as that country had the largest samples of children and adults.



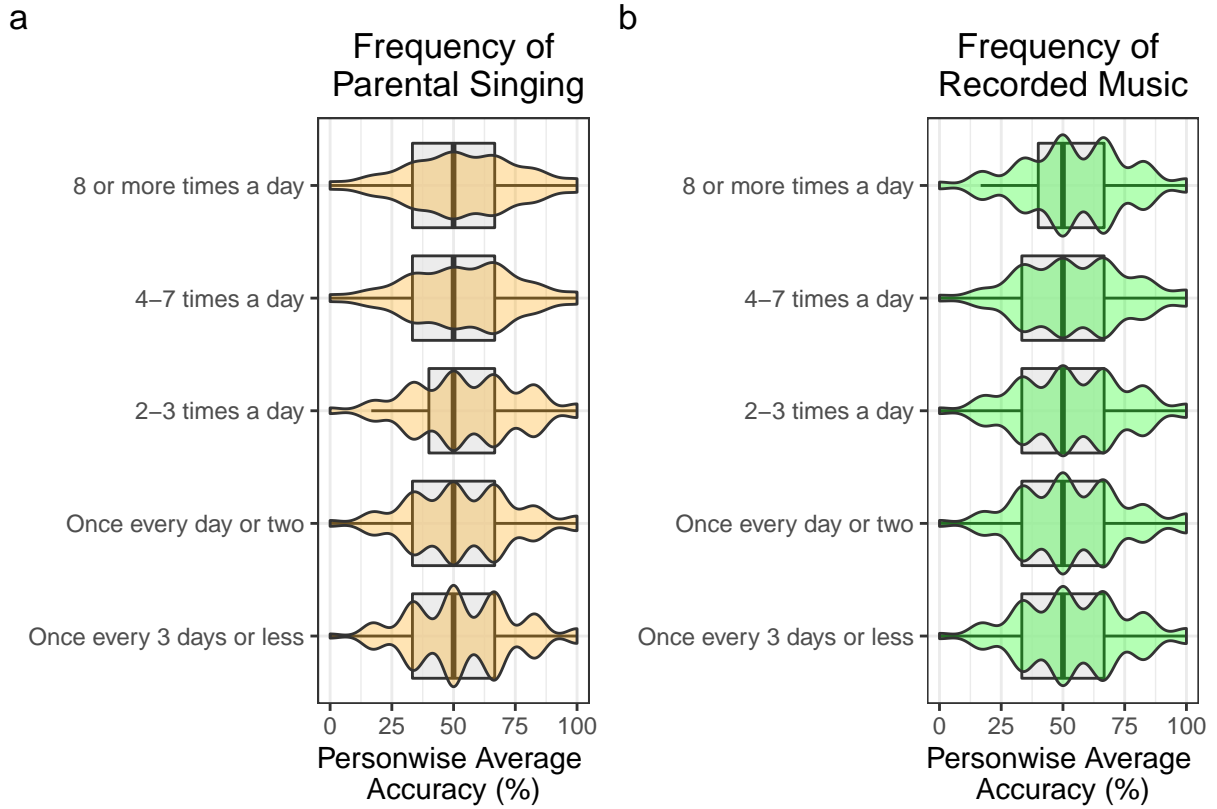
SI Figure 5. Children’s response times, aggregated across all trials, as a function of age. Response times were comparable across song types. Note that the y -axis is on a log scale.



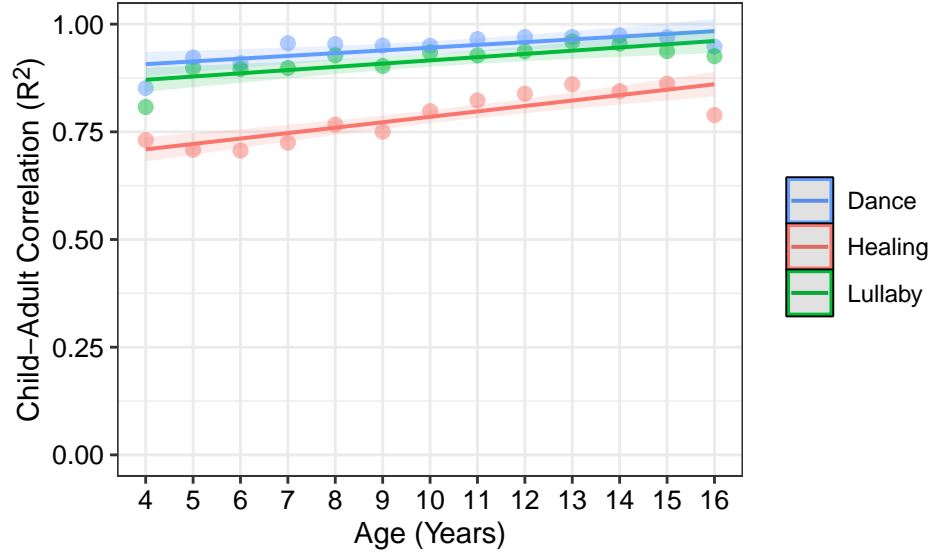
SI Figure 6. Response biases and criterion scores for the children were stable across ages. **(a)** The stacked-bar graphs represent the proportion of responses for each of the three song categories over age. The colors of the bars represent the response types; evidently, they were not evenly used. **(b)** We estimated criterion values for each age group, where higher criterion denotes more conservative guessing and lower criterion values denote more liberal guessing. The circles represent the criterion estimates; the lines depict a linear regression for each song type; and the shaded regions represent the 95% confidence intervals from each regression. None of the regressions were statistically significant ($p > 0.05$). Participants required the strongest evidence to make a lullaby guess and required the least evidence to make a dance guess.



SI Figure 7. A simulation shows that the pooled d-prime provides a less biased estimate of sensitivity than does an average of participant-level d-prime estimates. The typical approach for calculating cohort-level sensitivity is to first calculate d-prime scores for each participant, and then take their average. However, our experiment had a small number of trials per participant (in part because young children have short attention spans); this meant that there were limited numbers of signal and noise trials for each participant, rendering the participant-level estimates difficult to interpret. In cases like these, Macmillan & Kaplan (1985) advocate for a collapsed/pooled d-prime approach. To test whether our particularly large sample size (~5000 children) would compensate for the deficiencies of sparse within-subject sampling, we ran a simulation to test whether averaged d-prime or pooled d-prime would be optimal. The simulation used the same number of participants in our study and simulated the ability of each approach to estimate a true effect size over a range of within-subject trial counts. The two plots show the estimated size of error in each method (on the y -axes), at differing numbers of signal trials per participant (on the x -axes). The results show that the averaging d-prime approach underestimates the true effect by ~20%, and the pooled approach is unbiased. We therefore opted to use the pooled d-prime approach in the main analyses.



SI Figure 8. Children’s accuracy of classifying songs was not related to their in-home musical experiences: across different levels of (a) the frequency of parental singing and (b) the frequency of recorded music listening, person-wise accuracy did not change. The thick lines represent medians; the boxes represent interquartile ranges; the lines represent ranges; and the shaded areas depict kernel density estimates.



SI Figure 9. The similarity between children’s and adult’s inferences is high and modestly increases through childhood. We measured the similarity between children’s and adult’s inferences as the proportion of variance explained (R^2) by regressing the average song-wise inferences of children on the adults, for each of the 13 yearlong age bins. We then constructed a regression model that predicted these R^2 values from the interaction between the guessed category for each song and age of the children, producing the regression lines shown in the figure for each response type (dance, lullaby, or healing song). The shaded areas representing 95% confidence interval and the points representing the raw values. The main effect of age was statistically significant ($p < 0.001$) but the interaction of age with song-type was not ($p = 0.09$).

Supplementary Tables

Children		Adults	
Language	<i>n</i>	Language	<i>n</i>
English	2,820	English	52,630
Turkish	305	Spanish	11,324
Spanish	286	Turkish	5,099
German	123	German	4,350
Chinese/Mandarin	119	French	3,122
Chinese/Cantonese/Yue	63	Polish	3,048
French	62	Portuguese	2,917
Italian	61	Dutch	2,754
Polish	48	Danish	2,590
Russian	48	Russian	2,288
Indonesian	46	Swedish	2,281
Vietnamese	44	ChineseMandarin	1,750
Portuguese	42	Italian	1,645
Arabic	33	Indonesian	1,093
Dutch	33	ChineseCantoneseYue	1,065
Tamil	29	Norwegian	939
Swedish	26	Finnish	896
Tagalog	26	Greek	831
Hebrew	25	Romanian	788
Romanian	25	Tagalog	768
Hindi	24	Arabic	716
Catalan	18	Hindi	654
Chinese/Taiwanese/Min	17	Hungarian	610
Nan/Hokkien			
Greek	17	Croatian	533
Korean	17	Vietnamese	509
Norwegian	17	Czech	500
Khmer	16	Japanese	463
Malayalam	16	Other	431
Other	16	Serbian	421
Czech	15	Hebrew	382

SI Table 1 The native languages of participants, reported in the child and adult cohorts. For brevity, only the 30 most-commonly reported languages are displayed here.

Children		Adults	
Country	<i>n</i>	Country	<i>n</i>
United States	1,943	United States	32,848
Canada	350	United Kingdom	7,119
Turkey	302	Canada	4,956
United Kingdom	257	Turkey	4,661
Australia	194	Australia	4,164
Germany	118	Germany	3,978
Spain	82	Spain	3,042
Singapore	75	Denmark	2,700
China	72	Argentina	2,627
New Zealand	69	Poland	2,607
Argentina	63	Sweden	2,365
Italy	63	France	2,057
Hong Kong	59	Brazil	1,861
Indonesia	58	India	1,735
India	52	Mexico	1,711
Malaysia	51	Italy	1,502
France	42	The Netherlands	1,499
Philippines	42	Russia	1,415
Poland	42	Singapore	1,394
Mexico	41	Belgium	1,087
Sweden	37	Indonesia	1,041
Russian Federation	36	Norway	1,006
Brazil	32	Colombia	996
Netherlands	32	Philippines	970
Vietnam	32	Chile	948
Colombia	29	Finland	934
Taiwan	28	China	883
Austria	25	Kingdom of the Netherlands	834
Israel	25	Portugal	810
Chile	24	Malaysia	765

SI Table 2 The countries where participants were located, reported in the child and adult cohorts. For brevity, only the 30 most-commonly reported countries are displayed here.

Song ID	Type	Society	Country (present-day)	Language
1	Lullaby	Lardil	Australia	Aranda
2	Healing	Anggor	Papua New Guinea	Anggor
3	Lullaby	Blackfoot	USA	Blackfoot
4	Healing	Uttar Pradesh	India	Andamans
5	Dance	Emberá	Colombia	Emberá
6	Healing	Akan	Côte d'Ivoire	Akan
7	Lullaby	Nahua	Mexico	Maya (Yucatan Peninsula)
8	Healing	Yapese	Federated States of Micronesia	Chuuk
9	Dance	Yolngu	Australia	Aranda
10	Dance	Aymara	Chile	Aymara
11	Dance	Blackfoot	USA	Blackfoot
12	Lullaby	Saami	Norway	Saami
13	Dance	Bahia Brazilians	Brazil	Bahia Brazilians
14	Healing	Seri	Mexico	Huichol
15	Healing	Ye'kuana	Venezuela	Javaé
16	Dance	Aka	Central African Republic	Aka
17	Healing	Lunda	Democratic Republic of the Congo	Bemba
18	Dance	Greeks	Greece	Georgia
19	Dance	Sweden	Sweden	Saami
20	Healing	Iroquois	Canada	Iroquois
21	Dance	Iroquois	New York	Iroquois
22	Healing	Maya (Yucatan Peninsula)	Mexico	Maya (Yucatan Peninsula)
23	Lullaby	Goajiro	Venezuela	Chachi
24	Dance	Gourara	Algeria	Gourara
25	Dance	Ojibwa	USA	Copper Inuit
26	Lullaby	Javaé	Brazil	Javaé
27	Healing	Bahia Brazilians	Brazil	Bahia Brazilians
28	Lullaby	Chewa	Malawi	Bemba
29	Lullaby	Aka	Central African Republic	Aka
30	Healing	Georgia	East Georgia	Georgia
31	Dance	Ainu	Japan	Ainu
32	Lullaby	Highland Scots	United Kingdom	Highland Scots
33	Lullaby	Seri	Mexico	Huichol
34	Dance	Garo	India	Andamans
35	Lullaby	Kurds	Iran	Kurds
36	Lullaby	Fut	Cameroon	Akan
37	Dance	Hawaiians	USA	Hawaiians
38	Lullaby	Samoans	Samoa	Hawaiians
39	Healing	Kuna	Panama	Emberá
40	Lullaby	Marathi	India	Andamans
41	Healing	Pawnee	USA	Blackfoot
42	Lullaby	Guarani	Bolivia	Bahia Brazilians
43	Healing	Meratus	Indonesia	Central Thai
44	Lullaby	Ainu	Japan	Ainu
45	Dance	Maasai	Tanzania	Amhara
46	Healing	Korea	South Korea	Ainu
47	Healing	Turkmen	Iran	Kurds
48	Lullaby	Chukchee	Russia	Chukchee
49	Dance	Lozi	Zambia	Bemba
50	Lullaby	Kanaks	New Caledonia	Anggor
51	Lullaby	Greeks	Greece	Georgia
52	Dance	Tlingit	USA	Haida
53	Lullaby	Tuareg	Mali	Gourara
54	Healing	Kwakwaka'wakw	Canada	Haida
55	Lullaby	Nyangatom	Ethiopia	Amhara
56	Dance	Chachi	Ecuador	Chachi
57	Dance	Chuuk	Federated States of Micronesia	Chuuk
58	Healing	Quechan	USA	Hopi
59	Healing	Hawaiians	USA	Hawaiians
60	Dance	Kanaks	New Caledonia	Anggor
61	Healing	Tunisians	Tunisia	Gourara
62	Dance	Kurds	Iran	Kurds
63	Lullaby	Haida	Canada	Haida
64	Dance	Mataco	Argentina	Alacaluf
65	Dance	Saramaka	Suriname	Javaé
66	Dance	Mentawaians	Indonesia	Central Thai
67	Dance	Ewe	Ghana	Akan
68	Lullaby	Kuna	Panama	Emberá
69	Lullaby	Ona	Argentina	Alacaluf

(continued)

Song ID	Type	Society	Country (present-day)	Language
70	Healing	Ojibwa	USA	Copper Inuit
71	Dance	Tzeltal	Mexico	Maya (Yucatan Peninsula)
72	Lullaby	Iglulik Inuit	Canada	Copper Inuit
73	Lullaby	Q'ero Quichua	Peru	Aymara
74	Lullaby	Iroquois	Canada	Iroquois
75	Healing	Nanai	Russia	Chukchee
76	Dance	Highland Scots	United Kingdom	Highland Scots
77	Dance	Chukchee	Russia	Chukchee
78	Healing	Mataco	Argentina	Alacaluf
79	Healing	Walbiri	Central and North-Central Australia	Aranda
80	Healing	Kogi	Colombia	Chachi
81	Dance	Hopi	USA	Hopi
82	Lullaby	Hopi	USA	Hopi
83	Healing	Otavalo Quichua	Ecuador	Aymara
84	Healing	Ganda	Uganda	Amhara
85	Lullaby	Yapese	Federated States of Micronesia	Chuuk
86	Dance	Yaqui	Mexico	Huichol
87	Healing	Mbuti	Congo	Aka
88	Lullaby	Phunoi	Laos	Central Thai

SI Table 3 Summary information about the songs children heard.

Feature	Type	Selection	Definition	Scoring
Accent	Expert Annotation	Dance, Lullaby	The differentiation of musical pulses, usually by volume or emphasis of articulation. A fluid, gentle song will have few accents and a correspondingly low value.	"none" (0); "a little" (0.5); or "a lot" (1).
Tempo	Expert Annotation	Dance, Lullaby, Healing	The rate of salient rhythmic pulses (measured by having annotators tap the beat), adjusted to correspond with 'quarter note' equivalent; measured in beats per minute; the perceived speed of the music. A fast song will have a high value.	range: 61-205
Metrical stability	Expert Annotation	Dance Lullaby, Healing	The perceived clarity of the metrical rhythmic structure (groupings of perceived strong and weak beats).	Scale from 1 - 6, ranging from "No macrometer" (1) to "Totally clear macrometer" (6)
Melodic simplicity	Transcription	Lullaby, Healing	Variety versus monotony of the melody, measured as the proportion of notes in the transcription corresponding to the most common pitch-class.	range: 0.15 - 1
Vibrato	Expert Annotation	Lullaby, Healing	Presence of vibrato in the singing. Vibrato is a rapid, slight oscillation in pitch that is typically perceived as adding intensity.	"not present" (0); "present" (1)
Syncopation	Expert Annotation	Healing	The degree of of syncopation in the song (roughly equateable with 'rhythmic complexity'). Syncopation is when rhythmic accents align with metrically weak positions of a metrical beat structure (e.g., having accents half way between when people tap their foot rather than directly aligned with the foot taps).	"none" (0); "a little" (0.5); or "a lot" (1).
Maj (-) / min (-) tonality	Transcription	Healing, Lullaby	Tonal quality or mode of the transcription (major or minor) based on the Krumhansl-Schmuckler key-finding algorithm.	"Major" (0); "Minor" (1)
Ornamentation	Expert Annotation	Lullaby	Melodic embellishment by the singer, usually via rapid 'passing notes' that decorate a simpler underlying melodic structure.	"not present" (0); "present" (1)
Note Density	Transcription	Lullaby	The average number of notes per second in a given song	range: 0.44 - 6.82
Note Duration	Transcription	Dance, Lullaby	The average duration per note, in seconds.	range: 0.17 - 2.04
Intervallic distinctiveness	Transcription	Lullaby, Healing	Fraction of melodic intervals that belong to the second most common interval divided by the fraction of melodic intervals belonging to the most common interval; 0 if there are not two distinct most common melodic intervals.	range: 0 - 1
Duple rhythm	Expert Annotation	Dance, Lullaby	The presence of duple subdivisions of the beat	range: 0 - 1

SI Table 4 Musical features selected by LASSO procedure. The features were derived from either *Transcription* or *Expert Annotation* datasets from Mehr et al. (2019). The "Selection" column denotes which of the three song types the feature was selected for. The "Scoring" column denotes the original scale of the feature score (note, however, that all features were z-scored in the analysis).

Predictor	Children				Adults (relative to Children)			
	β	SE	t	p	β	SE	t	p
Intercept	0.346	0.019	18.697	<0.001	0.075	0.026	2.881	0.004
Accentuation	0.137	0.021	6.377	<0.001	0.026	0.030	0.861	0.391
Tempo	0.138	0.022	6.313	<0.001	0.000	0.031	-0.008	0.994
Metrical clarity	0.145	0.020	7.283	<0.001	-0.003	0.028	-0.124	0.902
Duple rhythm	0.061	0.019	3.156	0.002	0.008	0.027	0.288	0.774
Note duration	0.063	0.021	2.959	0.004	0.006	0.030	0.195	0.846

SI Table 5 Musical features influencing children’s “dance” guesses, compared to those of adults. The LASSO-selected correlates for children are listed on the left side of the table; those of adults are on the right side and are tested relative to the children (i.e., from the interaction terms in the model). No musical features show a significant difference between children and adults, indicating that the two groups’ inferences are guided similarly by the musical features studied here.

Predictor	Children				Adults (relative to Children)			
	β	SE	t	p	β	SE	t	p
Intercept	0.274	0.017	15.772	<0.001	-0.058	0.025	-2.354	0.020
Accentuation	-0.131	0.022	-5.921	<0.001	0.021	0.031	0.663	0.508
Melodic simplicity	-0.091	0.022	-4.164	<0.001	0.012	0.031	0.395	0.694
Tempo	-0.053	0.023	-2.311	0.022	0.004	0.032	0.137	0.892
Note duration	-0.096	0.024	-3.991	<0.001	0.020	0.034	0.585	0.560
Maj(-)/min(+) tonality	0.047	0.019	2.511	0.013	-0.005	0.026	-0.207	0.836
Ornamentation	-0.054	0.023	-2.401	0.018	0.002	0.032	0.073	0.942
Intervallic distinctiveness	-0.047	0.020	-2.349	0.020	-0.001	0.028	-0.043	0.966
Note density	-0.068	0.027	-2.508	0.013	0.015	0.039	0.393	0.695
Metrical clarity	-0.064	0.021	-3.027	0.003	0.013	0.030	0.438	0.662
Vibrato	-0.030	0.022	-1.386	0.168	0.019	0.031	0.596	0.552
Duple rhythm	-0.043	0.019	-2.300	0.023	0.007	0.026	0.269	0.789

SI Table 6 Musical features influencing children’s “lullaby” guesses, compared to those of adults. The LASSO-selected correlates for children are listed on the left side of the table; those of adults are on the right side and are tested relative to the children (i.e., from the interaction terms in the model). No musical features show a significant difference between children and adults, indicating that the two groups’ inferences are guided similarly by the musical features studied here.

Predictor	Children				Adults (relative to Children)			
	β	SE	t	p	β	SE	t	p
Intercept	0.379	0.016	24.159	<0.001	-0.018	0.022	-0.790	0.431
Metrical clarity	-0.070	0.018	-3.984	<0.001	-0.011	0.025	-0.447	0.655
Tempo	-0.058	0.018	-3.176	0.002	-0.017	0.026	-0.640	0.523
Maj(-)/min(+) tonality	-0.056	0.016	-3.502	<0.001	0.017	0.023	0.740	0.461
Syncopation	-0.018	0.018	-1.011	0.314	-0.014	0.025	-0.559	0.577
Vibrato	0.057	0.017	3.448	<0.001	-0.031	0.023	-1.339	0.182
Melodic simplicity	0.064	0.017	3.849	<0.001	-0.021	0.024	-0.867	0.387

SI Table 7 Musical features influencing children’s “healing” guesses, compared to those of adults. The LASSO-selected correlates for children are listed on the left side of the table; those of adults are on the right side and are tested relative to the children (i.e., from the interaction terms in the model). No musical features show a significant difference between children and adults, indicating that the two groups’ inferences are guided similarly by the musical features studied here.

Predictor	Children				Objective (relative to Children)			
	β	SE	t	p	β	SE	t	p
Intercept	0.422	0.014	31.204	<0.001	-0.082	0.019	-4.312	<0.001
Accentuation	0.163	0.016	10.407	<0.001	-0.101	0.022	-4.532	<0.001
Tempo	0.137	0.016	8.636	<0.001	-0.054	0.023	-2.403	0.017
Metrical clarity	0.142	0.015	9.741	<0.001	-0.071	0.021	-3.447	<0.001
Duple rhythm	0.068	0.014	4.882	<0.001	-0.072	0.020	-3.610	<0.001
Note duration	0.069	0.016	4.433	<0.001	-0.069	0.022	-3.139	0.002

SI Table 8 Musical features influencing children’s “dance” guesses, compared to the features that objectively distinguish song types from one another. The LASSO-selected correlates for children are listed on the left side of the table; those that reliably correlate with song types across world regions are on the right side and are tested relative to the children (i.e., from the interaction terms in the model). Contrasting SI Tables 5-7, quite a few features differ significantly, indicating that in many cases, children’s intuitions (erroneously) deviate from the objective differences between song types.

Predictor	Children				Objective (relative to Children)			
	β	SE	t	p	β	SE	t	p
Intercept	0.216	0.011	19.128	<0.001	0.124	0.016	7.768	<0.001
Accentuation	-0.110	0.014	-7.660	<0.001	-0.016	0.020	-0.766	0.445
Melodic simplicity	-0.079	0.014	-5.543	<0.001	0.078	0.020	3.890	<0.001
Tempo	-0.048	0.015	-3.256	0.001	0.015	0.021	0.713	0.477
Note duration	-0.076	0.016	-4.864	<0.001	0.079	0.022	3.566	<0.001
Maj(-)/min(+) tonality	0.041	0.012	3.410	<0.001	-0.032	0.017	-1.855	0.066
Melodic ornamentation	-0.052	0.015	-3.533	<0.001	0.049	0.021	2.384	0.018
Intervallic distinctiveness	-0.048	0.013	-3.705	<0.001	0.048	0.018	2.629	0.009
Note density	-0.053	0.018	-3.001	0.003	0.052	0.025	2.059	0.041
Metrical clarity	-0.051	0.014	-3.703	<0.001	0.017	0.019	0.891	0.374
Vibrato	-0.012	0.014	-0.835	0.405	0.014	0.020	0.704	0.482
Duple meter	-0.036	0.012	-2.952	0.004	0.040	0.017	2.318	0.022

SI Table 9 Musical features influencing children’s “lullaby” guesses, compared to the features that objectively distinguish song types from one another. The LASSO-selected correlates for children are listed on the left side of the table; those that reliably correlate with song types across world regions are on the right side and are tested relative to the children (i.e., from the interaction terms in the model). Contrasting SI Tables 5-7, quite a few features differ significantly, indicating that in many cases, children’s intuitions (erroneously) deviate from the objective differences between song types.

Predictor	Children				Objective (relative to Children)			
	β	SE	t	p	β	SE	t	p
Intercept	0.362	0.011	31.795	<0.001	-0.042	0.016	-2.604	0.010
Metrical clarity	-0.082	0.013	-6.370	<0.001	0.052	0.018	2.841	0.005
Tempo	-0.075	0.013	-5.632	<0.001	0.043	0.019	2.283	0.024
Maj(-)/min(+) tonality	-0.040	0.012	-3.388	<0.001	0.027	0.017	1.624	0.106
Sycopation	-0.032	0.013	-2.486	0.014	0.057	0.018	3.108	0.002
Vibrato	0.026	0.012	2.146	0.033	-0.015	0.017	-0.881	0.380
Melodic simplicity	0.044	0.012	3.619	<0.001	-0.029	0.017	-1.681	0.095

SI Table 10 Musical features influencing children’s “healing” guesses, compared to the features that objectively distinguish song types from one another. The LASSO-selected correlates for children are listed on the left side of the table; those that reliably correlate with song types across world regions are on the right side and are tested relative to the children (i.e., from the interaction terms in the model). Contrasting SI Tables 5-7, quite a few features differ significantly, indicating that in many cases, children’s intuitions (erroneously) deviate from the objective differences between song types.

References

- Bainbridge, C. M., Bertolo, M., Youngers, J., Atwood, S., Yurdum, L., Simson, J., Lopez, K., Xing, F., Martin, A., & Mehr, S. A. (2021). Infants relax in response to unfamiliar foreign lullabies. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-020-00963-z>
- Balch, W. R., & Lewis, B. S. (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. N., Irwin, J. M., & Tidball, G. (1996). Infants' Reliance on a Social Criterion for Establishing Word-Object Relations. *Child Development*, 20.
- Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, 17(1), 43–64. <https://doi.org/10.2307/40285811>
- Bergeson, T., & Trehub, S. (1999). Mothers' singing to infants and preschool children. *Infant Behavior and Development*, 22(1).
- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, 113(39), 10818–10823. <https://doi.org/10.1073/pnas.1605782113>
- Bohn, M., & Frank, M. C. (2019). The Pervasive Role of Pragmatics in Early Language. *Annual Review of Developmental Psychology*, 29.
- Bonneville-Roussy, A., Rentfrow, P. J., Xu, M. K., & Potter, J. (2013). Music through the ages: Trends in musical engagement and preferences from adolescence through middle adulthood. *Journal of Personality and Social Psychology*, 105(4), 703–717. <https://doi.org/10.1037/a0033770>
- Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and Early Language Acquisition. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00327>
- Cheever, T., Taylor, A., Finkelstein, R., Edwards, E., Thomas, L., Bradt, J., Holochwost, S. J., Johnson, J. K., Limb, C., Patel, A. D., Tottenham, N., Iyengar, S., Rutter, D., Fleming, R., & Collins, F. S. (2018). NIH/Kennedy Center workshop on music and the brain: Finding harmony. *Neuron*, 97(6), 1214–1218. <https://doi.org/10.1016/j.neuron.2018.02.004>
- Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, 8(1), 8659. <https://doi.org/10.1038/s41598-018-26889-1>
- Cirelli, L. K., Jurewicz, Z. B., & Trehub, S. E. (2019). Effects of maternal singing style on mother–infant arousal and behavior. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_01402
- Cowen, A. S., Laukka, P., Elfenbein, H. A., Liu, R., & Keltner, D. (2019). The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nature Human Behaviour*, 3(4), 369–382. <https://doi.org/10.1038/s41562-019-0533-6>
- Darwin, C. (1871). *The descent of man*. Watts & Co.
- Drake, C., & Heni, J. B. E. (2003). Synchronizing with Music: Intercultural Differences. *Annals of the New York Academy of Sciences*, 999(1), 429–437. <https://doi.org/10.1196/annals.1284.053>
- 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). <https://doi.org/10.1098/ospf.io/rspb.2017.0990>
- Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, 11.
- Fitch, W. T., Neubauer, J., & Herzel, H. (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, 63(3), 407–418. <https://doi.org/10.1006/anbe.2001.1912>
- Friedman, J., Hastie, T., & Tibshirani, R. (2016). *Lasso and elastic-net regularized generalized linear models*. *Rpackage version 2.0-5*.
- 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. <https://doi.org/10.3758/s13423-012-0296-9>
- Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A.-Y., & DeCasper, A. J. (2011). A melodic contour

- repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. *PLoS ONE*, 6(2). <https://doi.org/10.1371/journal.pone.0017304>
- Hale, S. (1990). A Global Developmental Trend in Cognitive Processing Speed. *Child Development*, 61(3), 653–663.
- Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, 16(1), 48–55.
- Hartshorne, J. K., de Leeuw, J., Goodman, N., Jennings, M., & O'Donnell, T. J. (2019). A thousand studies for the price of one: Accelerating psychological science with Pushkin. *Behavior Research Methods*, 51, 1782–1803. <https://doi.org/10.3758/s13428-018-1155-z>
- Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317(5843), 1360–1366. <https://doi.org/10.1126/science.1146282>
- Hilton, C. B., Asano, R., & Boeckx, C. A. (2020). Why musical hierarchies? *Behavioral and Brain Sciences*.
- Hilton, C. B., & Mehr, S. A. (2022). Citizen science can help to alleviate the generalizability crisis. *Behavioral and Brain Sciences*.
- Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Bainbridge, C. M., Atwood, S., Simson, J., Knox, D., Glowacki, L., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., Schachner, A., ... Mehr, S. A. (2021). Acoustic regularities in infant-directed speech and song across cultures. *bioRxiv*. <https://doi.org/10.1101/2020.04.09.032995>
- Holbrook, M. B., & 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. <https://doi.org/10.1177/0305735690182004>
- Huber, B., & Gajos, K. Z. (2020). Conducting online virtual environment experiments with uncompensated, unsupervised samples. *PLOS ONE*, 15(1), e0227629. <https://doi.org/10.1371/journal.pone.0227629>
- Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of Musical Tempo and Mode on Arousal, Mood, and Spatial Abilities. *Music Perception*, 20(2), 151–171. <https://doi.org/10.1525/mp.2002.20.2.151>
- Ilie, G., & Thompson, W. F. (2006). A Comparison of Acoustic Cues in Music and Speech for Three Dimensions of Affect. *Music Perception*, 23(4), 319–330. <https://doi.org/10.1525/mp.2006.23.4.319>
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109(1), 54–65. <https://doi.org/10.1016/j.cognition.2008.07.015>
- Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & McDermott, J. H. (2019). Universal and non-universal features of musical pitch perception revealed by singing. *Current Biology*, 29(19), 3229–3243. <https://doi.org/10.1016/j.cub.2019.08.020>
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54–75. <https://doi.org/10.1037/a0024208>
- Kamiloğlu, R. G., Slocombe, K. E., Haun, D. B., & Sauter, D. A. (2020). Human listeners' perception of behavioural context and core affect dimensions in chimpanzee vocalizations. *Proceedings of the Royal Society B*, 287(1929), 20201148.
- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, 104(30), 12577–12580.
- Kiselev, S., Espy, K. A., & Sheffield, T. (2009). Age-related differences in reaction time task performance in young children. *Journal of Experimental Child Psychology*, 102(2), 150–166. <https://doi.org/10.1016/j.jecp.2008.02.002>
- Krumhansl, C. L. (2010). Plink: "Thin Slices" of Music. *Music Perception*, 27(5), 337–354. <https://doi.org/10.1525/mp.2010.27.5.337>
- Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, 58(6), 927–935. <https://doi.org/10.3758/BF03205494>
- Li, W., Germine, L. T., Mehr, S. A., Srinivasan, M., & Hartshorne, J. K. (submitted). *Building a generalizable and replicable developmental psychology with citizen science*.
- Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The Origins of Social Categorization. *Trends in Cognitive Sciences*, 21(7), 556–568. <https://doi.org/10.1016/j.tics.2017.04.004>
- Liu, J., Hilton, C. B., Bergelson, E., & Mehr, S. A. (2021). Language experience shapes music processing

- across 40 tonal, pitch-accented, and non-tonal languages. *bioRxiv*. <https://doi.org/10.1101/2021.10.18.464888>
- Lomax, A. (1968). *Folk song style and culture*. American Association for the Advancement of Science.
- Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*, 52(6), 599–608. <https://doi.org/10.3758/BF03211696>
- Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, Experience, and Music Perception. *Psychological Science*, 1(4), 272–276. <https://doi.org/10.1111/j.1467-9280.1990.tb00213.x>
- Ma, W., & Thompson, W. F. (2015). Human emotions track changes in the acoustic environment. *Proceedings of the National Academy of Sciences*, 112(47), 14563–14568. <https://doi.org/10.1073/pnas.1515087112>
- Macmillan, N. A., & Kaplan, H. L. (1985). Detection Theory Analysis of Group Data: Estimating Sensitivity From Average Hit and False-Alarm Rates. *Psychological Bulletin*, 15.
- Margulis, E. H., Wong, P. C. M., Turnbull, C., Kubit, B. M., & McAuley, J. D. (2022). Narratives imagined in response to instrumental music reveal culture-bounded intersubjectivity. *Proceedings of the National Academy of Sciences*, 119(4), e2110406119. <https://doi.org/10.1073/pnas.2110406119>
- Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12months. *Cognition*, 123(1), 50–60. <https://doi.org/10.1016/j.cognition.2011.12.003>
- Mehr, S. A. (2014). Music in the home: New evidence for an intergenerational link. *Journal of Research in Music Education*, 62(1), 78–88. <https://doi.org/10.1177/0022429413520008>
- Mehr, S. A., & Krasnow, M. M. (2017). Parent-offspring conflict and the evolution of infant-directed song. *Evolution and Human Behavior*, 38(5), 674–684. <https://doi.org/10.1016/j.evolhumbehav.2016.12.005>
- 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. <https://doi.org/10.1017/S0140525X20000345>
- Mehr, S. A., Krasnow, M., Bryant, G. A., & Hagen, E. H. (2021). Origins of music in credible signaling. *Behavioral and Brain Sciences*. <https://doi.org/10.31234/osf.io/nrqb3>
- 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. <https://doi.org/10.1126/science.aax0868>
- 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. <https://doi.org/10.1016/j.cub.2017.12.042>
- Mehr, S. A., Song, L. A., & Spelke, E. S. (2016). For 5-month-old infants, melodies are social. *Psychological Science*, 27(4), 486–501. <https://doi.org/10.1177/0956797615626691>
- Mehr, S. A., & Spelke, E. S. (2017). Shared musical knowledge in 11-month-old infants. *Developmental Science*, 21(2).
- Mendoza, J. K., & Fausey, C. M. (2021). Everyday music in infancy. *Developmental Science*. <https://doi.org/10.31234/osf.io/sqatb>
- 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.
- Nettl, B. (2005). *The Study of Ethnomusicology* (2nd Edition). University of Illinois Press.
- 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.
- Nuckolls, J. B. (1999). The Case for Sound Symbolism. *Annual Review of Anthropology*, 28(1), 225–252. <https://doi.org/10.1146/annurev.anthro.28.1.225>
- 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*, 10(2), 58–71. <https://doi.org/10.1002/evan.1014>
- Patten, K. J., McBeath, M. K., & Baxter, L. C. (2018). Harmonicity: Behavioral and Neural Evidence for Functionality in Auditory Scene Analysis. *Auditory Perception & Cognition*, 1(3-4), 150–172. <https://doi.org/10.1080/25742442.2019.1609307>
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758–4763. <https://doi.org/10.1073/pnas.0909074107>
- Peretz, I., Gaudreau, D., & Bonnel, A.-M. (1998). Exposure effects on music preference and recognition.

- Memory & Cognition*, 26(5), 884–902. <https://doi.org/10.3758/BF03201171>
- Perlman, M., Paul, J., & Lupyan, G. (2021). Vocal communication of magnitude across language, age, and auditory experience. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001103>
- Phillips-Silver, J. (2005). Feeling the Beat: Movement Influences Infant Rhythm Perception. *Science*, 308(5727), 1430–1430. <https://doi.org/10.1126/science.1110922>
- Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, 110(41), E3965–E3972. <https://doi.org/10.1073/pnas.1304326110>
- Reinecke, K., & Gajos, K. Z. (2015). *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*.
- Rohrmeier, M., & Rebuschat, P. (2012). Implicit Learning and Acquisition of Music. *Topics in Cognitive Science*, 4(4), 525–553. <https://doi.org/10.1111/j.1756-8765.2012.01223.x>
- Saffran, J. R., Loman, M. M., & Robertson, R. R. W. (2000). Infant memory for musical experiences. *Cognition*, 77(1), B15–B23. [https://doi.org/10.1016/S0010-0277\(00\)00095-0](https://doi.org/10.1016/S0010-0277(00)00095-0)
- Schellenberg, E. G., Peretz, I., & Viellard, S. (2008). Liking for happy- and sad-sounding music: Effects of exposure. *Cognition & Emotion*, 22(2), 218–237. <https://doi.org/10.1080/02699930701350753>
- Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology*, 32(1), 76–92. <https://doi.org/10.1177/0022022101032001009>
- Schubert, E. (2004). Modeling Perceived Emotion With Continuous Musical Features. *Music Perception*, 21(4), 561–585. <https://doi.org/10.1525/mp.2004.21.4.561>
- Sheskin, M., Scott, K., Mills, C. M., Bergelson, E., Bonawitz, E., Spelke, E. S., Fei-Fei, L., Keil, F. C., Gweon, H., Tenenbaum, J. B., Jara-Ettinger, J., Adolph, K. E., Rhodes, M., Frank, M. C., Mehr, S. A., & Schulz, L. (2020). Online developmental science to foster innovation, access, and impact. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2020.06.004>
- Sievers, B., Lee, C., Haslett, W., & Wheatley, T. (2019). A multi-sensory code for emotional arousal. *Proceedings of the Royal Society B*, 286(1906).
- 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. <https://doi.org/10.1073/pnas.1209023110>
- Singh, M. (2018). The cultural evolution of shamanism. *Behavioral and Brain Sciences*, 41, 1–62.
- 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(05), 675–691. <https://doi.org/10.1017/S0140525X05000129>
- 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. <https://doi.org/10.1111/nyas.12622>
- 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. (1993a). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, 16(2), 193–211. [https://doi.org/10.1016/0163-6383\(93\)80017-3](https://doi.org/10.1016/0163-6383(93)80017-3)
- Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, 16(3), 285–295.
- Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, 17(6), 872–879. <https://doi.org/10.1111/desc.12170>
- Weninger, F., Eyben, F., Schuller, B. W., Mortillaro, M., & Scherer, K. R. (2013). On the Acoustics of Emotion in Audio: What Speech, Music, and Sound have in Common. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00292>
- 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. <https://doi.org/10.1073/pnas.0809035106>
- 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*, 9(4). <https://doi.org/10.1371/journal.pone.0094446>
- Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2017). Older but not younger infants

- associate own-race faces with happy music and other-race faces with sad music. *Developmental Science*. <https://doi.org/10.1111/desc.12537>
- Yamamoto, M., Naga, S., & Shimizu, J. (2007). Positive musical effects on two types of negative stressful conditions. *Psychology of Music*, 35(2), 249–275. <https://doi.org/10.1177/0305735607070375>
- Yan, R., Jessani, G., Spelke, E., Villiers, P. de, Villiers, J. de, & Mehr, S. (2021). Across demographics and recent history, most parents sing to their infants and toddlers daily. *Philosophical Transactions of the Royal Society B: Biological Sciences*. <https://doi.org/10.31234/osf.io/fy5bh>
- Yu, C. S.-P., McBeath, M. K., & Glenberg, A. M. (2021). The gleam-glum effect: /i:/ versus /λ/ phonemes generically carry emotional valence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <https://doi.org/10.1037/xlm0001017>
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, 107(13), 5768–5773. <https://doi.org/10.1073/pnas.1000121107>