Title: Humans share acoustic preferences with other animals

Authors: Logan S James[1,2,3,4,5]*, Sarah C Woolley[1,2], Jon T Sakata[1,2], Courtney B Hilton[5,6], Michael J Ryan[3,4], Samuel A Mehr[5,6]

Affiliations:

[1]Department of Biology, McGill University; Montreal, Canada.

[2]Centre for Research on Brain, Language and Music; Montreal, Canada.

[3]Department of Integrative Biology, University of Texas at Austin; Austin, USA.

[4]Smithsonian Tropical Research Institute; Gamboa, Panama.

[5]Child Study Center, Yale University; New Haven, USA.

[6]School of Psychology, University of Auckland; Auckland, New Zealand.

*Corresponding author. Email: loganjames@utexas.edu

Abstract: Many animals produce sounds during courtship, and receivers prefer some sounds over others. Shared ancestry and convergent evolution may generate similarities in preference across species and could underlie Darwin's conjecture that some animals "have nearly the same taste for the beautiful as we have". Here, we experimentally demonstrate that humans share acoustic preferences with a wide range of animals, that the strength of human preferences correlates with that in other animals, and that judgments are faster when in agreement with animals. Furthermore, we found evidence that music listening experience predicted agreement with animals. These results are consistent with theories arguing that biases in sensory and cognitive processing sculpt acoustic preferences and confirm Darwin's century-old hunch about the conservation of aesthetics and beauty in nature.

Main Text:

Diverse animals produce signals to attract mates. These include visual color patterns and movements, olfactory plumes, acoustic songs and calls, among many more, and span a wide range of animal clades including arthropods, mollusks, and all major vertebrate groups – fish, amphibians, reptiles, birds, and mammals (*1*). Receivers generally exhibit preferences among these signals, and honest signaling theories hypothesize that these preferences provide a benefit to the receiver (*2–5*). However, detailed investigations of sexual traits often fail to find strong evidence of honest signaling (*6*) and document relatively small effects (*5*), suggesting that preferences may exist for another function or could even arise incidentally (*7–11*).

Evidence for preferences arising outside of a benefit for receivers exists across many systems, including visual signals in fish and acoustic adornments in frogs, where male signals are thought to match (or 'exploit') pre-existing sensory biases (*12–19*). Indeed, preferences resulting from unrelated aspects of receiver psychology could lead to signal preferences that are shared across taxa, resulting in hidden preferences in species where signalers have not (yet) evolved signals that exploit those preferences (*20–22*). Revealing hidden preferences is difficult; consequently, little is known about the evolutionary breadth of shared preferences. However, it is evident that the dazzling colors of butterflies, aromas of flowers, and songs of birds are attractive not just to their intended receivers, but to humans as well. Indeed, Darwin noted that some animals "have nearly the same taste for the beautiful as we have" (*23*, *24*).

Here, we capitalize on decades of research on acoustic preferences across a taxonomically broad range of species to conduct a global citizen-science experiment to test the degree to which human preferences align with those observed in non-human animals.

Results

We analyzed the auditory preferences of a global sample of humans (n = 3,804 participants; Fig S1) across 110 pairs of acoustic stimuli from non-human animals (hereafter, animals) with demonstrated conspecific preferences (covering 16 species (*25–50*); Table S1; n = 43,759 responses). Humans were more likely than chance to select the stimulus preferred by animals across all pairs of stimuli (GLMM: $z = 2.0$, $p = 0.0484$; Fig. 1A).

The chance of human participants selecting the animal-preferred stimulus positively correlated with the strength of preference exhibited by animals (GLMM: $\chi^2_1 = 5.8$, $p = 0.0160$; Fig. 1B), and humans responded faster when they exhibited preferences that were consistent with animal preferences (LMM: $\chi^2_1 = 10.0$, $p = 0.0016$; Fig. 1C). Further, humans reliably selected the same stimulus when asked to rate the same pair of stimuli twice (LMM: $t_9 = 6.5$, $p = 0.0002$), even for stimuli where there was not significant agreement across participants ($t_{41} = 4.0$, $p = 0.0002$; Fig. 1D); this indicates that participants rarely selected at random.

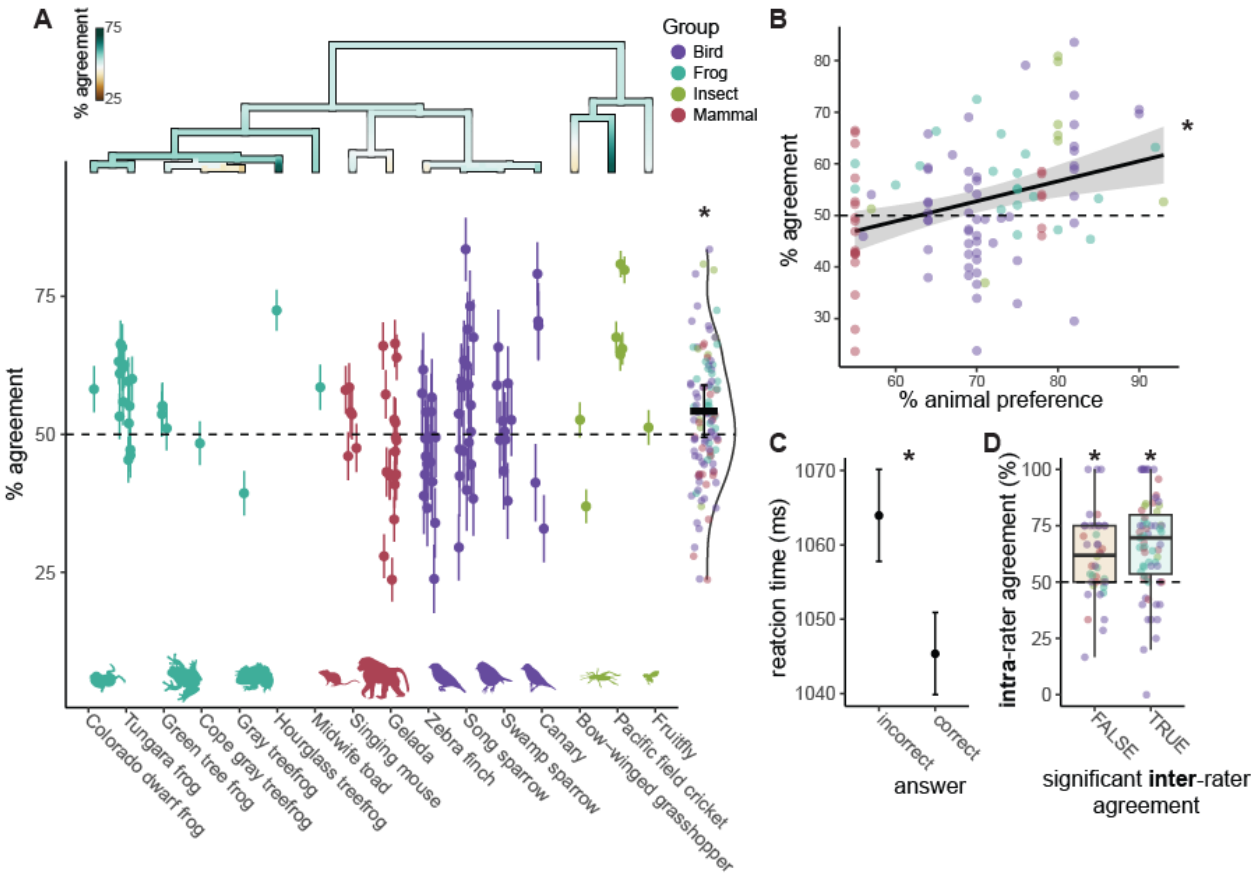

Fig. 1. Humans and other animals share acoustic preferences *(**A**) Each dot depicts the mean and 95% confidence interval for human responses to a pair of stimuli. Above is a phylogenetic tree, with color depicting the average agreement between humans and the species across all stimulus pairs (teal colors depict agreement, brown colors disagreement). Right are all data pooled with the horizontal bar and whiskers depicting the mean ± 95% confidence interval across all stimuli. (**B**) There is a significant relationship between the strength of the preference within the animal studies (x-axis) and the percent of humans that agreed with animals (y-axis). Each dot depicts a pair of stimuli, and the trendline is a simple linear correlation ± SE. (**C**) Reaction times were significantly faster for trials when the participant selected the stimulus that was preferred by animals (dots and error bars depict the mean ± SE across all trials). Note that, for analysis, we used z-transformed square root values, which were back transformed for visualization. (**D**) Intra-rater agreement is above chance both for stimulus pairs with significant inter-rater agreement (right) as well as for pairs at chance inter-rater agreement (left). Each dot depicts the mean intra-rater agreement for a stimulus pair. For all panels, colors depict the four larger phylogenetic groups (insects, birds, mammals, and frogs). * $p < 0.05$.*

Given the breadth of species, there was substantial diversity in the acoustic structure of the stimulus pairs allowing us to ask whether variation across sounds could predict aspects of preference. We categorized the traits that varied between the individual stimuli within each pair (Table S1), and found that the distinguishing trait-category significantly affected whether participants agreed with animals (GLMM: $\chi^2_9 = 22.2$, $p = 0.0083$; Fig. 2A). Therefore, we independently assessed the degree to which each trait-category predicted human-animal agreement. Among stimuli that differed in a single acoustic parameter, only frequency

significantly predicted the degree of agreement (GLMM: $z = 5.0$, $p < 0.0001$). Among traits defined by characteristics spanning many features (see Methods), we found significant agreement for sounds with acoustic adornments (e.g., "trills", "clicks", and "chucks"; $z = 4.2$, $p < 0.0001$) and a non-significant trend for agreement for individual differences (e.g., "high" vs "low" quality; $z = 1.8$, $p = 0.0705$). We did not detect significant effects for other traits ($p > 0.10$). While some acoustic traits predicted agreement for specific stimuli, we also analyzed measurable acoustic features across all stimulus pairs in the dataset, and found little evidence that differences in single acoustic features between stimuli in the pair could predict animal and human agreement when analyzed across all stimuli (Fig. 2B; Supplementary text; Table S2; Fig. S2).

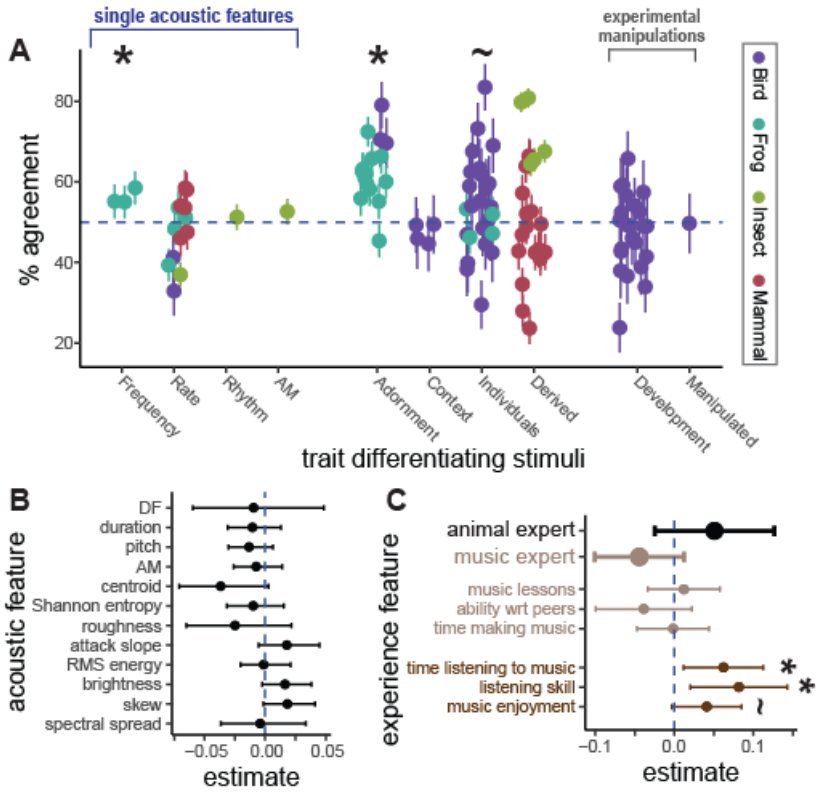

Fig. 2: Acoustic features and experience have little impact on agreement with non-human animals. *(**A**) Results for stimuli separated by the trait that differentiated the stimuli in the pair. (**B,C**) Model estimates and 95% confidence intervals for the impact of acoustic features (B) and participant demographics (C) on agreement with animals. For (B), all acoustic features were z-transformed to facilitate comparison across features. For (C), all values were binarized to facilitate comparison across variables, but asterisks depict significance from models with the full data. The larger dots in (C) indicate our main analyses, and smaller dots indicate exploratory analyses. For all panels, * $p < 0.05$, ~ $p < 0.10$.*

Experience can impact perception and preferences, and we had two *a priori* predictions of how experience could modulate human alignment with animal preferences: experience with animal sounds and musical expertise. Neither humans with experience identifying animals by their sounds ("animal experts"; n = 334, ~9% of participants that responded to the question), nor

expert musicians (n = 651, ~18% of participants that responded to the question) significantly agreed with animals more than non-experts (GLMMs: animal experts: $\chi^2_1 = 1.7$, $p = 0.1878$; music experts: $\chi^2_1 = 2.4$, $p = 0.1235$; Fig. 2C).

Similarly, we did not find evidence that additional measurements of music production expertise shaped the preferences of participants. In particular, whether participants had taken music lessons, their musical ability relative to their peers (if they did take lessons), and their daily time making music did not affect the likelihood of agreement with animal preferences (Table S3). However, music listening did predict human preferences. Participants that indicated more time listening to music, greater listening skill, or higher ratings of musical enjoyment agreed with animals more than other participants (time listening: $\chi^2_1 = 11.1$, $p = 0.0008$; music skill: $\chi^2_1 = 6.7$, $p = 0.0094$; music enjoyment: $\chi^2_1 = 3.3$, $p = 0.0675$; Fig. 2C, S1; Table S3).

Discussion

Our global survey discovered that humans share acoustic preferences with other animals, spanning insects, frogs, birds, and non-human mammals. This was demonstrated both in the choices that participants made and their timing of responses. In addition, the overall strength of preference across participants correlated with the strength of preference of animals.

Whereas we did not find an effect of experience with animal sounds (*51*, *52*) or music production expertise on alignment with animal preference, we did find that higher levels of music listening experience and expertise predicted whether humans agreed with animals. For instance, participants that rated their music listening skills higher (e.g., rated themselves skilled at "things like remembering melodies, hearing out of tune notes, or hearing a beat that is out of sync with the music") significantly agreed with animals more than those with lower listening skill ratings, raising the possibility that such participants had increased attention or were better able to discriminate between the sounds in a pair.

The notion that multiple species may respond similarly to the same stimuli is consistent with general form-function relationships in acoustic signals across species (*53*). For example, sound "roughness" is observed in alarm and distress calls and has been found to be aversive across numerous species, including humans and other mammals, frogs, and birds (*54–61*). Humans appear to be sensitive to relationships between form and function in other species (*62–64*), and we produce such relationships in our own vocalizations (*65–70*). Moreover, some animals are sensitive to aspects of distress calls from other species and many animals respond to heterospecific alarm calls (*71*, *72*). However, the majority of studies across species have primarily focused on sound used in aggressive or submissive contexts, rather than sounds used for attraction and preference.

There is less evidence about features that robustly predict preferred sounds (*60*, *61*, *73*), which is consistent with our finding that no single acoustic feature was effective for predicting agreement between humans and other animals. Relatively ambiguous measures of sound (e.g., "complexity") are often found to explain preferences within a specific species or context (*74–78*), including human preferences for songs from different species of birds (*79*), but these measures are difficult to compute and compare across very different types of sounds. Relatedly, measures that have been used to describe the timbre of music like "brightness" (*80*) cannot be directly applied to a range of animal sounds such as the clicks of insects.

Nevertheless, the present results raise the possibility that form-function relationships exist in acoustic signals used for attraction, and that humans are sensitive to these relationships in species other than our own. Such relationships may also support the perception of attractiveness in our own communicative vocalizations, including speech and music (*61*, *73*, *81*, *82*). The specific details of these aesthetic relationships are not yet clear, however, suggesting an opportunity for future research to uncover the underlying mechanisms. Of particular interest are the acoustic structures that induce preference and the neural and behavioral responses to such sounds across species.

Altogether, these results confirm the notion that a single species may show similar preferences as a taxonomically broad range of species across a diverse set of acoustic stimuli. Furthermore, while Darwin limited his notion of shared preferences to the visual coloration of birds, our data indicate a more expansive shared 'taste for the beautiful' (*23*, *24*) and remind us that much of the beauty we find in nature was originally intended for receivers other than us.

References and Notes

1. J. W. Bradbury, S. L. Vehrencamp, *Principles of Animal Communication* (Sinauer Associates, Sunderland, MA, 2011).

2. W. D. Hamilton, M. Zuk, Heritable true fitness and bright birds: A role for parasites? *Science* **218**, 384–387 (1982).

3. R. A. Johnstone, Sexual selection, honest advertisement and the handicap principle: Reviewing the evidence. *Biol Rev* **70**, 1–65 (1995).

4. S. L. Balenger, M. Zuk, Testing the Hamilton-Zuk Hypothesis: Past, present, and future. *Integr Comp Biol* **54**, 601–613 (2014).

5. L. R. Dougherty, Meta-analysis reveals that animal sexual signalling behaviour is honest and resource based. *Nat Ecol Evol* **5**, 688–699 (2021).

6. P. R. Y. Backwell, J. H. Christy, S. R. Telford, M. D. Jennions, J. Passmore, Dishonest signalling in a fiddler crab. *Proc R Soc Lond B* **267**, 719–724 (2000).

7. M. J. West-Eberhard, Sexual selection, social competition, and evolution. *Proc Am Philos Soc* **123**, 222–234 (1979).

8. J. A. Endler, Signals, signal conditions, and the direction of evolution. *Am Nat* **139**, S125–S153 (1992).

9. T. Guilford, M. S. Dawkins, Receiver psychology and the design of animal signals. *TINS* **16**, 430–436 (1993).

10. M. J. Ryan, Sexual selection, receiver biases, and the evolution of sex differences. *Science* **281**, 1999–2003 (1998).

11. M. J. Ryan, M. E. Cummings, Perceptual biases and mate choice. *Annu Rev Ecol Evol Syst* **44**, 437–459 (2013).

12. M. J. Ryan, A. S. Rand, The sensory basis of sexual selection for complex calls in the tungara frog, *Physalaemus pustulosus* (Sexual selection for sensory exploitation). *Evolution* **44**, 305 (1990).

13. M. J. Ryan, J. H. Fox, W. Wilczynski, A. S. Rand, Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. *Nature* **343**, 66–67 (1990).

14. F. H. Rodd, K. A. Hughes, G. F. Grether, C. T. Baril, A possible non-sexual origin of mate preference: are male guppies mimicking fruit? *Proc R Soc Lond B* **269**, 475–481 (2002).

15. C. Smith, I. Barber, R. J. Wootton, L. Chittka, A receiver bias in the origin of three–spined stickleback mate choice. *Proc R Soc Lond B* **271**, 949–955 (2004).

16. C. M. Garcia, E. Ramirez, Evidence that sensory traps can evolve into honest signals. *Nature* **434**, 501–505 (2005).

17. J. P. Renoult, T. C. Mendelson, Processing bias: extending sensory drive to include efficacy and efficiency in information processing. *Proc R Soc Lond B* **286**, 20190165 (2019).

18. S. V. Hulse, J. P. Renoult, T. C. Mendelson, Sexual signaling pattern correlates with habitat pattern in visually ornamented fishes. *Nat Commun* **11** (2020).

19. B. M. B. Downer-Bartholomew, F. H. Rodd, Female preference for color-enhanced males: a test of the sensory bias model in medaka, a drab fish. *Behav Ecol* **33**, 252–262 (2022).

20. M. J. Ryan, Sexual selection, sensory systems and sensory exploitation. *Oxford Surveys in Evolutionary Biology* **7**, 157–195 (1990).

21. J. Endler, A. Basolo, Sensory ecology, receiver biases and sexual selection. *TREE* **13**, 415–420 (1998).

22. G. G. Rosenthal, M. J. Ryan, Sexual selection and the ascent of women: Mate choice research since Darwin. *Science* **375**, eabi6308 (2022).

23. M. J. Ryan, *A Taste for the Beautiful: The Evolution of Attraction* (Princeton University Press, Princeton, NJ, 2018).

24. C. Darwin, *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).

25. E. Vallet, M. Kreutzer, Female canaries are sexually responsive to special song phrases. *Anim Behav* **49**, 1603–1610 (1995).

26. E. Vallet, I. Beme, M. Kreutzer, Two-note syllables in canary songs elicit high levels of sexual display. *Anim Behav* **55**, 291–297 (1998).

27. S. Leitner, C. K. Catchpole, Female canaries that respond and discriminate more between male songs of different quality have a larger song control nucleus (HVC) in the brain. *J Neurobiol* **52**, 294–301 (2002).

28. S. Leitner, C. Voigt, R. Metzdorf, C. K. Catchpole, Immediate early gene (ZENK, Arc) expression in the auditory forebrain of female canaries varies in response to male song quality. *J Neurobiol* **64**, 275–284 (2005).

29. S. Nowicki, W. A. Searcy, S. Peters, Quality of song learning affects female response to male bird song. *Proc R Soc Lond B* **269**, 1949–1954 (2002).

30. W. A. Searcy, S. Peters, S. Kipper, S. Nowicki, Female response to song reflects male developmental history in swamp sparrows. *Behav Ecol Sociobiol* **64**, 1343–1349 (2010).

31. Y. Chen, O. Clark, S. C. Woolley, Courtship song preferences in female zebra finches are shaped by developmental auditory experience. *Proc R Soc Lond B* **284**, 20170054 (2017).

32. C. Lauay, N. M. Gerlach, E. Adkins-Regan, T. J. DeVoogd, Female zebra finches require early song exposure to prefer high-quality song as adults. *Anim Behav* **68**, 1249–1255 (2004).

33. H. J. Barr, E. M. Wall, S. C. Woolley, Dopamine in the songbird auditory cortex shapes auditory preference. *Curr Biol*, 1–13 (2021).

34. M. L. Tomaszycki, E. Adkins-Regan, Experimental alteration of male song quality and output affects female mate choice and pair bond formation in zebra finches. *Anim Behav* **70**, 785–794 (2005).

35. M. J. Ryan, A. S. Rand, Sexual selection and signal evolution: the ghost of biases past. *Phil Trans R Soc B* **340**, 187–195 (1993).

36. M. S. Reichert, G. Höbel, Modality interactions alter the shape of acoustic mate preference functions in gray treefrogs. *Evolution* **69**, 2384–2398 (2015).

37. K. L. Laird, P. Clements, K. L. Hunter, R. C. Taylor, Multimodal signaling improves mating success in the green tree frog (*Hyla cinerea*), but may not help small males. *Behav Ecol Sociobiol* **70**, 1517–1525 (2016).

38. K. D. Wells, J. J. Schwartz, Vocal communication in a neotropical treefrog, *Hyla ebraccata*: Advertisement calls. *Anim Behav* **32**, 405–420 (1984).

39. R. Márquez, Female choice in the midwife toads (*Alytes obstetricans* and *A. cisternasii*). *Behav* **132**, 151–161 (1995).

40. R. Márquez, J. Bosch, X. Eekhout, Intensity of female preference quantified through playback setpoints: call frequency versus call rate in midwife toads. *Anim Behav* **75**, 159–166 (2008).

41. K. L. Akre, H. E. Farris, A. M. Lea, R. A. Page, M. J. Ryan, Signal perception in frogs and bats and the evolution of mating signals. *Science* **333**, 751–752 (2011).

42. M. J. Ryan, K. L. Akre, A. T. Baugh, X. E. Bernal, A. M. Lea, C. Leslie, M. B. Still, D. C. Wylie, A. S. Rand, Nineteen years of consistently positive and strong female mate preferences despite individual variation. *Am Nat* **194**, 125–134 (2019).

43. M. J. Ryan, *The Túngara Frog - a Study in Sexual Selection and Communication* (University of Chicago Press, 1985).

44. M. J. Ryan, A. S. Rand, Sexual selection in female perceptual space: How female túngara frogs perceive and respond to complex population variation in acoustic mating signals. *Evolution* **57**, 2608–2618 (2003).

45. M. S. Reichert, B. Ronacher, Noise affects the shape of female preference functions for acoustic signals: Female preferences in noise. *Evolution* **69**, 381–394 (2015).

46. M. G. Ritchie, E. J. Halsey, J. M. Gleason, *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in *D. melanogaster* song. *Anim Behav* **58**, 649–657 (1999).

47. R. M. Tinghitella, E. D. Broder, G. A. Gurule-Small, C. J. Hallagan, J. D. Wilson, Purring crickets: The evolution of a novel sexual signal. *Am Nat* **192**, 773–782 (2018).

48. R. M. Tinghitella, E. D. Broder, J. H. Gallagher, A. W. Wikle, D. M. Zonana, Responses of intended and unintended receivers to a novel sexual signal suggest clandestine communication. *Nat Commun* **12**, 797 (2021).

49. M. L. Gustison, T. J. Bergman, Vocal complexity influences female responses to gelada male calls. *Sci Rep* **6**, 19680 (2016).

50. B. Pasch, A. S. George, P. Campbell, S. M. Phelps, Androgen-dependent male vocal performance influences female preference in Neotropical singing mice. *Anim Behav* **82**, 177–183 (2011).

51. R. Rosenstein, M. L. Goldblatt, Birdsong complexity and preference: The role of birdwatching expertise. *Soc Animals*, 1–19 (2024).

52. J.-P. Chartrand, S. Filion-Bilodeau, P. Belin, Brain response to birdsongs in bird experts: *NeuroReport* **18**, 335–340 (2007).

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

54. T. Götz, V. M. Janik, Aversiveness of sounds in phocid seals: psycho-physiological factors, learning processes and motivation. *J Exp Biol* **213**, 1536–1548 (2010).

55. J. Soltis, T. E. Blowers,, A. Savage, Measuring positive and negative affect in the voiced sounds of African elephants (*Loxodonta africana*). *J Acoust Soc Am* **129**, 1059–1066 (2011).

56. J. C. Hechavarría, M. Jerome Beetz, F. García-Rosales, M. Kössl, Bats distress vocalizations carry fast amplitude modulations that could represent an acoustic correlate of roughness. *Sci Rep* **10**, 7332 (2020).

57. A. Anikin, K. Pisanski, M. Massenet, D. Reby, Harsh is large: nonlinear vocal phenomena lower voice pitch and exaggerate body size. *Proc R Soc Lond B* **288**, 20210872 (2021).

58. A. Marx, R. Lenkei, P. Pérez Fraga, V. Bakos, E. Kubinyi, T. Faragó, Occurrences of non-linear phenomena and vocal harshness in dog whines as indicators of stress and ageing. *Sci Rep* **11**, 4468 (2021).

59. N. Di Stefano, C. Spence, Roughness perception: A multisensory/crossmodal perspective. *Atten Percept Psychophys* **84**, 2087–2114 (2022).

60. S. Kumar, H. M. Forster, P. Bailey, T. D. Griffiths, Mapping unpleasantness of sounds to their auditory representation. *J Acoust Soc Am* **124**, 3810–3817 (2008).

61. J. H. McDermott, "Auditory preferences and aesthetics" in *Neuroscience of Preference and Choice* (Academic Press, 2012), pp. 227–256.

62. P. Filippi, J. V. Congdon, J. Hoang, D. L. Bowling, S. A. Reber, A. Pašukonis, M. Hoeschele, S. Ocklenburg, B. de Boer, C. B. Sturdy, A. Newen, O. Güntürkün, Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: evidence for acoustic universals. *Proc R Soc Lond B* **284**, 20170990 (2017).

63. T. Kelly, D. Reby, F. Levréro, S. Keenan, E. Gustafsson, A. Koutseff, N. Mathevon, Adult human perception of distress in the cries of bonobo, chimpanzee, and human infants. *Biol J Linn Soc* **120**, 919–930 (2017).

64. A. Terrade, M. Massenet, L. Pernel, N. Mathevon, A. Anikin, D. Reby, Nonlinear phenomena make animal calls alarming for human listeners. *iScience* **28**, 112600 (2025).

65. L. H. Arnal, A. Flinker, A. Kleinschmidt, A.-L. Giraud, D. Poeppel, Human screams occupy a privileged niche in the communication soundscape. *Curr Biol* **25**, 2051–2056 (2015).

66. D. E. Blasi, S. Wichmann, H. Hammarström, P. F. Stadler, M. H. Christiansen, Sound–meaning association biases evidenced across thousands of languages. *Proc Natl Acad Sci USA* **113**, 10818–10823 (2016).

67. S. A. Mehr, M. Singh, H. York, L. Glowacki, M. M. Krasnow, Form and function in human song. *Curr Biol* **28**, 356-368.e5 (2018).

68. S. A. Mehr, M. Singh, D. Knox, D. M. Ketter, D. Pickens-Jones, S. Atwood, C. Lucas, N. Jacoby, A. A. Egner, E. J. Hopkins, R. M. Howard, J. K. Hartshorne, M. V. Jennings, J. Simson, C. M. Bainbridge, S. Pinker, T. J. O'Donnell, M. M. Krasnow, L. Glowacki, Universality and diversity in human song. *Science* **366**, eaax0868 (2019).

69. C. B. Hilton, C. J. Moser, M. Bertolo, H. Lee-Rubin, D. Amir, C. M. Bainbridge, J. Simson, D. Knox, L. Glowacki, E. Alemu, A. Galbarczyk, G. Jasienska, C. T. Ross, M. B. Neff, A. Martin, L. K. Cirelli, S. E. Trehub, J. Song, M. Kim, A. Schachner, T. A. Vardy, Q. D. Atkinson, A. Salenius, J. Andelin, J. Antfolk, P. Madhivanan, A. Siddaiah, C. D. Placek, G. D. Salali, S. Keestra, M. Singh, S. A. Collins, J. Q. Patton, C. Scaff, J. Stieglitz, S. C. Cutipa, C. Moya, R. R. Sagar, M. Anyawire, A. Mabulla, B. M. Wood, M. M. Krasnow, S. A. Mehr, Acoustic regularities in infant-directed speech and song across cultures. *Nat Hum Behav* **6**, 1545–1556 (2022).

70. M. Singh, S. A. Mehr, Universality, domain-specificity and development of psychological responses to music. *Nat Rev Psychol* **2**, 333–346 (2023).

71. S. Lingle, T. Riede, Deer mothers are sensitive to infant distress vocalizations of diverse mammalian species. *Am Nat* **184**, 510–522 (2014).

72. R. D. Magrath, T. M. Haff, P. M. Fallow, A. N. Radford, Eavesdropping on heterospecific alarm calls: from mechanisms to consequences. *Biol Rev* **90**, 560–586 (2015).

73. K. Pisanski, D. R. Feinberg, "Vocal Attractiveness" in *The Oxford Handbook of Voice Perception*, S. Frühholz, P. Belin, Eds. (Oxford University Press, 2018; https://academic.oup.com/edited-volume/38687/chapter/335931723), pp. 606–626.

74. N. Choi, M. Adams, K. Fowler-Finn, E. Knowlton, M. Rosenthal, A. Rundus, R. D. Santer, D. Wilgers, E. A. Hebets, Increased signal complexity is associated with increased mating success. *Biol Lett* **18**, 20220052 (2022).

75. M. J. Ryan, Female mate choice in a neotropical frog. *Science* **209**, 523–525 (1980).

76. M. S. Reichert, J. Finck, B. Ronacher, Exploring the hidden landscape of female preferences for complex signals. *Evolution* **71**, 1009–1024 (2017).

77. D. J. Mountjoy, R. E. Lemon, Female choice for complex song in the European starling: a field experiment. *Behav Ecol Sociobiol* **38**, 65–71 (1996).

78. H. M. Lampe, G. Saetre, Female pied flycatchers prefer males with larger song repertoires. *Proc R Soc Lond B* **262**, 163–167 (1995).

79. O. C. Da Silva, A. M. Melo Santos, N. Schiel, A. Souto, Like music to our ears: The complexity of bird vocalizations as a key factor of attractiveness. *Anthrozoös* **35**, 91–104 (2022).

80. S. McAdams, "Musical timbre perception" in *The Psychology of Music*, D. Deutsch, Ed. (Academic Press, London, UK, 2013), pp. 36–68.

81. D. R. Feinberg, B. C. Jones, A. C. Little, D. M. Burt, D. I. Perrett, Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Anim Behav* **69**, 561–568 (2005).

82. S. A. Mehr, M. M. Krasnow, G. A. Bryant, E. H. Hagen, Origins of music in credible signaling. *Behav Brain Sci* **44**, e60 (2021).

83. B. Long, J. Simson, A. Buxó-Lugo, D. G. Watson, S. A. Mehr, How games can make behavioural science better. *Nature* **613**, 433–436 (2023).

84. D. Bates, M. Maechler, B. Bolker, S. Walker, Package "lme4." *R Package* (2007).

85. J. Fox, M. Friendly, S. Weisberg, Hypothesis tests for multivariate linear models using the car Package. *The R Journal* **5**, 39 (2013).

86. K. L. Akre, M. J. Ryan, Female túngara frogs elicit more complex mating signals from males. *Behav Ecol* **22**, 846–853 (2011).

Acknowledgments: We thank Elizabeth Adkins-Regan, Jaime Bosch, Katherine Buchanan, Bruce Byers, Morgan Gustison, Albertine Leitão, Stefan Leitner, Stephen Nowicki, Bret Pasch, Susan Peters, Michael Reichert, Michael Ritchie, Ryan Taylor, Robin Tinghitella, Michelle Tomaszycki, Eric-Marie Vallet, and the Macaulay Library for generously providing acoustic stimuli; Jan Simson, Isaac Kinley, and Sophia MacNeil for help with code and data management; and the members of the Music Lab for feedback on the design of the experiment.

Funding:

Smithsonian Institute Postdoctoral Fellowship (LSJ)

Fonds de recherche du Québec – Nature et technologies PR-299652 (SCW, JTS)

US National Institutes of Health DP5OD024566 (SAM)

Royal Society of New Zealand Te Apārangi RDF-UOA2103 (SAM)

Author contributions:

Conceptualization: LSJ, SCW, JTS, MJR, SAM

Methodology: LSJ, JTS, CBH, SAM

Data curation: LSJ, CBH, SAM

Investigation: LSJ

Visualization: LSJ

Software: LSJ, CBH

Funding acquisition: LSJ, SCW, JTS, SAM

Project administration: SAM

Supervision: SCW, JTS, MJR, SAM

Writing – original draft: LSJ, SCW, JTS, CBH, MJR, SAM

Writing – review & editing: LSJ, SCW, JTS, CBH, MJR, SAM

Competing interests: Authors declare that they have no competing interests

Data and materials availability: All data and code are available through GitHub: https://github.com/themusiclab/animal-sounds.

Materials and Methods

Participants

Participants were visitors to the citizen-science website https://themusiclab.org) and we collected data for over a year between 2023-09-06 and 2025-04-14 (Fig. S1).

We used the following inclusion criteria:

1. Reported age from 18-90
2. Reported wearing headphones and/or "in a very quiet place"
3. Reported no hearing impairment
4. Reported fluency in English (either as native or second language)

Participants gave informed consent under an ethics protocol approved by the Yale University Human Research Protection Program (protocol 2000033433), and the McGill University Research Ethics Board (initial approval October 24, 2022; protocol 22-09-042).

Stimuli

We obtained stimuli from various sources (Table S1). These files were requested from scientists (see Acknowledgements) based on previous studies of conspecific acoustic preferences within a species. We searched for studies that conducted preference assays. We did not seek out stimuli where the primary difference between stimuli was solely in amplitude or duration. All stimuli were produced by male animals with females as the intended receivers.

The peak amplitude of stimuli was normalized using Audacity. The singing mouse stimuli were near the upper frequency threshold for sound detection in humans; therefore, these stimuli were shifted down two octaves using Audacity.

Experimental design

Participants were presented with 16 pairs of stimuli (4 from each broad category of frogs, birds, mammals, and insects) selected randomly from all possible pairs and presented in random order. Because stimuli were selected at random, some participants received the same pair of stimuli twice, allowing us to measure intra-rater reliability.

At the end of the experiment, participants were told which of the four categories they most agreed with and were presented with information about the sounds of that animal category. This was to allow participants to share their result and encourage others to participate, a form of 'gamification' (*83*). However, we did not indicate whether participants got the 'correct' answer during the experiment nor did we identify which trials were 'correct' at the end, preventing participants from learning across the course of the experiment.

Acoustic analyses

We measured 12 acoustic features using MATLAB. Three were calculated based on the waveform (dominant frequency, duration, and the amplitude modulation variance). The remaining 9 were measured using the MIRtoolbox (Music Information Retrieval): pitch (pitch height, or "best" pitch), spectral centroid, brightness, roughness, Shannon entropy, attack slope, RMS energy (global energy), spectral skewness, and spectral spread.

For each pair of stimuli, we calculated the difference in each feature (preferred stimulus – not preferred stimulus). Distributions of features were screened, and the values for roughness were

square-root transformed prior to the difference calculation to improve normality. Additionally, we removed outlying values from these analyses. For example, measures that rely on frequency (dominant frequency, pitch, spectral centroid, attack slope) were excluded for stimuli with little spectral information (e.g., insect clicks), and all birdsong recordings had a maximum or near maximum value for brightness and were excluded.

Demographic analyses

While our experiment was designed to discover broad consistencies with non-human animal preferences, we collected demographic data on participants to evaluate factors that could shape auditory preferences for non-human animal signals. Not all participants answered all demographic items; as such, we report the proportion of responses for each item as opposed to raw *N*s.

We collected these data for several different reasons. In addition to general demographic information (e.g., age, gender), in some cases we wondered whether different categories of participants might have different degrees of performance on the task, such as those with "expert" experience with animal sounds or experience making music. We used answers of "*Yes*" to the question "*Do you have a hobby or job where you identify animals by their sounds? For example birdwatching, herping, zookeeping, etc.*" to identify experts on animal sounds as a binary categorical variable. To identify experts in music, we used the question: "*Think of your skill at making music (using a musical instrument or singing). How would you rate your own skill?*". Any participant that answered "*I'm an expert*" or "*I have a lot of skill*" was considered an expert in a binary categorical variable.

In addition to these focal data, we obtained some additional data for exploratory analyses of how experience impacts participant responses.

1. Gender (Male, Female or Other), as a categorical variable.
2. Age, in years (numeric variable)
3. Education (numeric variable with 8 possible responses from "*1: Some elementary/middle school*" to "*8: Completed graduate school*")
4. Income (numeric variable with 9 possible responses from "*1: Under $10,000*" to "*9: Over $150,000*")
5. Daily time spent making music (numeric variable with 9 possible responses from "*1: no time*" up to "*9: > 4 hours*")
6. Daily time spent listening to music (numeric variable with 9 possible responses from "*1: no time*" up to "*9: > 4 hours*")
7. Whether they had ever taken music lessons (binary variable)
8. Music skill as a child in relation to peers (numeric variable with 5 possible responses from "*1: They were a lot better than me*" to "*5: I was a lot better than them*")
9. Listening skill (*"How good do you think your music* ***listening*** *skills are? (things like remembering melodies, hearing out of tune notes, or hearing a beat that is out of sync with the music)"* slider scale from 0 to 1000)
10. Music enjoyment ("*In general, how much do you enjoy music?"* slider scale from 0 to 1000)

For our analyses of age, we had relatively few participants at higher ages leading to very large eigenvalue in our models (see below); therefore, we limited our analyses of age to participants less than 45 years of age.

The responses from both slider scales (Listening skill and Music enjoyment) were very non-normally distributed. The default answer for both sliders was 500 (out of 1000), therefore, for listening skill we binarized the data with those moving the slider to the right as "skilled" compared to those that moved the slider left. We omitted answers of exactly 500 (i.e., participants did not move the slider; 548 out of 3256 participants). For music enjoyment, about half the participants moved the slider all the way to the right, so we binarized the data between those participants ("enjoyers") compared to all other participants. Again, we omitted responses of exactly 500 (219 out of 3803 participants).

We used the data as described for the statistical models (see below). However, for visual comparison of the effect sizes across each variable, we binarized each of the numeric variables at the midpoint of the scale (e.g., those that answered 1-4 vs 5-8 in the education scale).

Strength of animal preference

From previous literature, we extracted data on the strength of the conspecific preference within animals. For studies with binary choices, we considered the percent of animals to pick the preferred stimulus as the strength of preference (which is most akin to our measure of human preference). Otherwise, we compared the magnitude of the behavioral result across the two stimulus types (preferred vs unpreferred) and calculated the percent of responses for the preferred stimulus. For example, in single phonotaxis tests in crickets, 78.3% approached the preferred songs while 19.1% of crickets approached the unpreferred songs (*47*). Thus, out of all responses, 80% were for the preferred stimulus, indicating the strength of preference (78.3/(78.3 + 19.1)). As another example, swamp sparrow females produced a mean of 6.9 copulation solicitation displays when hearing preferred songs, and 3.9 displays when hearing unpreferred songs (*30*). Thus, the strength of preference was 6.9/(6.9 + 3.9) = 64%. When these values were not explicitly reported in the published papers, we used PlotDigitizer (https://plotdigitizer.com/) to extract values from published figures.

Statistics

To test whether participants selected the animal-preferred stimulus more than expected by chance, we ran an intercept-only generalized linear mixed effects model (GLMM) with a binary response variable (choice: 1 = agree with animals, 0 = disagree) and a binomial link function (package: 'lme4' (*84*)). We included Stimulus Pair ID, Species ID, and Participant ID as random effects. To determine which specific stimulus pairs had significant preferences, we also ran similar intercept-only GLMMs except on each pair of stimuli with Participant ID as the sole random effect. These models all measure whether overall responses differ from chance (0.5).

To assess whether the strength of preference in humans was similar to that in other animals, we conducted a GLMM with the strength of non-human animal preference for each stimulus pair as the independent variable, the same binary response variable as above, and with Species ID, Stimulus Pair ID, and Participant ID as random effects.

To analyze how variation in reaction time covaried with participant choices, we conducted an LMM with the reaction time (square root and z-transformed to improve model assumptions, limited to responses < 5 sec) as the dependent variable, whether the participant's decision agreed with animals as the independent variable, and Stimulus Pair ID, Species ID, and Participant ID as random effects. To test for intra-rater reliability, we conducted intercept-only LMMs with the

average percent of concordant choices within participants for each stimulus pair as the dependent variable and Species ID as a random effect, either for all stimulus pairs, or limited to those for which humans did not display an overall significant preference for one stimulus over the other (i.e., non-significant inter-rater reliability).

To assess whether acoustic features could predict aspects of preference, we conducted three models per acoustic feature. Our primary measure asked whether the difference in a feature between the two stimuli could predict agreement between humans and animals. Therefore, we conducted linear mixed effects models with acoustic difference (animal's preferred stimulus – less preferred stimulus) as the independent variable, average percent agreement as the dependent variable, and Species ID as a random effect. We also asked whether the value of any acoustic feature predicted whether humans or non-human animals preferred a stimulus in a pair. For this, we conducted mixed-effects models with an acoustic feature as the dependent variable, whether the stimulus was preferred by animals or by humans (i.e., > 50% human choice) as an independent binary categorical factor, and Stimulus Pair ID and Species ID as random effects.

To assess the degree to which demographic information about participants influenced their responses, we used similar GLMMs as above, with a binary response variable (choice: 1 = agree with animals, 0 = disagree) and a binomial link function. Our independent variable was the demographic variable of interest (see above), and Stimulus Pair ID, Species ID, and Participant ID as random effects in all models.

We conducted Type II Wald chi-square tests (package: 'car'(*85*)) to assess the significance of the GLMMs and Type III analyses of variance with Satterthwaite's method for the LMMs.

Supplementary Text

Acoustic feature analyses

We ran a series of exploratory acoustic analyses to determine whether any specific acoustic feature could help explain preferences across species. To this end, we calculated 12 spectral features, some of which are known to generally be important in acoustic communication (e.g., dominant frequency, Shannon entropy, and amplitude variance), and others which are known to be important in human music (e.g., roughness and spectral brightness; see Methods). For each feature, we calculated the difference of the feature between each pair of stimuli. We then ran three models to answer: (1) Can the feature predict which stimulus was preferred across animals? (2) Can the feature predict which stimulus was preferred by humans? (3) Can the magnitude and direction of differences in the feature predict the agreement between animals and humans?

Overall, we found little evidence that any of these spectral features could predict preferences (Table S2). For two acoustic features, the difference between the stimuli in the pair significantly related to which stimulus was preferred by humans and/or animals: brightness and spectral centroid. Animals generally preferred the stimulus with greater brightness (LMM: $F_{1,34.0} = 17.1$, $p = 0.0002$) and greater spectral centroid ($F_{1,102} = 9.7$, $p = 0.0024$; Fig. S2). Humans also generally preferred the stimulus with greater brightness ($F_{1,64.8} = 4.5$, $p = 0.0373$), but did not demonstrate significant preferences for spectral centroid ($F_{1,102} = 2.0$, $p = 0.1605$). For neither feature was there a significant relationship between the difference in the feature and the agreement between humans and animals (Brightness: $F_{1,32.7} = 2.2$, $p = 0.1518$; Spectral centroid: $F_{1,101} = 3.3$, $p = 0.0725$). However, we note that we were unable to include brightness

measurements for birds (measurement was maxed out) and most insects (no spectral information). For the remaining 10 features, none of the three analyses were significant.

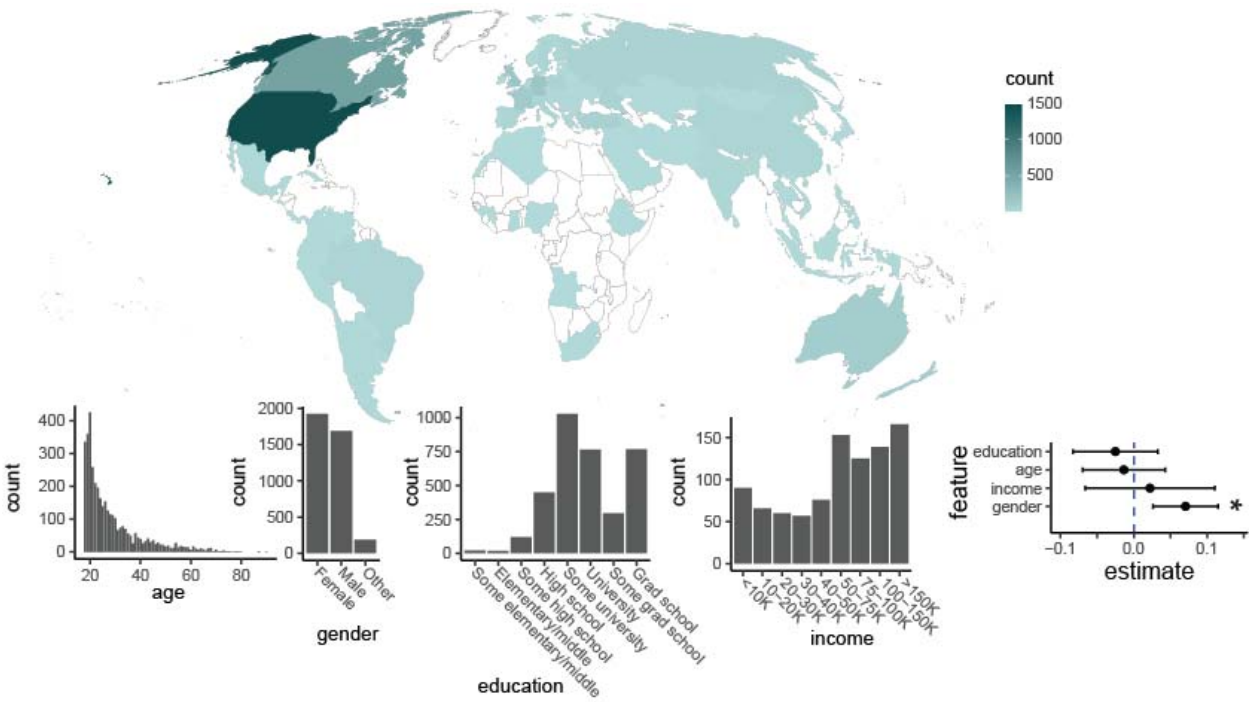

Fig. S1. Participant demographics.

Top: Global distribution of participants. Bottom: Summary bar plots depicting the age (in years), gender, education, and income levels of participants. Bottom right depicts the estimated effect of each demographic feature on participants' agreement with other animals. We found that only gender significantly predicted agreement with animals (male participants agreed with animals more than female participants: $\chi^2_2 = 9.9$, $p = 0.0071$; see Table S3). Note that for the purposes of visualization, each feature was binarized for the effect size estimate, but statistical analyses used all levels (numeric for age, factors for gender, and ordered factors for education and income).

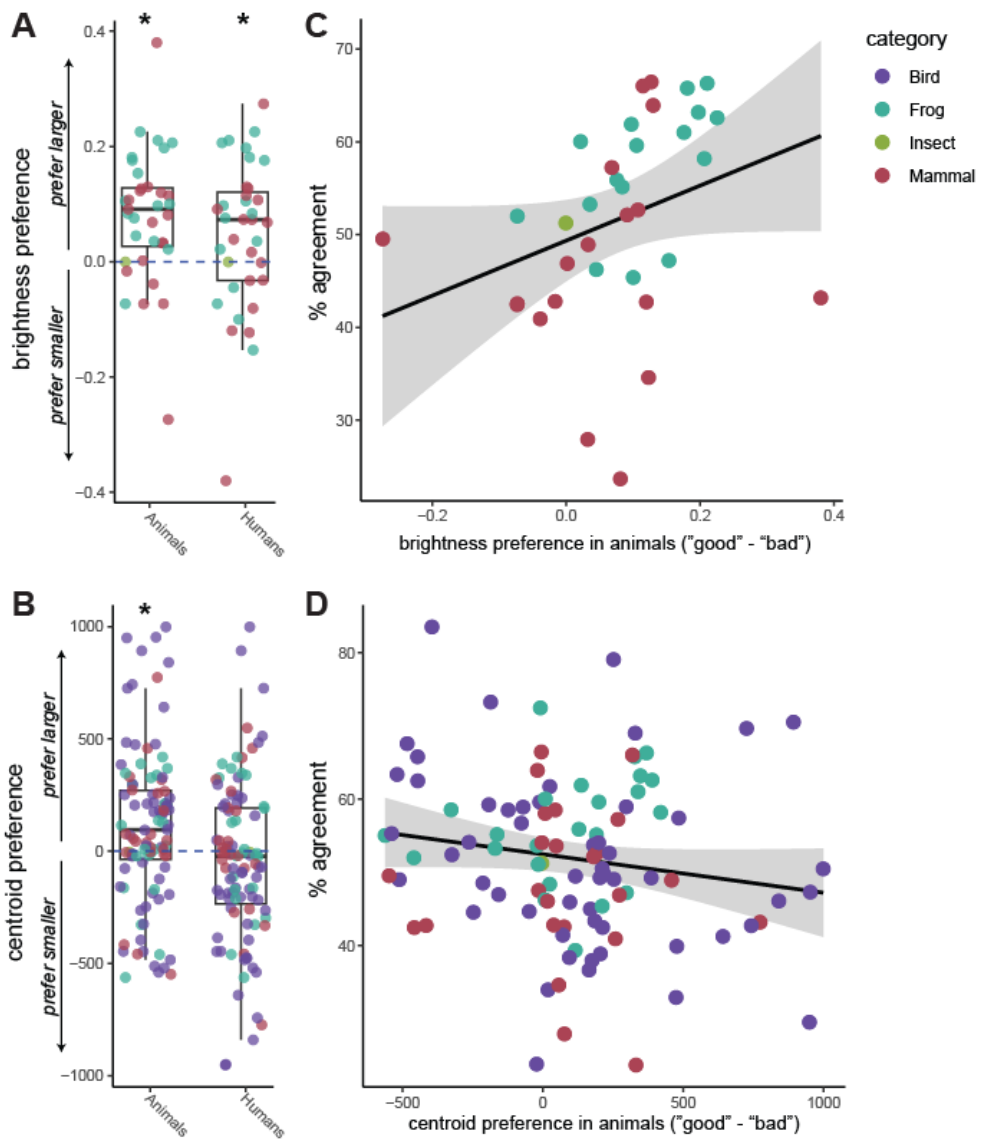

Fig. S2. Acoustic features.

Across stimulus pairs, animals tended to prefer the stimulus that had greater brightness (**A**) and greater spectral centroid (**B**), and humans also tended to prefer the stimulus with greater brightness (A). However, the magnitude and direction of difference of the features between the stimuli did not predict the percent agreement between humans and animals for either feature (**C, D**).

Table S1. Description of stimuli. The common name of the non-human animal species, the number of pairs of stimuli tested for that species, the main trait that the pairs differed on and direction of difference for the preferred stimulus, whether we were able to obtain and use the exact same stimuli that were tested on the animals themselves ("Same stim"), and references.

Species	N	Trait	Direction	Same stim	Citation
Canary	3	Adornment		Y	(*25–28*)
Canary	2	Rate	Faster	Y	(*25–28*)
Song sparrow	10	Individuals		Y	(*29*)
Song sparrow	10	Individuals		Y	(*29*)
Swamp sparrow	10	Development		Y	(*30*)
Zebra finch	4	Context		Y	(*31*)
Zebra finch	12	Development		N	(*32*); from Adkins-Regan
Zebra finch	2	Individuals		Y	(*33*)
Zebra finch	1	Manipulated		Y	(*34*)
Colorado dwarf frog	1	Adornment		Y	(*35*)
Cope gray tree frog	1	Rate	Faster	N	from Reichert
Gray tree frog	1	Rate	Average	Y	(*36*)
Green tree frog	2	Frequency	Lower	Y	(*37*)
Green tree frog	2	Rate	Faster	Y	(*37*)
Hourglass tree frog	1	Adornment		N	(*38*); from Macaulay Library
Midwife toad	1	Frequency	Lower	Y	(*39*, *40*)
Túngara frog	11	Adornment		Y	(*42*, *86*)
Túngara frog	4	Individuals		Y	(*44*)
Bow-winged grasshopper	1	Amplitude modulation	More	Y	(*45*)
Bow-winged grasshopper	1	Rate	Faster	Y	(*45*)
Fruit fly	1	Rhythm	Arrhythmic	Y	(*46*)
Pacific field cricket	5	Derived	Less	Y	(*47*, *48*)
Gelada	18	Derived	More	Y	(*49*)
Singing mouse	6	Rate	Faster	Y	(*50*)

Table S2. Acoustic results. This table summarizes the analyses into the effects of acoustic features on preferences. The "Diff" comparison asks whether the difference in a feature between the stimuli in the pair predicted agreement between humans and non-human animals. The "Animals" and "Humans" comparisons ask whether that group demonstrates a directional preference for the feature.

feature	comparison	DF	DF den	F	p
DF	Diff	1	101.1	0.1	0.7482
DF	Animals	1	106.0	2.4	0.1258
DF	Humans	1	106.0	0.1	0.8017
duration	Diff	1	95.7	0.8	0.3850
duration	Animals	1	109.0	3.1	0.0793
duration	Humans	1	109.0	0.7	0.4063
pitch	Diff	1	94.7	1.7	0.1901
pitch	Animals	1	194.5	0.1	0.7336
pitch	Humans	1	193.9	2.7	0.1046
AM var	Diff	1	108.0	0.5	0.5016
AM var	Animals	1	109.0	0.1	0.8183
AM var	Humans	1	109.0	1.1	0.2939
spectral centroid	Diff	1	101.0	3.3	0.0725
spectral centroid	Animals	1	102.0	9.7	**0.0024**
spectral centroid	Humans	1	102.0	2.0	0.1605
Shannon entropy	Diff	1	59.1	0.6	0.4485
Shannon entropy	Animals	1	202.9	1.6	0.2096
Shannon entropy	Humans	1	202.9	0.1	0.7791
roughness	Diff	1	99.7	1.1	0.2987
roughness	Animals	1	106.0	0.2	0.6407
roughness	Humans	1	106.0	3.6	0.0616
attack slope	Diff	1	101.0	1.7	0.1975
attack slope	Animals	1	102.0	0.4	0.5196
attack slope	Humans	1	102.0	0.5	0.5020
RMS	Diff	1	107.4	0.0	0.9133
RMS	Animals	1	203.2	0.4	0.5239
RMS	Humans	1	203.2	1.2	0.2677
brightness	Diff	1	32.7	2.2	0.1518
brightness	Animals	1	34.0	17.1	**0.0002**
brightness	Humans	1	64.8	4.5	**0.0373**
spectral skewness	Diff	1	88.3	2.5	0.1200
spectral skewness	Animals	1	202.9	2.7	0.1031
spectral skewness	Humans	1	202.9	1.9	0.1744
spectral spread	Diff	1	101.7	0.0	0.8323
spectral spread	Animals	1	104.0	0.2	0.6382
spectral spread	Humans	1	104.0	0.2	0.6336

Table S3. Demographic results. Summary of the effects of demographic variables on the similarity between human and non-human animal preferences.

Feature	Chi-squared	DF	p
animal sounds expert	1.74	1	0.1878
music expert	2.37	1	0.1235
music lessons	0.27	1	0.6049
ability wrt peers	0.14	1	0.7050
time making music	0.56	1	0.4534
music enjoyment	3.34	1	0.0675
time listening to music	11.14	1	**0.0008**
listening skill	6.75	1	**0.0093**
education	1.81	1	0.1784
gender	9.89	2	**0.0071**
income	0.09	1	0.7704
age	0.02	1	0.8804