

cambridge.org/bbs

## Target Articles

**Cite this article:** Savage PE, Loui P, Tarr B, Schachner A, Glowacki L, Mithen S, Fitch WT. (2021) Music as a coevolved system for social bonding. *Behavioral and Brain Sciences* 44, e59: 1–22. doi:10.1017/S0140525X20000333

Target Article Accepted: 17 August 2020

Target Article Manuscript Online: 20 August 2020  
Commentaries Accepted: 18 December 2020

### Keywords:

comparative; cooperation; cultural evolution; harmony; language; music; prediction; reward; synchrony; vocal learning

### What is Open Peer Commentary? What

follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (pp. 18 and 39) and an Authors' Response (pp. 132 and 140). See [bbsonline.org](https://bbsonline.org) for more information.

**Cite this article:** Mehr SA, Krasnow MM, Bryant GA, Hagen EH. (2021) Origins of music in credible signaling. *Behavioral and Brain Sciences* 44, e60: 23–39. doi:10.1017/S0140525X20000345

Target Article Accepted: 15 August 2020

Target Article Manuscript Online: 26 August 2020  
Commentaries Accepted: 18 December 2020

### Keywords:

coalitions; credible signaling; cultural evolution; infancy; music; natural selection; parent–offspring conflict; territoriality

We are delighted to present an unusual BBS publication. In early 2018, we received a double submission: two papers exploring the same topic from different perspectives – “Origins of Music in Credible Signaling,” by Samuel A. Mehr, Max M. Krasnow, Gregory A. Bryant, and Edward H. Hagen; and “Music as a Co-evolved System for Social Bonding,” by Patrick E. Savage, Psyche Loui, Bronwyn Tarr, Adena Schachner, Luke Glowacki, Steven Mithen, and W. Tecumseh Fitch. Each paper was reviewed in parallel, but independently, and both ultimately accepted. Our intention was to encourage consideration of how complex subjects like music might be investigated in different ways, integrating the perspectives of different laboratories and multiple commentators.

Thus, invited commentators might respond to the Mehr et al. article, the Savage et al. article, or both. Most chose both, as hoped. Unlike the usual BBS article presentation, the two target articles, two commentary groups and responses are interleaved. Follow the links above to find the companion target article and for the [index](#) of commentaries and responses. – The Editors

Patrick E. Savage<sup>a</sup>, Psyche Loui<sup>b</sup>, Bronwyn Tarr<sup>c</sup>, Adena Schachner<sup>d</sup>,  
Luke Glowacki<sup>e</sup>, Steven Mithen<sup>f</sup> and W. Tecumseh Fitch<sup>g</sup>

<sup>a</sup>Faculty of Environment and Information Studies, Keio University, Fujisawa 252-0882, Japan; <sup>b</sup>College of Arts, Media and Design, Northeastern University, Boston, MA 02115, USA; <sup>c</sup>Institute of Cognitive and Evolutionary Anthropology & Department of Experimental Psychology, University of Oxford, Oxford OX2 6PN, UK; <sup>d</sup>Department of Psychology, University of California San Diego, La Jolla, CA 92093, USA; <sup>e</sup>Department of Anthropology, Boston University, Boston, MA 02215, USA; <sup>f</sup>Department of Archaeology, University of Reading, Reading RG6 6AB, UK and <sup>g</sup>Department of Behavioral and Cognitive Biology, University of Vienna, Vienna 1090, Austria.

[psavage@sfc.keio.ac.jp](mailto:psavage@sfc.keio.ac.jp); <http://PatrickESavage.com>

[p.loui@northeastern.edu](mailto:p.loui@northeastern.edu); <http://www.psycheloui.com>

[bronwyn.tarr@anthro.ox.ac.uk](mailto:bronwyn.tarr@anthro.ox.ac.uk); [bronwyn.tarr01@gmail.com](mailto:bronwyn.tarr01@gmail.com); <https://www.anthro.ox.ac.uk/people/dr-bronwyn-tarr-schachner@ucsd.edu>; <https://madlab.ucsd.edu>

[glowacki@fas.harvard.edu](mailto:glowacki@fas.harvard.edu); <https://www.hsb-lab.org/>

[s.j.mithen@reading.ac.uk](mailto:s.j.mithen@reading.ac.uk); <http://www.reading.ac.uk/archaeology/about/staff/s-j-mithen.aspx>

[tecumseh.fitch@univie.ac.at](mailto:tecumseh.fitch@univie.ac.at); <https://homepage.univie.ac.at/tecumseh.fitch/>

## Abstract

Why do humans make music? Theories of the evolution of musicality have focused mainly on the value of music for specific adaptive contexts such as mate selection, parental care, coalition signaling, and group cohesion. Synthesizing and extending previous proposals, we argue that social bonding is an overarching function that unifies all of these theories, and that musicality enabled social bonding at larger scales than grooming and other bonding mechanisms available in ancestral primate societies. We combine cross-disciplinary evidence from archeology, anthropology, biology, musicology, psychology, and neuroscience into a unified framework that accounts for the biological and cultural evolution of music. We argue that the evolution of musicality involves gene–culture coevolution, through which proto-musical behaviors that initially arose and spread as cultural inventions had feedback effects on biological evolution because of their impact on social bonding. We emphasize the deep links between production, perception, prediction, and social reward arising from repetition, synchronization, and harmonization of rhythms and pitches, and summarize empirical evidence for these links at the levels of brain networks, physiological mechanisms, and behaviors across cultures and across species. Finally, we address potential criticisms and make testable predictions for future research, including neurobiological bases of musicality and relationships between human music, language, animal song, and other domains. The music and social bonding hypothesis provides the most comprehensive theory to date of the biological and cultural evolution of music.

## 1. Introduction

Darwin famously considered music to be a puzzle for evolutionary theory. Music is universal across human cultures (Brown & Jordania, 2013; Mehr et al., 2019; Savage, Brown, Sakai, & Currie, 2015), yet its function seems mysterious, because “neither the enjoyment nor the capacity of producing musical notes are faculties of the least use to man in reference to his daily habits of life...” (Darwin, 1871, p. 33). Darwin went on to speculate that music first evolved “for the sake of charming the opposite sex,” after which language “derived from previously developed musical powers.”

Since Darwin there has been no shortage of hypotheses about why and how music evolved (cf. Honing, Cate, Peretz, & Trehub, 2015; Wallin, Merker, & Brown, 2000). The null hypothesis is that music is an evolutionarily “useless” by-product of other evolved capacities, with no adaptive function and involving no direct selection for musical capacities (Pinker, 1997, p. 528). Others hypothesize that musicality evolved for specific adaptive purposes, including signaling mate quality (Miller, 2000), advertising coalitions (Hagen & Bryant, 2003; Merker, 2000), or soothing infants (Dissanayake, 2000; Falk, 2004; Mehr & Krasnow, 2017). Many authors have discussed the evolutionary value of music in facilitating group cohesion (e.g., Benzon, 2001; Brown, 2000a, 2007; Cross & Morley, 2009; Dissanayake, 2009; Dunbar, 2012a; Freeman, 2000; Gioia, 2019; Huron, 2001; Loersch & Arbuckle, 2013; McNeil, 1995; Merker, Morley, & Zuidema, 2018; Mithen, 2005; Oesch, 2019; Patel, 2018; Roederer, 1984; Schulkin & Raglan, 2014; Trainor, 2018; Trehub, Becker, & Morley, 2018), sometimes suggesting that musicality may have arisen via group selection (especially Brown, 2000a). Although such proposals succeed in explaining some properties (or genres) of music, we argue that

PATRICK SAVAGE is Associate Professor in the Faculty of Environment and Information Studies at Keio University in Japan, where he directs the CompMusic Lab for comparative and computational musicology. He holds a Ph.D. in Ethnomusicology from the Tokyo University of the Arts and an M.Sc. in Psychology, Neuroscience, and Behavior from McMaster University. Previously, he was a postdoc in the School of Anthropology and Museum Ethnography at the University of Oxford. His 2017 Ph.D. dissertation entitled “Measuring the cultural evolution of music” was awarded the Ikushi Prize from the Japanese Emperor, and he has won national competitions in Japanese folk song. Twitter: @PatrickESavage

PSYCHE LOUI is Assistant Professor of Creativity and Creative Practice in the Department of Music at Northeastern University, where she directs the Music, Imaging, and Neural Dynamics (MIND) Lab, which studies the cognitive neuroscience of music. She holds a Ph.D. in Psychology from the University of California at Berkeley and a B.S. in Psychology and Music from Duke University, and completed her postdoctoral training in neurology at the Harvard Medical School. Loui has published over 70 peer-reviewed journal articles and book chapters on music and the brain. Her work has received media coverage by the *New York Times*, *Boston Globe*, BBC, Associated Press, and other news outlets. Twitter: @psycheloui

BRONWYN TARR is Research Affiliate and Departmental Lecturer at the University of Oxford's Institute of Cognitive and Evolutionary Anthropology (ICEA). She is interested in the evolution of social behaviors and the psychological and physiological effects of dance. She holds a B.Sc. in Evolutionary Biology and Biodiversity, a B.Sc. in Zoology and Animal Behavior (both from the University of Cape Town), an M.Sc. in Environmental Change and Management, and a D.Phil. (Ph.D.) in Experimental Psychology (both from the University of Oxford). She was a Postdoctoral Research Associate at the Social and Evolutionary Neuroscience Research Group (University of Oxford), a visiting Research Fellow at the Institute for Advanced Study Toulouse, and she is currently based in the Social Body Lab (ICEA). Twitter: @BronwynTarr1

ADENA SCHACHNER is Assistant Professor of Psychology at the University of California, San Diego, where she directs the Mind and Development Laboratory. She holds a Ph.D. in Psychology from Harvard University and a B.A. in Cognitive Science from Yale University, and completed postdoctoral training at Boston University. Her research explores how children and adults reason about the social world, with a focus on the cognitive bases for links between music and social cognition. Dr. Schachner has been recognized with the Rising Star Award from the APS, and funded by the NSF, NIH, and APF. Twitter: @adenaschachner

LUKE GLOWACKI is a faculty member in the Department of Anthropology at Boston University, where he directs the Human Systems and Behavior Lab. He holds a Ph.D. in Human Evolutionary Biology from Harvard University and was previously a Research Fellow at the Institute for Advanced Study in Toulouse and faculty at Pennsylvania State University. Twitter: @HSB\_Lab

STEVEN MITHEN is Professor of Early Prehistory at the University of Reading, UK. Having studied at the universities of Sheffield, York, and Cambridge, his academic career at Reading since 1992 has involved managerial roles as Pro Vice Chancellor (International and Research) and Deputy Vice Chancellor. His teaching and research interests concern human evolution, hunter-gatherers, and the emergence of farming, with long-term field projects in southern Jordan and western Scotland. He was elected as a Fellow of the British Academy in 2003.

TECUMSEH FITCH is Professor of Cognitive Biology, and the head of the Department of Behavioral and Cognitive Biology, in the Faculty of Life Sciences at the University of Vienna in Austria. He is interested in the evolution of cognition and communication in both humans and other vertebrates, studied using a comparative approach, and particularly in the evolution of language and music. He holds a bachelor's in Biology and a Ph.D. in Cognitive Science, both from Brown University. After a postdoc at MIT and Harvard, he lectured at Harvard and the University of St. Andrews before moving to Vienna in 2009. He is a recipient of an ERC Advanced Grant and is an elected Fellow of the Association for Psychological Science. Twitter: @TecumsehFitch

no single account succeeds as a general explanatory framework for the evolution of human musicality. Our purpose in the current target article is to synthesize and extend previous proposals into a new, parsimonious framework that can explain and predict many aspects of human music-making.

Our argument is that *human musicality is a coevolved system for social bonding*. Crucially, following Honing (2018) and others, we clearly distinguish between music and musicality. “Music” encompasses the diverse cultural products generated by and for music making: songs, instruments, dance styles, and so on. In contrast, “musicality” encompasses the underlying biological capacities that allow us to perceive and produce music. Distinguishing these clearly is crucial because musical systems are diverse, culture-specific products of cultural development, whereas musicality comprises multiple biological mechanisms, shared across human cultures that enable musical production, perception, and enjoyment.<sup>1</sup> Musicality is not a monolithic trait evolved to solve one particular problem (coalition signaling, infant mood regulation, sexual attraction, and so on), but rather a set of capabilities that can be used in different ways to support multiple functions, all involving social affiliation, but no one of which is the “primary” or “original” function.

The key phrase “social bonding” refers to the formation, strengthening, and maintenance of affiliative connections (“bonds”) with certain conspecifics (i.e., the set of social processes that engender the bonded relationships that underpin prosocial behavior). As a group-living primate species, such bonds are psychologically and biologically central to human survival and reproduction (e.g., via enhanced predator protection, cooperative child-rearing, collaborative foraging, expansion, and defense of territories; Dunbar, 2012b; Dunbar & Shultz, 2010; Hrdy, 2009; Tomasello & Vaish, 2013). For the purpose of this paper, we use “social bonding” as an umbrella term to encompass both bonding processes (over short and longer time scales) and their effects. Consequently, we take “social bonding” to encompass a variety of social phenomena including social preferences, coalition formation, identity fusion, situational prosociality, and other phenomena that bring individuals together. The social functions of music share a general social utility: to forge and reinforce affiliative inter-individual relationships, for example, by synchronizing and harmonizing the moods, emotions, actions, or perspectives of two or more individuals. Crucially, we argue that music achieves this in a variety of situations where language is less effective, and on a scale greater than that achievable by the ancestral bonding mechanisms (ABMs) available to other primates (e.g., grooming). We argue that social bonding promotes, and is the consequence of, interactions not only during music

making, but also subsequently via long-lasting changes in affiliative dispositions of group members toward one another, and their associated longer-term prosociality. Because social interactions involve multiple levels of group structure, our conception of social bonding also includes darker phenomena such as out-group exclusion that bring certain individuals closer together by pushing away others (see sect. 6.4).

The final keyword here is “*coevolved*”: we argue that culturally evolving systems of music (Savage, 2019a) have developed in tandem with the human capacity for musicality through a process of gene–culture coevolution. We build on recent arguments by Patel (2018) and Podlipniak (2017), who suggest that music arose initially as a cultural “invention” that created the context for later selection enhancing human musicality. In much the same way that the use of fire by early hominins provided the preconditions for biological adaptations to cooked food (Wrangham, 2009), or the invention of dairy farming in some European and African cultures created selection for lactase persistence (Tishkoff et al., 2007), early instantiations of music provided the selective preconditions for later neurobiological changes underlying human musicality. Notably, both Patel and Podlipniak identified social bonding as a candidate function driving such gene–culture coevolution, with Patel (2018, p. 118) noting the possibility that “musical behavior first arose as a human invention and then had (unanticipated) beneficial effects on social cohesion.” We argue that because music had multiple adaptive effects on social bonding, this led to subsequent selection (both genetic and cultural) for the ability and motivation to make particular forms of music – music that has features that most effectively function to promote social bonding. This combination of cultural and biological selection led to the particular features and ubiquity of modern human music and musicality.

Our article closely examines this claim, and provides a framework for understanding the biological and cultural evolution of music, taking this argument as foundational. We provide a detailed cross-disciplinary review of the evidence for specific mechanisms by which music functions to enhance social bonding, and consider how some of the mechanisms underlying musicality may have coevolved with music. Similar to Patel, we take for granted the large and sophisticated literature on gene–culture coevolution in general, and will not review it here (cf. Boyd & Richerson, 1985; Cavalli-Sforza, & Feldman, 1981; Durham, 1991; Henrich, 2016; Jablonka & Lamb, 2005; Laland, Odling-Smee, & Feldman, 2000, 2010; Richerson et al., 2010; Tomlinson, 2018). However, we do not see the “invention” of music as a unitary event later followed by genetic adaptation, but rather as an iterated process where different proto-musical components of musicality arose over an extended period as behavioral innovations that, because of initial positive effects, generated new cognitive and social niches for subsequent biological adaptations, themselves yielding new innovations, and so on in a virtuous spiral. We thus posit essentially an iterated Baldwin effect (Baldwin, 1896; Bateson, 2004; Griffiths, 2003; Podlipniak, 2017), or more generally, prolonged cognitive “niche construction” (Laland et al., 2000). This mechanism is closely related to many contemporary models of language evolution involving a series of “protolanguages” (Arbib, 2005; Fitch, 2010, 2017). Although hypotheses about the specific ordering of events involved (e.g., Dunbar, 2012a; Mithen, 2005) are useful, it is not our purpose here to propose a specific sequence, but rather to advance a new conception of the entire process.

In their target article, Mehr, Krasnow, Bryant, and Hagen present a contrasting hypothesis for the origins of music. Their hypothesis synthesizes and extends their previous proposals (Hagen & Bryant, 2003; Mehr & Krasnow, 2017) into a generalized “credible signaling” hypothesis that incorporates signaling of both coalition strength and parental attention. They also present critiques of the social bonding hypothesis and other candidate hypotheses. The BBS editors decided that publishing these two target articles with contrasting hypotheses would stimulate productive commentary beyond that usually possible for only a single target article. Both target articles originated from the same symposium on “The Origins of Music in Human Society,”<sup>2</sup> but differ in multiple ways in addition to the focus on social bonding versus credible signaling. In particular, Mehr et al. take an approach grounded in evolutionary psychology, focused on demonstrating domain-specificity and evidence for adaptation. In contrast, our approach emphasizes cultural evolutionary theory, including in particular gene–culture coevolution and cognitive niche construction (cf. Laland & Brown, 2011). We take a pluralistic approach to adaptation and modularity, involve experts from diverse disciplines to synthesize evidence into a single framework, and propose testable predictions for future research. We expand on more detailed contrasts between the two articles in sect. 6.

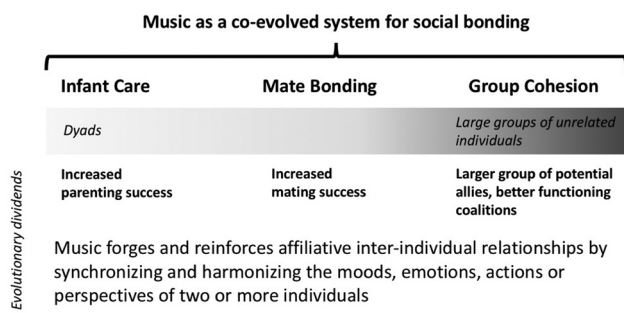
The following sections lay out the details and implications of the music and social bonding (MSB) hypothesis. Section 2 describes the proposed evolutionary functions and coevolutionary process. Section 3 details cross-disciplinary evidence supporting the MSB hypothesis. Section 4 specifies the neurobiological mechanisms proposed to underlie music’s social bonding functions. Section 5 describes testable predictions that follow from the MSB hypothesis. Section 6 addresses a number of potential criticisms of our hypothesis, and sect. 7 provides a brief conclusion.

## 2. Social bonding as a unifying function in the evolution of musicality

The music and social bonding (MSB) hypothesis posits that core biological components of human musicality evolved as mechanisms supporting social bonding. Musicality relies on multiple neurocognitive components, which likely evolved at different times and for different reasons: musicality is more a cognitive toolkit than a single tool (Fitch, 2015a). Most of the tools in this musical toolkit function to facilitate social bonding, but some may also be used for non-social purposes such as individual mood regulation (see sect. 6.5).

We avoid arguing for one specific single adaptive function for music (e.g., coalition advertisement, courtship, or infant mood regulation) because we think it unlikely that a single “main” evolutionary function for complex, multi-component abilities such as language or music exists. Imagine asking the parallel question “what is vision for?” and coming up with a hypothesis set including “spotting predators,” “judging mate quality,” “finding food,” and “avoiding obstacles.” It seems clear that these are all functions of vision, and all provide potential causal explanations for adaptive improvements in vision during evolution. But the desire to identify ONE function as primary seems misguided. A better approach is mechanistic: we ask “what are lenses for?,” and answer in engineering terms: lenses are for focusing an image on the retina, to enable accurate visual perception. Whether the





**Figure 1.** We propose that supposedly competing hypotheses for the evolution of human music, including mate bonding, infant care, and group cohesion (within both small coalitions and larger groups), are complementary sub-components of a broader social bonding function.

image is of a predator, mate, or food is not critical, because of improved visual resolution will aid them all.

Turning to music, “social bonding” provides an umbrella explanation analogous to “vision is for seeing.” Particular design features of music (singing discrete pitches, generating an isochronous beat, and use of repetitive patterns based on small-integer ratios) function mechanistically to enhance predictability, aiding synchronization and harmonization when multiple people sing, dance, and play instruments together. Coherent and harmonious merging of sounds and movements during group activity leads to positive feelings of prediction, fulfillment of expectation, and mutual accomplishment. These, through activation of the dopaminergic reward system and other pathways, have affiliative emotional and rewarding effects immediately and also long after music-making ceases (see sect. 4). Crucially, the resulting strengthened social bonds are operative over multiple types and sizes of groups, ranging from dyads (e.g., parent and infant, potential mates) to bands of small coalitions and large groups of unrelated individuals (Fig. 1). Social bonding through music thus produces its ultimate evolutionary dividends in multiple complementary ways, including a larger group of potential allies, increased child rearing success, increased mating success, and better-functioning coalitions.

## 2.1 Ancestral bonding mechanisms

Why was social bonding adaptive for our ancestors, and in what ways does music improve or increase social bonding? Group living comes with costs (e.g., increased local competition for food and mates) and benefits (e.g., safety in numbers and cooperative hunting/defense). Animals that live in groups, particularly primates, have evolved mechanisms that help balance these costs and benefits by forging strong affiliative bonds: good quality, persistent, differentiated inter-individual commitments that require investment of time and energy (Dunbar, 1991). Strong social bonds enhance individuals’ prospects of receiving support through coalitions, which, in certain primate species, influence dominance rank and reproductive performance (Silk, 2007). These coalitions form the backbone of successful cooperative hunting, child rearing, and joint defense against predators or competitors (Dunbar & Shultz, 2010). Ecological factors typically constrain the size of a group, but larger groups of well-coordinated, strongly bonded humans enabled exploitation of new forms of resources (e.g., larger prey), and more reliable protection from predators (Dunbar, 2012b).

ABMs in other primates include grooming, play, and – in some species – non-procreational sex. These ABMs are essentially dyadic (or for play, very small groups mostly limited to young animals), and require substantial time commitments even in small groups if all individuals in the group are to invest in all others. Although vocal duets are present in tropical birds and some primates (Farabaugh, 1982; Haimoff, 1986; Mann, Dingess, Barker, Graves, & Slater, 2009; Thorpe, 1972), group vocal choruses that are both differentiated and coordinated appear nearly unique to humans (but see Mann, Dingess, & Slater [2006] for the fascinating example of the group-chorusing plain-tailed wren).

As Dunbar (1993) has argued, the steady increases in group size, complexity, and fluidity that occurred during hominin evolution put increasing strain on ABM-based social bonds. Beyond group sizes of 20 or so, dyadic bonding based on ABMs such as grooming became unsustainably time-consuming, so supra-dyadic bonding mechanisms were needed. Dunbar (2012a) suggests that another ABM in great apes and humans was laughter (Davila Ross, Owren, & Zimmerman, 2009), which facilitates social bonds among reasonably large groups. However, there are limits to a bonding mechanism based on laughter: Unlike music, which people can intentionally choose to engage in at any time, large group laughter can be difficult to elicit and to sustain for long periods. Music may have provided our ancestors with a novel system that, like laughter, allowed for simultaneous bonding with a larger group of individuals, but across a broader set of times and contexts, and for longer periods of time than otherwise possible (Dunbar, 2012a; Launay, Tarr, & Dunbar, 2016). This new system augmented the smaller-scale ABMs that became less robust in larger groups. Specific design features of human musicality – particularly our capacity and proclivity to produce repetitive, synchronized, harmonized music for extended periods – provided a flexible toolkit for bonding, allowing our ancestors to achieve social bonding on a large scale.

## 2.2 Design features of musicality

### 2.2.1 Rhythm and dance

Most music has two distinctive rhythmic components: an isochronous (equal-timed) beat, and a metric structure (a hierarchical arrangement of sonic events into small groups with differentially accented constituents; Arom, 1991; London, 2004; Savage et al., 2015). These features together provide a predictable, repetitive structure underlying extended, coordinated, and varied group performances, while allowing room for variation and improvisation. Isochronicity and metric structure make the performance predictable, which facilitates planning synchronized and coordinated movements (e.g., dancing). Although synchronization solely to the beat (e.g., in marching or unison chanting) allows large groups to integrate, it tends to submerge individual contributions. Meter solves this problem by allowing many individuals to contribute, out of phase, to the same integrated rhythm. Neither of these core design features of musicality appears well-designed for solo performances, but they support the synchronized and coordinated musical sounds and dance movements of groups that are widespread features of human musical systems (Savage et al., 2015).

Dancing is another intrinsically rhythmic component of human musicality (cf. Fitch, 2015a, 2015b; Laland, Wilkins, & Clayton, 2016). Even newborn infants perceive a musical beat (Winkler, Háden, Ladinig, Sziller, & Honing, 2009), and dance develops early: Infants hearing music produce spontaneous rhythmic

movements during their first year, although the ability to entrain these movements reliably to a beat takes several years to develop (Kim & Schachner, 2020; McAuley, Jones, Holub, Johnston, & Miller, 2006; Merker, Madison, & Eckerdal, 2009; Zentner & Eerola, 2010). The capacity to perceive and move to a beat is a core component of musicality, rare among vertebrates (Patel, 2014; Schachner, Brady, Pepperberg, & Hauser, 2009) but universal across human cultures (Brown, 1991). Dance provides an energetic mode of musical participation that is accessible to large numbers of individuals regardless of age, familiarity with the music, or instrumental/singing virtuosity. In addition to its visual effects, dance can also generate an auditory signal, for example, because of foot stamping or hand clapping, and certain styles of dance (such as tap dancing) create their own sonic accompaniment. These factors suggest that dance is a core part of music-making (“musicking”) and not a separate domain (Tarr, 2017).

Dance thus expands the potential circle of rhythmically coordinated participants in musical interactions. The inclusive aspect of human musicality provided by dance is predicted by the MSB hypothesis, but poses a challenge to hypotheses seeing music primarily as a signal of virtuosity. Hereafter, we consider dance a core component of musical performance.

### 2.2.2 Melody, harmony, and vocal learning

The human capacity for song entails vocal production learning: the ability to imitate and learn vocal patterns beyond our species-typical repertoire of screams, laughter, and so on. By about 2 or 3 years of age (often earlier), children reproduce songs that their caregivers sing to them, with intact pitch range and contours (Trehub, 2016). Young children commonly exhibit greater fluency in song than in speech (e.g., singing Twinkle Twinkle Little Star from beginning to end with fractured, word-like sounds). This vocal learning ability is highly developed in humans relative to other primates, and the neurobiological mechanisms of its evolution are relatively well-understood, in part because of its convergent evolution in songbirds and other non-human species (Fitch, 2015a; Janik & Slater, 1999; Jarvis, 2019; Syal & Finlay, 2011; see sect. 4.4 for details). Vocal learning forms a foundation for group participation in singing culture-specific songs.

In contrast to the continuously varying pitch of normal speech, the discrete pitches used in song and instrumental music generate predictable sequences that enable frequency matching between individuals during group music production (Merker, 2002; Savage et al., 2015). Unison performance in which multiple parts produce the same melodies at either the same frequencies (1:1 frequency ratio) or an octave apart (2:1 ratio) is so widespread among humans it is often not even considered a form of harmonization (although cf. Jacoby et al. [2019] for evidence that octave equivalence is not completely universal). Octave singing in particular represents the most universal form of musical harmony: different pitches performed simultaneously with maximally overlapping acoustic spectra (cf. Bowling & Purves, 2015). The common tendency for men and women to sing together in octaves is paralleled by the roughly octave difference in men and women’s average vocal pitch, based on vocal anatomy (Titze, 1989). This is an unusual feature among primates (and mammals more generally) not observed in chimpanzees (Grawunder et al., 2018) – a potential anatomical adaptation for vocal harmonization.

Harmonious overlapping of acoustic spectra also shapes another common design feature: Musical systems around the world restrict pitches to scales containing a limited number of

discrete pitch classes (rarely more than seven; Savage et al., 2015). These pitch classes often reflect small-integer frequency relationships which sound consonant together (e.g., the 3:2 frequency ratio underlying musical fifths, 4:3 ratios for fourths, and so on; Bowling, Purves, & Gill, 2018; Gill & Purves, 2009; Kuroyanagi et al., 2019; McDermott, Lehr, & Oxenham, 2010; Terhardt, 1984). By producing pitches that adhere to scales, groups of singing individuals effectively minimize uncertainty in fundamental frequency, thus maximizing harmony via spectral alignment (Sethares, 2004). Coordinating with other individuals musically, by aligning acoustic spectra, can sound pleasing and promote bonding. The specific mechanisms and causal relationships behind this effect remain contested (Bowling, Hoeschele, Gill, & Fitch, 2017, 2018; Bowling & Purves, 2015; Harrison & Pearce, 2020; Jacoby et al., 2019; Large, Kim, Flaig, Bharucha, & Krumhansl, 2016; McBride & Tlusty, 2020; McDermott et al., 2010, 2016; Merker et al., 2018; Pfordresher & Brown, 2017). Nevertheless, scales facilitate harmony, where multiple voices/instruments combine consonantly – another design feature supporting group coordination but not solo performance.

### 2.2.3 Repetitive structure

The synchronization of rhythms and harmonization of pitches described above is facilitated and enhanced by the widespread use of repetitive musical structures (Savage et al., 2015). Structural building blocks can range from short rhythmic and/or melodic motives of only a few notes, to entire phrases, to large-scale sections or entire works. The level of repetition in music is one of its most striking differences from language (Fitch, 2006; Margulis, 2014), and multiple repetitions of a recording of a spoken phrase cause it to sound sung rather than spoken (Deutsch, Henthorn, & Lapidis, 2011). Repetition enhances memorization and predictability, allowing multiple performers to engage in long periods of coordinated music-making, with all-night music-and-dance rituals common from contemporary Western nightclub culture to ethnographic descriptions of small-scale societies (Merriam, 1964; Thornton, 1995). In contrast, language and ABMs such as laughter are more difficult to sustain for long periods, making them less suitable for the kind of sustained inclusive interactions that promote the strongest social bonds. However, extreme repetition can lead to boredom and to a dearth of memorable distinguishing features, preventing music from serving as a cue of social identity (see below). Both human and bird songs tend to balance repetition and novelty in the form of repetition with variation (Kroodtsma, 1978; Lomax, 1968).

### 2.2.4 Music and social identity

A final potential design feature of culturally-transmitted group music concerns its role in flexibly and hierarchically indicating kinship and group identity (Stokes, 1994; Turino, 2008). Because songs are variable, complex, and memorable, two people knowing the same song likely acquired this knowledge via social learning – and thus are likely to share a common socio-cultural history. Thus, shared knowledge of musical repertoire provides information about shared socio-cultural background (Schachner et al., Preprint; Soley & Spelke, 2016). Musicality may have coevolved in support of this social bonding function: Cultural innovations created a wide variety of musical styles and features, and musical knowledge became a cue to social history and cultural group membership. This created selective feedback favoring individuals who tended to perceive music as a cue to group membership, as they would have more accurate ideas about others’ social group membership.

This hypothesized combination of cultural and biological evolution would lead to an evolved bias to use music as a cue to guide and facilitate social interactions, consistent with findings that shared musical knowledge serves as a social cue from early in childhood through adulthood (see sects. 3.3 and 3.4).

Synchronized and harmonized group performances help cement group identity, and eventually allow skilled participation in ritualized performances to serve as a hard-to-fake indicator of group membership. Furthermore, the existence of diverse pieces and sub-styles allows subgroups to express their uniqueness within a broader shared musical repertoire or style. Such expressions of identity at multiple hierarchical levels are useful because human biological and cultural evolution has been characterized by increasing complexity of social structure, as exemplified by the large-scale nation-states characteristic of modern human societies (Turchin et al., 2018). Thus, group musical performance – including dance – facilitates lasting, culturally evolving indicators of group identity and bonds – akin to passwords or shibboleths (cf. Feekes, 1982; Fitch, 2004) – that extend beyond individual recognition and memory, aiding intercultural marriage and trade.

### 2.3 Gene–culture coevolution

These specific design features and their interactions – dancing to an isochronous beat with a metrical hierarchy, singing learned melodies based on discrete scales in harmony, using predictable, repetitive musical structures, and using musical performances as cues for social identity – are widespread throughout the world's musical systems (Savage et al., 2015; see sect. 3.1). These features have clear functions for group performance, but little or no function in solo performance (hence their rarity in birdsong, whale song, and certain solo human music genres such as lament; Frigyesi, 1993; Tolbert, 1990). These design features are therefore predicted *a priori* by the MSB hypothesis, but not by solo signaling hypotheses such as sexual selection for mate attraction (Miller, 2000) or maternal singing to infants (Mehr & Krasnow, 2017; Mehr et al., target article). Although these features promote coordination in dyadic music (e.g., duets) and memorability/communicative power in solo music (e.g., lullabies; Cirelli & Trehub, 2020; Corbeil, Trehub, & Peretz, 2016), their added value in supporting extended, coordinated group performances is most evident for larger groups.

MSB posits an extended timeline in which different core mechanisms of musicality arose through a coevolutionary “virtuous spiral.” Although many of the specific design features above could in principle function independent of the others, and would prove adaptive independently at any proto-musical stage, over evolutionary time we hypothesize that isochronous beats coevolutionarily enabled meter and dance, and that pitched singing enabled scale-based melody and harmony. Each new feature added value in supporting extended, coordinated, harmonious group performance. Each feature may have been initially based on behavioral innovations involving synchronization of the ancestrally individualistic displays seen in other great apes (e.g., chimpanzee pant-hoot displays and fruit tree “carnival” displays, cf. Merker, 1999; Merker et al., 2018). However, each innovation opened a new cognitive/musical niche selecting for independent specialization of relevant neural circuitry (see sect. 4).

Early instantiations of music provided selective preconditions for later cognitive and neurobiological changes underlying human musicality, analogous to the well-documented examples of gene–culture coevolution involving fire and dairy farming.

Cultural innovations created a variety of proto-musical behaviors, with musical knowledge becoming a potential cue to social history and cultural (sub-)group membership. For example, this could have created selective feedback favoring individuals who used music as cues to group membership. Together, biological and cultural coevolution created a framework for the coordinated, harmonious, emotional group performances that are evident today throughout the world's musical cultures. The major inter-relationships among these components of human musicality are summarized in Figure 2 (but see sect. 6.3 for caveats regarding causality in our proposed coevolutionary mechanisms).

### 2.4 Benefits of social bonding

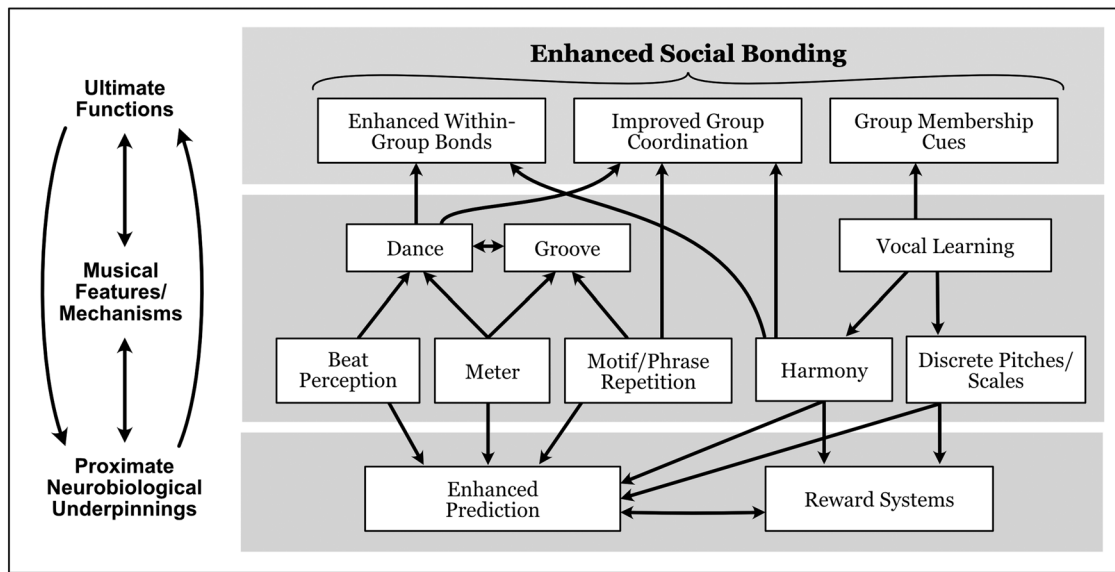
We hypothesize that musicality increased the number of “simple” relationships (e.g., “friends”), and increased the quality (depth and complexity) of existing relationships. The opportunity for many individuals to participate productively in social interaction through proto-musical behaviors facilitates an efficient bonding mechanism for groups of varying sizes, thereby conferring associated benefits (as outlined in sect. 2.1). However, we must consider the nature of the subsidiary relationships and social structures in which they operate. Many vertebrate species live in large groups (e.g., fish schools, bird flocks, and ungulate herds), but do not exhibit strong social bonds with more than a small number of individuals, and/or the relationships are undifferentiated. Indeed, the “number of differentiated relationships” (Bergman & Beehner, 2015) can vary independently from raw group size. For example, a monogamous pair with bi-parental care involves two differentiated relationships (sexual mate, and caregiving partner) or even three (adding joint territory defense), a situation typical in many birds. The social bonding design features we have identified can operate at multiple levels simultaneously, in the same way that a couple dancing at a party can intensify their own relationship, and their relationship with the broader social group.

### 2.5 Participatory versus presentational music

For most of hominin evolution, the only way to experience music was to make it oneself, or to observe others making music in real time. But as music-making technology culturally evolved, opportunities for solo listening increased (e.g., recording technology and personal music-playing devices) and individual virtuosity became increasingly emphasized. Cross-cultural analyses suggest that forms of music-making coevolved in parallel with social structures: larger-scale, more hierarchical societies tend to emphasize “presentational” music made by small numbers of performers for large numbers of passive (or virtual) audiences. Conversely, smaller-scale, more egalitarian societies tend to emphasize “participatory” music in which large groups sing, dance, and play instruments together with little or no distinction between performers and audience (Lomax, 1968; Turino, 2008). Once group size increases substantially, it may not be feasible for all individuals to participate actively in a coordinated manner, but music can facilitate bonding via passive (including digital) participation. This enables music (e.g., national anthems) to help construct social identities even among massive “imagined communities” (Anderson, 1991) whose members may never physically interact with one another.

The participatory mode of musical performance is hypothesized to be the ancestral one that operated over long time scales. It is imperative to avoid conflating pervasive technology-driven





**Figure 2.** Proposed coevolutionary relationships among multiple musical features and mechanisms, indicating their contributions to ultimate functions by facilitating social bonding in multiple ways, their proximate neurobiological underpinnings in prediction and reward systems, and feedback loops among these different levels.

aspects of contemporary musical practice (e.g., static audiences, solo listening, and control by global corporations) with the conditions under which humans experienced music during most of our evolutionary history. As a result, testing predictions of the MSB hypothesis should favor contexts such as drumming circles, campfire singalongs, and folk dances over solo-listening via headphones, or collective, static listening at a Mozart performance. Even in societies dominated by presentational music, participatory contexts retain their social and emotional potency, as highlighted by the collective singing of Italians from their balconies during the coronavirus lockdown (Grahn, Bauer, & Zamm, 2020; Horowitz, 2020; Kornhaber, 2020).

## 2.6 Summary

Summarizing, the MSB hypothesis argues that music is a derived bonding mechanism, akin to but augmenting previous ABMs such as grooming and laughter. This augmentation occurs via the provision of a shared framework for individual participants to establish and maintain strong bonds with more than one individual (or a small group of individuals) at a time, thus bridging the “bonding gap” problem posed during human evolution by increasing group size and complexity (Dunbar, 1993, 2012b). Proto-musical features may initially have arisen as behavioral innovations that later initiated a process of gene–culture coevolution. Crucially, the design features of musicality discussed above make music better suited than ABMs or language for coordinating behavior and facilitating social bonding in larger and more complex groups.

## 3. Cross-disciplinary evidence

Evidence in support of the MSB hypothesis comes from cross-cultural, historical/archeological, developmental, and social psychological research.

### 3.1. Cross-cultural evidence

One line of evidence for the MSB hypothesis comes from the study of cross-cultural musical universals (Brown & Jordania, 2013; Lomax, 1968; Mehr et al., 2019; Nettle, 2015; Savage, 2018, 2019b; Savage & Brown, 2013; Stevens & Byron, 2016; Trehub et al., 2018). Music, like language, is a human universal found in all known cultures (Brown, 1991; Mehr et al., 2019). Few if any specific musical features are found in all known musics, just as few specific linguistic features are found in all known languages (Evans & Levinson, 2009). However, researchers have identified dozens of “statistical universals” that predominate throughout diverse samples of the world’s music, relating both to functional context and to musical structure (Mehr et al., 2019; Savage et al., 2015; Table 1). These cross-cultural similarities suggest selection by biological and/or cultural evolution.

Crucial to our hypothesis, music performs similar social bonding functions across cultures. All of the 20 widespread functional contexts supported by at least one analysis in Mehr et al. (2019) summarized in Table 1 relate to social bonding, particularly through the ubiquitous use of music in communal ceremonies and rituals (e.g., healing, procession, mourning, storytelling, greeting visitors, praise/religion, and weddings). Even the secular use of music as art or entertainment is itself often a form of communal ritual. For example, aspects of Western art music concert attendance function to cement social bonds between participants and exclude non-participants in similar ways to other elite rituals throughout history (Nooshin, 2011; Small, 1998). Other non-ritual contexts have social bonding functions in bringing together parents and infants (lullabies and play songs), mates (love songs), or coordinating activities among multiple individuals (work songs and dance music). Finally, regulation of moods/emotions is one of the key components of our definition of social bonding (“...synchronizing and harmonizing the moods, emotions, actions, or perspectives of two or more individuals”). Even mood regulation via solo music can support social functions or evoke social contexts. For example, people may ease the pain of separation from

**Table 1 (Savage et al.).** Cross-culturally widespread musical structures and functions

| <b>Functional context</b> (from Mehr et al., 2019)  |  |
|---|--|
| (1) Dance   | (15) War <sup>a</sup>  |
| (2) Infancy   | (16) Praise <sup>a</sup>   |
| (3) Healing   | (17) Love <sup>a</sup>   |
| (4) Religious activity                              | (18) Group bonding <sup>a</sup>  |
| (5) Play  | (19) Marriage/weddings <sup>a</sup>  |
| (6) Procession                                      | (20) Art/creation <sup>a</sup>   |
| (7) Mourning  |  |
| (8) Ritual  |  |
| (9) Entertainment                                   |  |
| (10) Children                                       |  |
| (11) Mood/emotions                                  |  |
| (12) Work   |  |
| (13) Storytelling                                   |  |
| (14) Greeting visitors                              |  |
| <b>Musical structure</b> (from Savage et al., 2015) |  |
| (1) Group performance                               | (15) Voice use   |
| (2) Isochronous beat                                | (16) Modal register (chest voice)  |
| (3) <i>Metric hierarchy</i>                         | (17) <i>Word use</i>   |
| (4) <i>2- or 3-beat subdivisions</i>                | (18) Male performers   |
| (5) <i>2-beat subdivisions</i>                      | (19) Co-occurrence of: dance accompaniment, group performance, isochronous beats, percussion instruments, few duration values, motivic rhythms, repetitive phrases, syllabic singing |
| (6) Few durational values (<5)                      |  |
| (7) Motivic rhythms                                 |  |
| (8) Discrete pitches                                |  |
| (9) $\leq 7$ scale degrees                          |  |
| (10) <i>Unequal scales</i>                          |  |
| (11) Small intervals (<750 cents)                   |  |
| (12) Descending/arched contours                     |  |
| (13) Short phrases (<9 s)                           |  |
| (14) Instrument use                                 |  |

Functional contexts were found by Mehr et al. (2019) to be associated with singing in ethnographic descriptions of the 60 societies from the Human Relations Area Files Probability Sample (Lagacé, 1979). Musical structures were found by Savage et al. (2015) to predominate (items 1–18) or to co-occur (item 19) consistently in each of nine world regions across a sample of 304 audio recordings from the Garland Encyclopedia of World Music (Nettl, Stone, Porter, & Rice, 1998–2002). Nested relationships are indicated with indented italics.

<sup>a</sup>Indicates associations that were only significant using one of the two methods reported by Mehr et al. (2019) (Mehr et al. [2019] used two methods to examine universal associations with singing: “topic annotations from the Outline of Cultural Materials [‘OCM identifiers’] and automatic detection of related keywords.” The second method was needed “because some hypotheses correspond only loosely to the OCM identifiers (e.g., ‘love songs’ is only a partial fit to ARRANGING A MARRIAGE [the OCM identifier used] and not an exact fit to any other identifier).” Similarly, “group bonding” is only a partial fit to the OCM identifier “SOCIAL RELATIONSHIPS AND GROUPS,” which covers a broader range of social behaviors than simply “group bonding.” After adjusting for ethnographer bias and multiple comparisons, Mehr et al. found “support from both methods for 14 of the 20 hypothesized associations between music and a behavioral context, and support from one method for the remaining six.” See Mehr et al. [2019] for further details).

loved ones by listening to or playing music that evokes shared memories (Kornhaber, 2020), or use music to prepare their mood for an effective social interaction, allowing them to regulate their behavior and behave in the socially-expected manner (Erber, Wegner, & Theriault, 1996; Greenwood & Long, 2009).

Similarly, most of the widespread structural aspects of music support coordinated music-making. Throughout the world, humans tend to sing, play percussion instruments, and dance to simple, repetitive music in groups, and this is facilitated by the widespread use of simple-integer pitch and rhythm ratios, scales

based on a limited number of discrete pitches ( $\leq 7$ ), and isochronous beats grouped in multiples of two or three (Bowling & Purves, 2015; Jacoby & McDermott, 2017; Jacoby et al., Preprint; Kuroyanagi et al., 2019; Ravignani, Delgado, & Kirby, 2017; Savage et al., 2015). The widespread use of simple, discrete meters and scales also enables multiple people to memorize and coordinate their performances. These widespread musical properties have few direct parallels in language. Group coordination provides a common purpose that unifies the cross-cultural structural regularities of human music (Savage et al., 2015).



### 3.2 Fossil and archeological evidence

Although music itself leaves no fossil record, inferences can be drawn from evidence about the evolution of musicality, the role this played in early human society, and its relationship to other evolutionary developments such as brain size, language, group size, and sociality (Mithen, 2005; Morley, 2013). The fossil record for human evolution indicates that capacities for sophisticated and diverse vocalizations and body language, including dancing, were present before there is credible evidence for compositional language (as reviewed in Mithen, 2005). Archeological evidence from the Paleolithic indicates increasing group size and long-distance contacts (Gamble, 2010; Read & Van der Leeuw, 2015), suggesting that ABMs had become insufficient by at least 2 million years ago. The earliest surviving musical instruments – bone flutes – have been dated to over 35,000 years ago and are speculated to have functioned to support larger social networks (Conard, Malina, & Münzel, 2009). Prehistoric rock art often appears to be positioned with regard to the acoustic properties of either the cave or cliff face on which it is located (e.g., Fazenda et al., 2017; Rainio, Lahelma, Aikas, Lassfolk, & Okkonen, 2018), suggesting that music played a role in the social-bonding rituals associated with that art. Similarly, prehistoric and early historic architecture used for social-bonding ceremonies often appears to have been designed with regard to its acoustic properties and to facilitate music making (e.g., Göbekli Tepe: Notroff, Dietrich, & Schmidt, 2015; Stonehenge and other Neolithic monuments in Britain: Banfield, 2009; Watson & Keating, 1999; and Ancient Mayan temples: Sanchez, 2007).

### 3.3. Developmental evidence

Extensive evidence demonstrating spontaneous and early development of social functions of music also supports the MSB hypothesis. Adults around the world produce infant-directed songs, such as lullabies, with similar, cross-culturally recognizable acoustic features (Mehr, Singh, York, Glowacki, & Krasnow, 2018; Trehub, Unyk, & Trainor, 1993). Song is highly effective at emotional modulation in infants – reliably more effective than speech, with infants exhibiting longer visual fixations and greater reductions in stress and body movement to maternal singing than to speaking (Cirelli & Trehub, 2020; Corbeil et al., 2016; Ghazban, 2013; Nakata & Trehub, 2004; Trehub, 2016). Infants also respond differently to songs sung in different styles (e.g., lullaby vs. play-song; Cirelli, Jurewicz, & Trehub, 2019; Rock, Trainor, & Addison, 1999). Singing to infants thus appears to serve a communicative function, allowing parents to communicate specific emotional messages to infants before they can understand the semantic content of language (Rock et al., 1999; Trainor, Clark, Huntley, & Adams, 1997; Trehub et al., 1997). Singing and musical interactions also directly improve parent–infant social bonds: Interventions promoting singing and musical interaction between parents and infants strengthen parents' attachment to their infants, more so than nonmusical play (Vlismas, Malloch, & Burnham, 2013). Music thus facilitates both parent–infant communication and parent–infant bonding from early in life, before extensive experience or opportunities for learning.

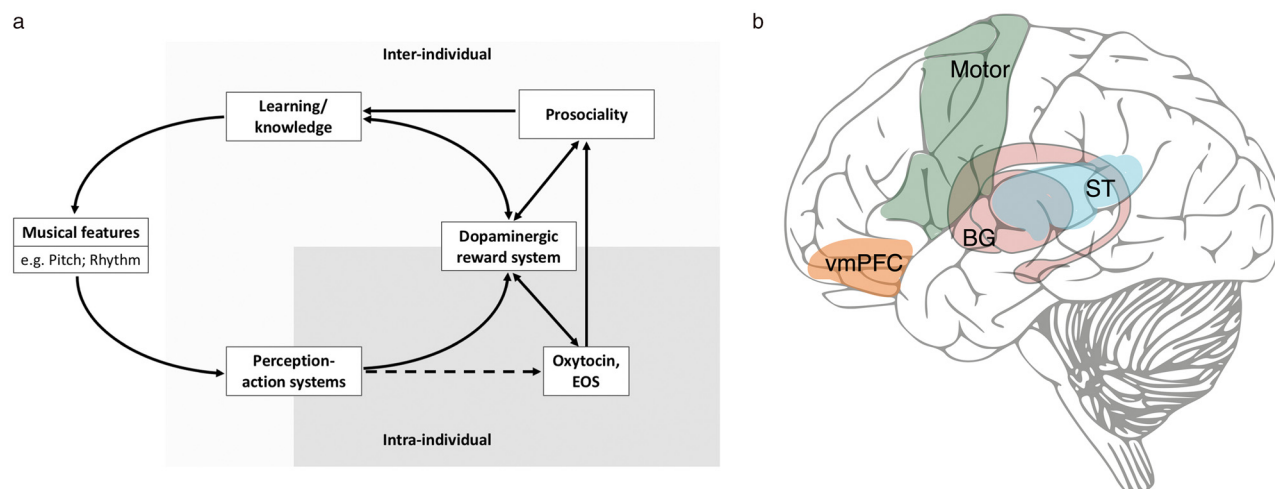
Beyond infancy, musical activities continue to promote bonding: Across a range of tasks, group musical involvement increases children's prosocial behavior. Thus, young children act more prosocially (in terms of sharing and fairness) after a musical game than a similar non-musical game (Kirschner & Tomasello,

2010); after group singing than group art or competitive games (Good & Russo, 2016); and after joint synchronized, rhythmic movement than non-synchronized movement (Rabinowitch & Meltzoff, 2017).

Children (like adults) choose to affiliate with members of their own social group (Bigler, Jones, & Lobliner, 1997). From early infancy, music serves as a marker of social group membership, allowing for the identification of preferred social partners (Cirelli, Trehub, & Trainor, 2018). Shared knowledge of specific songs serves as a particularly informative signal of common group membership: because of the wide range of forms a song can take, knowledge of a particular song implies common social or cultural background (Soley & Spelke, 2016). Infants accordingly treat shared musical knowledge as socially meaningful from early in life: 5-month-old infants prefer to look at people who sing melodies previously sung by a parent, over people who sing melodies previously sung by an unfamiliar adult (Mehr, Song, & Spelke, 2016). These early preferences appear to form the foundation for selective social affiliations based on music: At preschool age, children use knowledge of a familiar song as a social cue to select friends (Soley & Spelke, 2016), and by 14 months exhibit more prosocial behavior (helping) toward an unfamiliar woman who sings a familiar song (previously sung by a parent) than an unfamiliar song (Cirelli & Trehub, 2018). Together, these results suggest that musical knowledge shapes the formation of children's social bonds, and that the link between shared musical knowledge and social connection is rooted in early infancy.

### 3.4. Social psychological evidence

Behavioral experiments from social psychology support the MSB hypothesis, suggesting that musical behavior is not only associated with, but may causally support, social bonding. In particular, music provides a foundation for synchronized behavior in large groups (as argued above), and a number of experiments and meta-analyses show that rhythmic synchronization with other individuals promotes increased prosocial behavior (i.e., actions that increase others' well-being; Mogan, Fischer, & Bulbulia, 2017; Rennung & Göritz, 2016). Synchrony has been empirically linked to cooperation in economic games (Lang, Bahna, Shaver, Reddish, & Xygalatas, 2017; Launay, Dean, & Bailes, 2013; Reddish, Bulbulia, & Fischer, 2014; Wiltermuth & Heath, 2009), entitativity (feelings of being on the same team; Lakens & Stel, 2011; Reddish, Fischer, & Bulbulia, 2013), rapport and interpersonal liking (Hove & Risen, 2009; Miles, Nind, & Macrae, 2009; Valdesolo & Desteno, 2011), and helping behavior (Cirelli, Einarson, & Trainor, 2014; Kokal, Engel, Kirschner, & Keysers, 2011; Valdesolo & Desteno, 2011). Similarly, dancing in synchrony increases participants' feelings of connectedness to the group with which they are dancing, as well as their liking and assessment of similarity with co-dancers (Tarr, Launay, Cohen, & Dunbar, 2015; Tarr, Launay, & Dunbar, 2016). These prosocial effects of synchrony are robust in different contexts (Mogan et al., 2017). Although demand characteristics have been suggested as possible confounds underlying these effects (Atwood, Mehr, & Schachner, 2020; Rennung & Göritz, 2016), significant prosocial effects of synchrony remain after potential confounds of suggestion, competence, and shared intention are eliminated (e.g., in a virtual reality setting; Tarr, Slater, & Cohen, 2018). However, meta-analyses implied inconclusive results regarding the precise roles of "music" and of synchrony to an isochronous beat, as



**Figure 3.** (a) Proposed neurobiological mechanisms underlying music's social bonding functions, showing intra- and inter-individual levels. We propose that the dopaminergic reward system interacts with the endogenous opioid system (EOS) and the release of oxytocin, ultimately providing opportunities for individuals to synchronize their moods, emotions, actions, and/or perspectives through musical engagement (dashed arrow indicates need for more evidence to confirm that the perception/production of music stimulates this pathway). (b) Key neuroanatomical regions in the human brain underlying the MSB hypothesis. ST: superior temporal lobe structures important for auditory perception including Heschl's gyrus, planum temporale, superior temporal gyrus, superior temporal sulcus, middle temporal gyrus. Motor: frontal lobe structures crucial for action planning and execution including premotor and supplementary motor areas as well as primary motor cortex. BG: basal ganglia and related structures, including amygdala, striatum, ventral tegmental area/substantia nigra, caudate, putamen, globus pallidus, and nucleus accumbens. vmPFC: ventromedial prefrontal cortex.

opposed to more generally synchronized or coordinated non-musical behaviors such as gaze synchrony, affect synchrony, and motor synchrony (Mogan et al., 2017; Rennung & Göritz, 2016). In sect. 5, we propose clearer predictions and tests of specific mechanisms by which music promotes social bonding.

More broadly, behavioral studies indicate varied social bonding effects associated with music-based activities, even those that do not explicitly involve constant synchrony. Young children randomly assigned to activities incorporating music exhibit elevated levels of empathy compared to non-musical controls in longitudinal studies (Rabinowitch, Cross, & Burnard, 2013), and adults singing in regular group sessions develop feelings of social closeness toward co-participants more quickly than people engaged in other (non-musical) group activities (Pearce, Launay, & Dunbar, 2015). Feelings of inclusion, connectivity, and positive affect emerge in small and large singing groups, with participants in large choirs (>80 participants) reporting greater changes in these measures compared to smaller choirs (Weinstein, Launay, Pearce, Dunbar, & Stewart, 2016). These findings highlight the relevance of music-based activities for large-scale social bonding.

#### 4. Neurobiological mechanisms

The MSB hypothesis proposes that social bonding is the ultimate, functional explanation of the evolution of musicality. We now propose specific hypotheses about underlying neurobiological proximate mechanisms underpinning music's social effects (Fig. 3). In brief, music involves predictable combinations of rhythms and pitches, activating neural mechanisms for perception that are tightly coupled with the motor system. Learning to form predictions about these features activates the dopaminergic reward system, which synchronizes its activity with distal regions within the brain. Crucially, predictability also supports synchronization of homologous regions in other individuals' brains. This

"neural resonance" (synchronous brain activity across individuals) facilitates social bonding through shared experience, joint intentionality, and "self-other merging." Through the production of oxytocin and endogenous opioids, neural resonance also facilitates prosociality by associating the rewarding musical experience with specific co-experiencers. Furthermore, because these prosocial experiences are themselves rewarding, we seek them out by attending to and learning more musical features/experiences, updating our predictions (e.g., through statistical learning, by performing and/or experiencing new music), and closing the mechanistic cycle. This proposed mechanistic cycle is detailed below.

##### 4.1 Perception-action coupling

Perception-action coupling refers to anatomical and/or functional connectivity between brain regions involved in sensory perception (e.g., of pitch or rhythm) and those that are involved in movement (e.g., vocalization and dance). Specifically, auditory-motor coupling is a key neural mechanism that underlies social bonding through music because it enables individuals to synchronize and/or harmonize their own music and actions with others, which is crucial for coordinated group music making. Even during the perception of solo music, the tight coupling between perceptual and motor regions leads to spontaneous and obligatory activity in premotor and supplementary motor areas, classic motor areas that are also part of the action observation network that drives physical and observational learning (Cross, Kraemer, Hamilton, Kelley, & Grafton, 2008).

Rhythm and beat consistently activate the premotor area, supplementary motor area, and basal ganglia, regions commonly thought to belong to the motor system (Grahn & Brett, 2007). Furthermore, the auditory system is strongly coupled with areas in the motor system during rhythm perception (Grahn & Rowe, 2009), and rhythmic oscillatory activity in both the auditory and motor systems tracks the rhythm of music (Fujioka, Ross,

& Trainor, 2015). Some observations show that neural phase-locking activity is even higher in music than in speech (Vanden Bosch der Nederlanden, Joannisse, & Grahm, 2020). This process of “neuronal entrainment” (neural activity changing its frequency, amplitude, and/or phase in response to external stimulation) is a proposed mechanism through which rhythm in sensory stimuli affects the brain by coordinating activity between separate neuronal populations, such as between the auditory and motor systems (Jones, 2018; Morillon & Baillet, 2017). This neuronal entrainment enables selective attention to specific points in time (Lakatos et al., 2008; Large & Jones, 1999). In particular, auditory–motor coupling is strongest when perceiving “high groove” music that elicits the pleasurable drive toward action such as in dance (Janata, Tomic, & Haberman, 2012). Groovy music elicits the urge to dance by increasing the auditory cortex’s sensitivity and its coupling with the motor cortex (Stupacher, Hove, Novembre, Schutz-Bosbach, & Keller, 2013), which is particularly evident with medium levels of rhythmic complexity and expectation violation (Koelsch et al., 2019; Witek et al., 2014). In this respect, dance – or any movement to music – is inextricably linked to musical experiences. Note, however, that similar to many of the mechanisms proposed here, coding of value in sensory cortices (i.e., a stronger sensory response to more important or rewarding stimuli) is not unique to the auditory domain but is also evident in other sensory domains such as vision (Koelsch et al., 2019).

An important pathway underlying perception–action coupling is the arcuate fasciculus, a bundle of axonal connections between frontal lobe (including motor areas) and superior temporal lobe (including auditory areas). Abundant neuroimaging evidence supports the role of the arcuate fasciculus in music making, specifically in auditory perception–action coupling (Halwani, Loui, Rüber, & Schlaug, 2011; Loui, Alsop, & Schlaug, 2009, 2011; Moore, Schaefer, Bastin, Roberts, & Overy, 2017; Sammler, Grosbras, Anwander, Bestelmeyer, & Belin, 2015). This same pathway also plays a role in social functions: more emotionally empathic people have higher microstructural integrity within the arcuate fasciculus (Parkinson & Wheatley, 2014). In contrast, people on the autism spectrum, who have known impairments in social bonding, have less connectivity in the arcuate fasciculus (Fletcher et al., 2010; Wan, Demaine, Zipse, Norton, & Schlaug, 2010). By enabling perception–action coupling, the arcuate fasciculus thus provides one possible shared neural mechanism between music and social bonding.

#### 4.2 Prediction and the dopaminergic reward system

Musical perception–action coupling sets up repeated cycles of prediction, expectation violation, and resolution (Huron, 2006). In these hierarchical perception–action trajectories, the predictive context surrounding pitch and rhythm are established, violated, and then resolved (Clark, 2013; Fitch, von Graevenitz, & Nicolas, 2009). Successful predictions become rewarding to the brain by activating neurons of the dopaminergic system and its related areas (caudate, nucleus accumbens, amygdala, and ventromedial prefrontal cortex) that code for fundamental evolutionary rewards such as food and sex, and also learned rewards such as money (Friston, 2010; Knutson, Westdorp, Kaiser, & Hommer, 2000; Schultz, Dayan, & Montague, 1997). The same dopaminergic reward system is also active during the anticipation and perception of pleasurable music (Blood & Zatorre, 2001; Blood, Zatorre, Bermudez, & Evans, 1999; Cheung et al., 2019; Salimpoor,

Benovoy, Larcher, Dagher, & Zatorre, 2011, 2015; Zatorre, 2018; Zatorre & Salimpoor, 2013), supported by the functional coupling between auditory areas in the superior temporal lobe and reward-sensitive areas such as the nucleus accumbens (Salimpoor et al., 2013). Manipulating expectations for pitch-related musical features, such as consonance and dissonance, can modulate activity in the nucleus accumbens and amygdala. Thus, music can provide its own reward prediction error and motivate learning (Cheung et al., 2019; Gold et al., 2019). Additionally, people who frequently experience chills when listening to music show high white matter connectivity between auditory, social, and reward-processing areas (Sachs, Ellis, Schlaug, & Loui, 2016). Chills from music are also related specifically to increased binding to dopamine receptor D2 (Salimpoor et al., 2011). In contrast, people with musical anhedonia, who find music unrewarding, have decreased functional connectivity and altered structural connectivity between auditory and reward-related areas (Loui et al., 2017; Martínez-Molina, Mas-Herrero, Rodríguez-Fornells, Zatorre, & Marco-Pallarés, 2016; Mas-Herrero, Zatorre, Rodríguez-Fornells, & Marco-Pallarés, 2014).

Because humans are social animals, the predictions we make and the rewards we receive are often tied to social stimuli. Thus, the brain has to learn from social cues by associating social stimuli with reward predictions (Atzil, Gao, Fradkin, & Barrett, 2018). Indeed, the same areas in the dopaminergic reward system – the caudate, nucleus accumbens, and ventromedial prefrontal cortex – are causally linked to cooperative behavior as well as prediction and reward. The reward system is activated when we share information with others about ourselves (Tamir & Mitchell, 2012), when we view loved ones (Bartels & Zeki, 2004), and when mothers bond with their infants (Atzil et al., 2017). Prosocial behaviors commonly engage the reward system (Zaki & Mitchell, 2013); these include cooperating (Decety, Jackson, Sommerville, Chaminade, & Meltzoff, 2004), perspective taking (Mitchell, Banaji, & Macrae, 2005), and empathizing with others (Beadle, Paradiso, & Tranel, 2018). Together, these results suggest that the dopaminergic reward system is involved causally in the link between music and social bonding through the mechanism of prediction.

#### 4.3 Oxytocin and the endogenous opioid system (EOS)

We propose that opioids released in the EOS, and oxytocin, are also part of the mechanistic underpinnings linking prediction, reward, and social bonding (Chanda & Levitin, 2013; Launay et al., 2016; Tarr, Launay, & Dunbar, 2014). The nucleus accumbens and ventral tegmental area are key regions that overlap between the dopaminergic reward system and the EOS (Dölen, Darvishzadeh, Huang, & Malenka, 2013; Le Merrer, Becker, Befort, & Kieffer, 2009), and dopamine is thought to be a salience processing mechanism regulated by oxytocin (Love, 2014; Shamay-Tsoory & Abu-Akel, 2016).

The EOS likely plays a mechanistic role in music-related prosociality. This system has been implicated in the maintenance of social bonds in primate social networks (Keverne, Martensz, & Tuite, 1989; Maestripieri, 2010; Ragen, Maninger, Mendoza, Jarcho, & Bales, 2013; Schino & Troisi, 1992). Intervention studies in humans indicate that, compared to a placebo, naltrexone (an opioid blocker) can reduce feelings of social connections with others (e.g., Inagaki, 2018; Inagaki, Ray, Irwin, Way, & Eisenberger, 2016), and lower affiliative behavior and desire for interpersonal closeness (Tchalova & Macdonald, 2020). Listening to music influences mu-opiate receptor expression in



the EOS (Stefano, Zhu, Cadet, Salamon, & Mantione, 2004) and can reduce the need for pain medication<sup>3</sup> (e.g., Bernatzky, Presch, Anderson, & Panksepp, 2011; Lepage, Drolet, Girard, Grenier, & DeGagné, 2001). Elevated pain thresholds are experienced after singing (Pearce et al., 2015; Weinstein et al., 2016) and synchronized dancing (Tarr et al., 2015, 2016), but not after administration of naltrexone (Tarr, Launay, Benson, & Dunbar, 2017), suggesting that pain threshold is an appropriate proxy-measure of endorphin uptake in these experiments. There is some evidence of endorphin-mediated synchrony effects on cooperation (e.g., when dancing; Lang et al., 2017), further demonstrating links between music, the EOS, and social bonding.

Although more empirical research is needed, there is evidence that oxytocin levels are elevated after taking part in a singing class (Grape, Sandgren, Hansson, Ericson, & Theorell, 2003), or following a group jam session of improvised singing (Keeler et al., 2015). Elevated oxytocin levels have been correlated with increased generosity (Fujii, Schug, Nishina, Takahashi, & Okada, 2016; Zak, Stanton, & Ahmadi, 2007), empathy (Domes, Heinrichs, Michel, Berger, & Herpertz, 2007; Hurlmann et al., 2010), and possibly trust (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; Zak, Kurzban, & Matzner, 2005; but see Nave, Camerer, & McCullough [2015] and Declerck, Boone, Pauwels, Vogt, & Fehr [2020]). Furthermore, intranasal administration of oxytocin promotes in-group cooperation (e.g., De Dreu and Kret, 2016) and increases synchrony in dancing (Josef, Goldstein, Mayseless, Ayalon, & Shamay-tsoory, 2019) and finger-tapping behavior (Gebauer et al., 2016), suggesting a reciprocal feedback loop between music-based activity and social cohesion. Although evidence linking oxytocin specifically with music remains limited, and the strength of oxytocin's relationship with cooperation more generally is debated (particularly studies based on administering intranasal oxytocin; e.g., Walum, Waldman, & Young, 2016), current evidence suggests that music engages the oxytocin and EOS systems in ways that facilitate social bonding, as predicted by the MSB hypothesis. Combined with the reward system, these pathways offer a positive-feedback loop following music engagement, enabling groups of individuals to synchronize their moods, emotions, actions and/or perspectives, and providing motivation to continue engaging with others in social and musical contexts.

#### 4.4 Learning and vocal imitation

The capacity to learn and reproduce complex motor movements, including vocalizations (songs), is central to the cultural transmission of music. Although humans are the only primates capable of learning complex, novel vocalizations, this ability has evolved independently at least seven times in evolutionary history (Fitch & Jarvis, 2013; Nowicki & Searcy, 2014; Sval & Finlay, 2011), allowing us to make inferences about how and why it evolved. Some vocal learning clades (seals, baleen whales, and some songbirds) show a strong male bias in vocal learning abilities consistent with sexual selection. However, such a bias is absent in most other vocal learners (parrots, elephants, toothed whales, many tropical bird species, and humans), suggesting that sexual selection cannot be the only factor driving the evolution of vocal learning (Fitch, 2006). Instead, learned animal songs (solo or duet) appear to serve multiple evolutionary functions within the umbrella of social bonding, including mate attraction, cementing and affirming social bonds *within* pairs or groups, and territorial functions including advertising the bonded group's

ability to repel outsiders (Geissmann, 1999; Haimoff, 1986; Wickler, 1980).

In vocal learning species, vocal imitation and song production are likely based on similar neurobiological mechanisms (Mercado, Mantell, & Pfordresher, 2014). Learning to reproduce pitches and rhythms accurately engages reward mechanisms, as shown by evidence that dopamine neurons encode performance error in songbirds (Gadagkar et al., 2016). Furthermore, the presence of a conspecific (of the opposite sex in this case) leads the male zebra finch to decrease variability of sung syllables; this syllabic structure is attributed to perception–action circuits analogous to the human superior temporal and motor structures (Fitch & Jarvis, 2013; Sakata & Brainard, 2008). Once individuals learn to produce musical features, they not only reproduce learned patterns of features, but also deviate from predicted combinations of features, for example by inventing new melodies (Wiggins, Tyack, Scharff, & Rohrmeier, 2018).

### 5. Predictions for future research

The MSB hypothesis predicts that core design features of music make it particularly well-suited to facilitate social bonding, and particularly effective in the bonding of large, complex groups. This leads to the following testable predictions.

#### 5.1. Cross-domain predictions (e.g., music, language, ritual)

The MSB hypothesis predicts that music (including dance) is better-suited to social bonding of large, complex groups than ABMs (grooming and laughter), language, or other non-acoustic bonding mechanisms such as shared decorations or non-musical ritual behaviors (e.g., praying together without music). Music should be more effective and/or efficient relative to other methods as group size and complexity increase, such that while making music in pairs might only produce a small increase in dyadic bonding relative to conversation, making music in larger, more complex groups of people (dozens or hundreds organized into differentiated sub-groups) should be more effective for collective bonding than language, laughter, grooming, and so on.

In a social species such as humans, many activities can develop and enhance social bonding, but we predict that bonding via non-musical methods such as language, ritual, or sports should be enhanced by the addition of musical components (e.g., religious services with group singing will result in stronger bonding than those that only involve group prayer). Different musical components are predicted to have synergistic effects such that – all things being equal – including more of these components (e.g., synchronized, harmonized singing and dancing in groups) will tend to increase bonding more than activities that only use one or a few (e.g., conversations or recitation in pairs).<sup>4</sup> We also predict that participatory musical performances will have significantly stronger effects than either non-participatory (e.g., performance for a static audience) or solo musical experiences (e.g., listening alone to recordings). Group size and complexity should have independent effects (e.g., singing in large choirs should produce greater bonding than singing in small choirs).

These predictions can be tested in controlled experiments and/or field studies along the lines of those discussed in sect. 3. Designing studies that control for specific similarities and differences between closely related domains such as music, language, and dance is challenging but not impossible. For



example, to control for the fact that languages have their own (non-isochronous) rhythms, Savage et al. (2020) had groups of participants simultaneously recite the lyrics to “Twinkle, Twinkle, Little Star” to an isochronous beat or in non-isochronous free rhythm. Savage et al. (2020) also propose additional manipulations that would allow this paradigm to test other specific predictions of the MSB hypothesis regarding the social bonding effects of melody, harmony, and dance (cf. Fig. 3 in Savage et al., 2020).

### 5.2. Cross-cultural predictions

The MSB hypothesis predicts that music’s social bonding functions should be distributed widely in space and time. Hence, the kinds of predictions described in sect. 5.1 regarding music’s superior social bonding power in large groups should apply consistently across cultures and throughout history. Furthermore, it predicts that musical contexts and structures that promote social bonding (e.g., coordinated, participatory group performances) will be more common across cultures than music produced by and for individuals. At the same time, the relative importance of participatory versus presentational music-making is predicted to vary cross-culturally as a function of social structure (because of limitations on simultaneous coordinated performance discussed in sect. 2.5). Smaller-scale, more egalitarian cultures should thus perform and value participatory music more than larger-scale, hierarchical cultures where presentational music should be more common and valued. Participatory versus presentational distinctions are analogous to those found in “imagistic” (high-intensity, small-scale) versus “doctrinal” (low-intensity, large-scale) religious rituals, respectively (Whitehouse, 2004), and are predicted to covary cross-culturally with these modes of religiosity. Even in cultures where music is often consumed passively by individuals (e.g., in Western culture, over headphones on personal listening devices), MSB predicts that music will be more effective than non-musical alternatives for social bonding purposes (cf. Rentfrow & Gosling, 2006). These predictions about cross-cultural use of music for social bonding could be tested in cross-cultural behavioral experiments (cf. Henrich et al., 2005; Jacoby et al., *Preprint*; Polak et al., 2018) or analysis of cross-cultural databases of recordings, artifacts, ethnographies, or questionnaires (cf. Lomax, 1968; Mehr et al., 2019; Savage, 2019c; Savage et al., 2015; Whitehouse et al., *preprint*; Wood et al., *Preprint*).

### 5.3. Cross-species predictions

The MSB hypothesis proposes that human musicality has been shaped by biological and cultural selection, and that the features of music are particularly well-suited for social bonding functions because they support extended, coordinated group performances on a large scale. The MSB account does not claim that music’s social bonding function is a unique biological adaptation specific to human musicality. Instead, it argues that music-like behaviors should enhance existing bonding mechanisms in other species as well. Thus, it predicts that, rather than an all-or-nothing divide between human and non-human “music,” species will vary continuously in the degree to which they share specific features of human musicality. The social bonding functions associated with different components of musicality should operate similarly across species, depending on the specific subcomponent, its suitability for group coordination, and the importance of social bonding to their species.

Thus, melodic, learned songs among songbirds, whales, or other vocal learners are predicted to enhance social bonding in these species in a manner analogous to song in humans. These effects may be limited in many non-human species by their lack of ability and/or interest in performing in coordinated groups (e.g., some primates appear motivated to conduct group displays but are unable to synchronize to a beat, whereas some birds appear able to move to a beat but are unmotivated to do so in groups in the wild; Hoeschele, Merchant, Kikuchi, Hattori, & ten Cate, 2018). However, such effects should be pronounced in species that perform duets (e.g., many birds, and duetting primates such as gibbons or titi monkeys; Haimoff, 1986; Hall, 2004). Conversely, social primates that do not typically perform in coordinated groups may nonetheless experience social bonding effects of “group” music when exposed to versions of their own vocalizations that have been artificially manipulated to be in synchrony/harmony. Such production/perception dissociations and other nuances of musicality could be tested in controlled cross-species experiments (cf. Hoeschele et al., 2018; Merchant, Grah, Trainor, Rohrmeier, & Fitch, 2015).

The MSB hypothesis posits that music and musicality provided a major means by which humans could coordinate behavior on a larger scale than dyads or small groups, allowing for the formation of larger socio-cultural groups. If true, and if different species share components of musicality to differing degrees, then across species, production or proficiency in “musical” behaviors should predict both the number and complexity of social bonds. For example, gelada baboons live in unusually large and complex groups for primates, and they also exhibit rhythmic and melodic vocal features that are unique among primates (Bergman, 2013; Gustison, Aliza, & Bergman, 2012; Richman, 1978, 1987). Similar to geladas, many parrot species live in large fission–fusion social groups, and members of the parrot clade show vocal imitation, call convergence, duetting, and the capacity for rhythmic synchronization (Balsby & Scarl, 2008; Bradbury, 2001; Scarl & Bradbury, 2009; Schachner et al., 2009). In both of these clades, pairs or mating “harems” form stronger bonds than those they share with the larger groups in which they are embedded (cf. Balsby & Scarl, 2008; Wanker, Sugama, & Prinage, 2005). Other species that live in complex fission–fusion groups and could provide evidence of specific design features are elephants and some odontocetes (e.g., orcas and bottlenose dolphins). Such species live in large, complex fission–fusion groups, and are documented vocal learners, but their possession of other design features of music (e.g., synchronization) have not been tested rigorously.

For many species, evidence for design features of musicality would count as evidence *against* our hypotheses. Examples include solitary species (e.g., many reptiles), species for whom groups consist only of mothers and dependent young (e.g., many carnivores), or group living species that do not have differentiated social bonds with other group members (e.g., schooling fish, larger herds, and swarming insects).

The MSB hypothesis further predicts that if a species does not follow this pattern (e.g., by having a larger social group size than predicted by their features of musicality), then that species will have evolved other non-musical but effective means of coordinating behavior that likely do not appear in human behavior (e.g., reproductive suppression in naked mole rats or pheromonal queen control in eusocial insects; Alaux, Maisonnasse, & Le Conte, 2010; Dengler-Crish & Catania, 2007). Thus, although the social bonding design features seen in human musicality are not the only way to achieve large, well-bonded groups, they are

effective enough that we predict them to evolve convergently (cf. Fitch, 2006).

#### 5.4. Neurobiological predictions

The MSB hypothesis predicts that each of the mechanistic factors proposed above (Fig. 3) contributes to the effects of music on social bonding. Alterations of these mechanistic pathways should therefore produce specific, quantifiable results on bonding. For example, music's perceived social bonding functions should correlate with oxytocin/EOS production, and disrupting the oxytocin/EOS pathway via blocking oxytocin or opioid receptors should disrupt its social bonding effects. Furthermore, because of the dopaminergic reward system is at the center of prediction for musical features, populations with deficient dopaminergic activity may have impaired predictions, which could affect their ability to synchronize or harmonize with others. On the contrary, drugs that restore dopaminergic functions are hypothesized to restore these abilities, and because of the reciprocal nature of these interactions, activities that enhance predictions (such as dancing and harmonizing) may in turn restore dopaminergic functions. These predictions are being tested in the case of Parkinson's disease, which is a special population with deficient dopaminergic activity (Cameron et al., 2016; Grahn & Rowe, 2009).

Another prediction is that special populations with high sociability may respond well to musical features especially when coupled with social stimuli, as in the case of children with Williams syndrome (Järvinen-Pasley et al., 2010; Lense, Gordon, Key, & Dykens, 2014). At a neural level, music's social bonding function should correlate with the degree of neural connectivity between the perception-action and prediction-reward networks, and disruptions to this network (e.g., lesions or genetic syndromes) should accordingly disrupt music's social bonding effect. For example, people with musical anhedonia, who have disrupted connectivity between auditory prediction and reward networks (Belfi & Loui, 2020), are predicted to have weaker social bonds, and genetic differences (e.g., in DRD2) may predict variation in bonding experienced through musical activities. Although some of these predictions may be difficult to test ethically in humans through controlled experiments, many can be tested using neuroimaging combined with neuropsychological testing in special populations, as well as correlational, longitudinal, or intervention (including brain-stimulation) studies, genome-wide association studies, and/or animal models that share specific neurobiological endophenotypes (Finlay, Darlington, & Nicastro, 2001; Fitch & Jarvis, 2013; Gingras, Honing, Peretz, Trainor, & Fisher, 2018; Hoeschele et al., 2018; Niarchou et al., 2019).<sup>5</sup>

## 6. Potential criticisms

Having detailed our social bonding hypothesis and its predictions, we wish to preempt several potential criticisms.

### 6.1 Music, language, and domain-specificity

The key criticism that we anticipate regards the degree to which the evolution of musicality and social bonding are uniquely and causally linked. Few would deny that music can facilitate social bonding via neurobiological mechanisms that are evolutionarily adaptive. However, whether music is a domain-specific evolutionary adaptation for social bonding, as opposed to a byproduct of the evolution of other adaptations, is open to debate. Language,

in particular, has been proposed as an evolutionary adaptation that led to musicality as a byproduct (Pinker, 1997).<sup>6</sup> Importantly, many researchers have noted that, although there are clear differences in the structure and processing of music and language, there is extensive overlap ranging from structural content (e.g., “musilinguistic continua” between speech and song including intermediate forms like poetry and chant) to neurobiological substrates (e.g., similar neural substrates for processing of pitch, rhythm, and syntax; Brown, 2000b, 2017; Fitch, 2006; Patel, 2008; Peretz & Coltheart, 2003; Peretz, Vuvan, Lagrois, & Armony, 2018; Savage, Merritt, Rzesutek, & Brown, 2012). Indeed, many have proposed that the evolution of musicality may have paved the way for the evolution of language (Brown, 2000b, 2017; Darwin, 1871; Fitch, 2010; Mithen, 2005; Shilton, Breski, Dor, & Jablonka, 2020).

We accept that our present level of understanding is insufficient to demonstrate conclusively that music coevolved uniquely with social bonding independent from language or other social behaviors. Accordingly, in sect. 5, we proposed future investigations of such relationships. However, the fact that music and language are both found universally in all known societies (Brown, 1991; Mehr et al., 2019) suggests that both music and language independently fulfill more fundamental adaptive functions than technologies or cultural artifacts that are not cross-culturally universal.

We make no claim that the mechanisms discussed here are entirely specific to music, or that “musicality” is modular in either the cognitive or neuroscientific senses of this term. For example, prediction and predictive coding are ubiquitous features of vertebrate brains (Clark, 2013; Schultz & Dickinson, 2000), by no means specific to musicality. However, music affords a uniquely effective scaffolding framework, including rhythm and harmony, within which neural prediction (and occasional expectation violations) can unfold (Fitch et al., 2009; Hanslick, 1858; Huron, 2006; Koelsch, Vuust, & Friston, 2019). Similarly, synchrony is widespread in human sociality (including phenomena such as gaze synchrony, affect synchrony, the chameleon effect, and others), but the isochrony of musical rhythm provides an unusually effective affordance for synchronization. Furthermore, phenomena such as “groove” seem to be mainly evoked by musical stimuli, and therefore are relatively domain-specific. Thus, musicality encompasses multiple mechanisms that vary in their domain-specificity, but combines them into a uniquely effective package.

### 6.2 Group selection

Most previous social bonding theories of music evolution have relied on an evolutionary mechanism incorporating some form of group selection, in which genetic variants are selected for because of their effects on the reproductive success of entire groups (e.g., Brown, 2000a; Wiltermuth & Heath, 2009). Group selection has been largely dismissed for decades (Williams, 1966), and while it is re-emerging in the form of multi-level selection (Traulsen & Nowak, 2006; Wilson & Wilson, 2007) and cultural group selection (Richerson et al., 2016), it remains controversial (Pinker, 2012; see also commentary accompanying Richerson et al., 2016).

The MSB hypothesis does NOT require group selection (any more than grooming, play, or laughter do): fitness advantages accrue to individuals who are able to bond more effectively with larger numbers of individuals. Although there are often advantages to well-bonded groups for various activities (e.g.,

group hunting or foraging, jointly repelling enemies), even for such activities the key fitness advantages accrue to individuals.

### 6.3 Gene–culture coevolution and causality

Some evolutionary psychologists have been critical of social bonding theories of music evolution because they consider them circular arguments that fail to explain the ultimate causal mechanism by which music could have evolved as a biological adaptation:

Perhaps singing lullabies soothes babies; perhaps dancing relieves tension; perhaps shared stories bond the community. The question is, why would anyone have predicted, *a priori*, that people would be constituted in such a way that these things would happen? (Pinker, 2007, pp. 170–171)

Several have posited an adaptive function for music in enhancing “cohesion” or “bonding”.... But this reasoning is circular: it takes as a given the fact that music performance and listening produces reliable effects... and then argues that one or more parts of the music faculty evolved in order to produce these effects. But why should music produce these effects and not others? ... accounts invoking cohesion and/or bonding as an adaptive target provide neither a specific account of the ultimate functional mechanism by which music should increase cohesion, nor an account of how that cohesion would produce fitness advantages. And if cohesion is indeed fitness enhancing, why should individuals wait for music-making to produce that cohesion? Why not just be cohesive without music? (Mehr & Krasnow, 2017, p. 676)

music does not *directly* cause social cohesion: rather, it signals existing social cohesion that was obtained by other means (Mehr et al., target article, sect. 4.2.1, para. 14 [emphasis in original]; paraphrasing Hagen & Bryant, 2003, p. 30)

Our preceding account provides *a priori* arguments detailing why and how specific design features of human musicality have social bonding effects, the mechanisms underlying these effects, and how and why these may have evolved. In particular, we provided specific reasons that behaviors with the design features of music would have social bonding effects: because such behaviors allow people to predict, synchronize, share goals, distinguish individual contributions, experience shared positive emotions, and make social decisions more than other human behaviors (ABMs or language). This explains why music should produce “[social bonding] effects and not others”: behaviors that allow us to align in time and frequency, coordinate behaviors in large groups while distinguishing individual contributions, share emotions and goals, and choose appropriate social partners have tangible and predictable social bonding effects. Music is a particularly effective cognitive “technology” (Patel, 2008, 2018) that fulfills these design criteria, making musicality an effective toolkit for social bonding functions, shaped by both biological and cultural evolution.

Our hypothesis differs from most traditional social bonding theories because we do not argue that musicality necessarily originated as a biological adaptation. Instead, components of musicality may have arisen initially as cultural inventions and/or byproducts of other adaptations, later exapted and modified through gene–culture coevolution for their social bonding functions in a musical context (e.g., beat synchronization initially as a byproduct of the evolution of vocal learning, as argued by Patel, Iversen, Bregman, & Schulz [2009] and Schachner et al. [2009]; although cf. Merker et al. [2018] for an alternative interpretation). The initial social cohesion functions may not have begun as genetic adaptations. In this sense, we largely agree with Mehr et al., who write:

We also agree with the proponents of the social bonding hypothesis that musical abilities evolved because musical performances played an important role in cooperative sociality. But given the issues described above, we find it more likely that music evolved to credibly signal decisions to cooperate that were already reached by other means, not to determine them. (Mehr et al., target article, sect. 4.2, para. 2)

But in a social environment in which social bonding already enhanced individual reproductive fitness, the subsequent cultural evolution of musical behaviors would lead to biological selection on musicality (e.g., to promote motivation to engage in/attend to musical behaviors), because of the adaptive consequences of musicality for social bonding. In this way, just as social bonding is crucial in most primates, generating selection on the mechanisms that achieve it, social bonding functions of “proto-musical” mechanisms may have played important roles in hominin evolution long before today’s full-blown musicality evolved.

We emphasize that past adaptive function, although important, should not be the sole criterion by which to judge theories of the evolution of musicality. As previously argued at length (e.g., Fitch, 2006, 2015b; Honing et al., 2015), Tinbergen’s (1963) multi-factorial perspective, which seeks understanding of traits at the four interlinked explanatory levels of mechanism, ontogeny, phylogeny, and adaptive function, is a fruitful method for understanding the evolution of musicality. We may never know with certainty the precise ancestral adaptive conditions or specific genetic mutations involved in the evolution of musicality. Even so, the comparative method provides a key tool for empirically testing evolutionary hypotheses (Fitch, 2015b). Section 5 lists a variety of testable empirical predictions of the MSB hypothesis.

### 6.4. Parochial altruism and out-group exclusion

Enhanced social bonding between some individuals inevitably means a relative decrease between others. In-group social bonding has a dark side of increasing hostility toward out-groups (Gelfand, Caluori, Jackson, & Taylor, 2020; Whitehouse, 2018), as exemplified in the use of music in warfare by the Nazis and other groups throughout history (Brown & Volgsten, 2006). The traditional Māori haka “Ka Mate” is famously used by New Zealand’s national rugby team to simultaneously bind team-mates together and intimidate the opposing team through coordinated dancing and vocalization (Jackson & Hokowhitu, 2002). The ability of music to exclude out-group members might appear to be an argument against its function in bonding in-group members, but out-group exclusion is entirely consistent with the social bonding hypothesis. Because the creation or strengthening of a social bond between some (participating) individuals by definition excludes others, the observation that particular forms of music can cause emotional dissonance or fear in others is compatible with a social bonding function.

Earlier expositions of the social bonding hypothesis (Brown, 2000a; Freeman, 2000) noted that “bonding is always exclusionary” and “individuals who do not ‘belong’ become enemies ... The process is similar to sexual jealousy, which manifests the exclusionary nature of the pair bond” (Freeman, 2000, pp. 421–422). This observation is mirrored in the recent literature on oxytocin which, far from being an indiscriminate “love drug,” simultaneously exerts affiliative effects among in-group members and exclusionary effects toward out-group individuals (cf. Beery, 2015; Shamay-Tsoory & Abu-Akel, 2016). The use of music to exclude others is no argument against its social bonding origins.



### 6.5. Solo music, sexual selection, and individual signaling

Although coordinated group performances predominate throughout the world, various widespread musical genres are not necessarily performed in coordinated groups. In particular, lullabies and love songs are found throughout the world and are often performed by a lone singer (Mehr & Singh et al., 2018; Trehub et al., 1993). This is perfectly consistent with the MSB hypothesis, as lullabies and love songs are often dyadic: sung by a soloist to bond with another person (by soothing an infant or wooing a potential mate).

More generally, some may wonder why, if social bonding is so important to the evolution of musicality, do people enjoy playing or listening to music alone? We emphasize that even solo music listening can support social bonding goals (Trehub et al., 2018). A young adult meeting a new person in an online chat discusses music preferences more often than other topics, and based on music preferences alone, people draw social inferences about others (Rentfrow & Gosling, 2006). Thus, music preferences developed during solo listening can be used as social cues, displayed and evaluated when establishing new social bonds.

Solo listening may serve other, non-social functions (e.g., mood regulation, staying awake while driving; DeNora, 2000; North, Hargreaves, & Hargreaves, 2004; Sloboda, O'Neill, & Ivaldi, 2001). We do not argue that social bonding is the only possible function of music. By analogy, language's primary function may be to communicate information between people, but it is also useful in private thought, or to allow one to preserve thoughts for the future (particularly after the invention of writing). Similarly, the same auditory-motor-reward connections that make music so socially powerful also allow people to enjoy playing or listening to music alone. Often, solo music was experienced previously in a social context, which is re-evoked by solo listening/playing.

Related to the idea of virtuosic solo music-making is the distinction between social bonding and theories such as sexual selection or honest signaling that emphasize music as a signal of individual fitness. The MSB hypothesis does not reject such theories. Instead, it emphasizes that individual signaling theories are insufficient to explain all of the broader social functions of music, whereas social bonding provides more explanatory power (although we concede that the MSB hypothesis cannot explain all possible functions of music; Oesch, 2019). For example, in contemporary Western night clubs and traditional non-Western societies, all-night music and dance rituals function both to bond participants and as opportunities to find potential mates (Merriam, 1964; Thornton, 1995). In such contexts, dancing, singing, and/or playing instruments can function to bond with same and opposite-sex partners and to advertise evolutionary fitness to potential mates. Bonding and signaling hypotheses are not mutually exclusive, but rather complementary.

The complementarity of the MSB and alternative hypotheses makes it challenging to falsify the MSB hypothesis. However, we have provided a number of specific predictions, each of which is potentially falsifiable and would count as evidence against the MSB hypothesis, particularly if alternative hypotheses better predict the data. For example, our hypothesis and Hagen and Bryant's (2003) coalitional signaling hypothesis make predictions regarding synchrony: we argue that synchrony should enhance social bonding, whereas Hagen and Bryant argue that synchrony should enhance perceived coalitional quality. To differentiate between these and other competing hypotheses, our predictions regarding the effects of synchrony (or other aspects of

musicality) on social bonding could be compared directly against perceived coalition quality or other competing predictions (e.g., attractiveness; Miller, 2000, parental investment; Mehr & Krasnow, 2017; Mehr et al., target article) in future research. If synchrony increases perceived bonding relative to perceived coalition quality, attractiveness, or parental investment, it would constitute evidence favoring the MSB hypothesis over competing alternatives. Another example of predictions that differentiate among alternative hypotheses is the MSB prediction that social bonding functions will be common cross-culturally but the relative frequencies of specific genres and sub-functions (e.g., lullabies vs. love songs vs. group dancing) will vary across societies. In contrast, theories that focus on infant-directed song or sexual selection predict instead that these categories should be more common and consistent cross-culturally than the other categories of social bonding. Furthermore, phylogenetic or other cross-species analyses (e.g., Hoeschele et al., 2018; Schruth, Templeton, & Holman, *In press*; Shultz, Opie, & Atkinson, 2011) could allow us to quantify the relative effects of group size, sexual competition, parental investment strategies, or other factors on the evolution of vocal learning, beat perception, or other aspects of musicality. We encourage tests of MSB predictions against those of competing hypotheses.

## 7. Conclusion

Social bonding has long been acknowledged as an important function of contemporary music, but its causal role in the evolution of music has often been dismissed as a naïve application of group selection theory. Recent advances in gene-culture coevolution theory allow us to provide a more nuanced model of music evolution that does not rely on group selection. Our argument has focused on social bonding as the primary factor shaping the evolution of human musicality. This MSB hypothesis provides a framework for understanding the past evolution of musicality, and a starting point for the future cultural evolution of new forms of music that harness the social power of music to bring people together. Music may not be a "universal language" (Longfellow, 1835; Savage, 2019b), but music's universal power to bring people together across barriers of language, age, gender, and culture sheds light on its biological and cultural origins, and provides humanity with a set of tools to create a more harmonious future – both literally and figuratively.

**Acknowledgments.** This paper originated in an invited symposium entitled "The origins of music in human society" held in December 2017 at the Abbaye de Royaumont, France, organized by Paul Seabright, Francis Maréchal, Marie-Hélène Cassar, and Juliette Lobry. We gratefully acknowledge funding for the workshop from the French Agence Nationale de la Recherche (under the Investissement d'Avenir program, ANR-17-EURE-0010). We thank the organizers, sponsors, and other invited participants (in addition to the seven of us, these were: Simha Arom, Kofi Asante, Jean-Julien Aucouturier, Greg Bryant, François Cam, Didier Demolin, Edward Hagen, Nori Jacoby, Martin Lang, Jacques Launay, Samuel Mehr, Marcel Peres, Benjamin Purzycki, Lauriane Rat-Fischer, Nick Rawlins, Sandra Trehub, Connor Wood, and the Ensemble Organum) for this stimulating interdisciplinary workshop and discussion leading to this article and a separate article by Mehr, Bryant, Hagen, and Max Krasnow. We thank Greg Bryant, Shinya Fujii, Eva Jablonka, Nori Jacoby, Kimberly Kobayashi-Johnson, Samuel Mehr, Bjorn Merker, Aniruddh Patel, Peter Pfordresher, Manvir Singh, Dor Shilton, Parker Tichko, Sandra Trehub, the students of Keio University's CompMusic and NeuroMusic Labs, and of Oxford University's Institute of Cognitive and Evolutionary Anthropology, and four anonymous reviewers for comments on earlier versions of this manuscript.



**Author contributions.** PES proposed the article concept and author list with feedback from WTF and LG. PES, WTF, and PL drafted the abstract and outline, with feedback from LG, AS, and BT. WTF, PES, and PL drafted sects. 1 and 2; PL and BT drafted sect. 4; PES, WTF, and PL drafted sects. 5–7. All authors drafted sub-sects. relevant to their own expertise, which were then edited and synthesized into sects. 3 and 4 by PES, PL, BT, WTF, and AS. All authors edited and approved the final manuscript.

**Financial support.** PES was supported by Grant-in-Aid no. 19KK0064 from the Japan Society for the Promotion of Science and startup grants from Keio University (Keio Global Research Institute, Keio Research Institute at SFC, and Keio Gijuku Academic Development Fund). PL was supported by the National Science Foundation NSF-STTR no. 1720698 NSF-CAREER #1945436, NSF-STTR #2014870, the Grammy Foundation and startup funds from Northeastern University. BT was supported by funding from the French Agence Nationale de la Recherche (under the Investissement d'Avenir program, ANR-17-EURE-0010) while on a Visiting Fellowship at the Institute of Advanced Study Toulouse. AS was supported by the National Science Foundation under NSF-BCS no. 1749551. WTF was supported by Austrian Science Fund (FWF) DK Grant “Cognition & Communication” (W1262-B29).

**Conflict of interest.** None.

## Notes

1. For discussion of practical and ethical challenges involved in defining and comparing “music” and “musicality” in cross-culturally valid ways, see Brown (2020), Ewell (2020), Iyer and Born (2020), Jacoby et al. (2020), Nettl (2015), and Savage (2019b).
2. The symposium program is available at [https://www.iast.fr/sites/default/files/IAST/conf/royaumont/royaumont\\_program051217.pdf](https://www.iast.fr/sites/default/files/IAST/conf/royaumont/royaumont_program051217.pdf).
3. Endogenous opioids (e.g., beta-endorphins) likely interact with other systems in analgesic effects (e.g., Welch & Eads, 1999). Nevertheless, elevated pain threshold is a common proxy-measure of elevated beta-endorphin levels, because of the EOS's role in our pain-pleasure circuitry (Mueller et al., 2010), and the fact that direct measures are invasive and expensive.
4. However, other factors (e.g., ceiling effects, optimal degrees of complexity, and rhythm-melody interactions [Prince, Thompson, & Schmuckler, 2009]) may limit social bonding effects, leading to non-linear interactions when combining multiple musical components.
5. We have refrained from making detailed predictions about genetic bases of musicality because our current state of knowledge is limited (Gingras et al., 2018). However, we hope that new findings from initiatives such as the 2019 symposium on “Deciphering the biology of human musicality through state-of-the-art genomics” (<http://www.mcqg.uva.nl/musicality2019>) will enable researchers to investigate genetic dimensions of the gene-culture coevolutionary mechanisms we describe.
6. Others view language evolution as being driven largely by cultural evolution (e.g., Jablonka, Ginsburg, & Dor, 2012; Kirby, 2017).

## References

- Alaux, C., Maisonnasse, A., & Le Conte, Y. (2010). Pheromones in a superorganism: From gene to social regulation. *Vitamins & Hormones*, 83, 401–423.
- Anderson, B. R. O. (1991). *Imagined communities: Reflections on the origin and spread of nationalism*. Verso.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105–167.
- Arom, S. (1991). *African polyphony and polyrhythm: Musical structure and methodology*. Cambridge University Press.
- Atwood, S., Mehr, S. A., & Schachner, A. (2020). Expectancy effects threaten the inferential validity of synchrony-prosociality research. *PsyArXiv*. <https://doi.org/10.31234/osf.io/zjy8u>.
- Atzil, S., Gao, W., Fradkin, I., & Barrett, L. F. (2018). Growing a social brain. *Nature Human Behaviour*, 2(9), 624–636.
- Atzil, S., Touroutoglou, A., Rudy, T., Salcedo, S., Feldman, R., Hooker, J. M., ... Barrett, L. F. (2017). Dopamine in the medial amygdala network mediates human bonding. *Proceedings of the National Academy of Sciences of the USA*, 114(9), 2361–2366.
- Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30(354), 441–451.
- Balsby, T. J. S., & Scarl, J. C. (2008). Sex-specific responses to vocal convergence and divergence of contact calls in orange-fronted conures (*Aratinga canicularis*). *Proceedings of the Royal Society B: Biological Sciences*, 275, 2147–2154.
- Banfield, S. (Ed.). (2009). *The sounds of Stonehenge*. Archaeopress.
- Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *NeuroImage*, 21(3), 1155–1166.
- Bateson, P. (2004). The active role of behaviour in evolution. *Biology and Philosophy*, 19, 283–298.
- Beadle, J. N., Paradiso, S., & Tranel, D. (2018). Ventromedial prefrontal cortex is critical for helping others who are suffering. *Frontiers in Neurology* 9: article 288. <https://doi.org/10.3389/fneur.2018.00288>.
- Beery, A. K. (2015). Antisocial oxytocin: Complex effects on social behavior. *Current Opinion in Behavioral Sciences*, 6, 174–182. <https://doi.org/10.1016/j.cobeha.2015.11.006>.
- Belfi, A. M., & Loui, P. (2020). Musical anhedonia and rewards of music listening: Current advances and a proposed model. *Annals of the New York Academy of Sciences*, 1464, 99–114.
- Benzon, W. (2001). *Beethoven's anvil, music in mind and culture*. Basic Books.
- Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Current Biology*, 23, 268–269.
- Bergman, T. J., & Beehner, J. C. (2015). Measuring social complexity. *Animal Behaviour*, 103, 203–209.
- Bernatzky, G., Presch, M., Anderson, M., & Panksepp, J. (2011). Emotional foundations of music as a non-pharmacological pain management tool in modern medicine. *Neuroscience & Biobehavioral Reviews*, 35(9), 1989–1999.
- Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, 68(3), 530–543.
- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the USA*, 98(20), 11818–11823.
- Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, 2(4), 382–387.
- Bowling, D. L., Hoeschele, M., Gill, K. Z., & Fitch, W. T. (2017). The nature and nurture of musical consonance. *Music Perception*, 35(1), 118–121.
- Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences of the USA*, 112(36), 11155–11160.
- Bowling, D. L., Purves, D., & Gill, K. Z. (2018). Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences of the USA*, 115(1), 216–221.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.
- Bradbury, J. (2001). Vocal communication of wild parrots. *Journal of the Acoustical Society of America*, 115, 2373.
- Brown, D. (2020). An open letter on racism in music studies: Especially ethnomusicology and music education. My People Tell Stories. <https://www.mypeopletellstories.com/blog/open-letter>.
- Brown, D. E. (1991). *Human universals*. Temple University Press.
- Brown, S. (2000a). Evolutionary models of music: From sexual selection to group selection. In F. Tonneau & N. S. Thompson (Eds.), *Perspectives in ethology*, vol. 13: *Evolution, culture, and behavior* (pp. 231–281). Plenum Publishers.
- Brown, S. (2000b). The “musilanguage” model of musical evolution. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 271–300). MIT Press.
- Brown, S. (2007). Contagious heterophony: A new theory about the origins of music. *Musicae Scientiae*, 11(1), 3–26.
- Brown, S. (2017). A joint prosodic origin of language and music. *Frontiers in Psychology*, 8(1894), 1–20. <https://doi.org/10.3389/fpsyg.2017.01894>.
- Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, 41(2), 229–248.
- Brown, S., & Volgsten, U. (2006). *Music and manipulation: On the social uses and social control of music*. Berghahn Books.
- Cameron, D. J., Pickett, K. A., Earhart, G. M., & Grahn, J. A. (2016). The effect of dopaminergic medication on beat-based auditory timing in Parkinson's disease. *Frontiers in Neurology*, 7(19), 1–8. <https://doi.org/10.3389/fneur.2016.00019>.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton University Press.
- Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences*, 17(4), 179–193.
- Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., & Koelsch, S. (2019). Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity. *Current Biology*, 29(23), 4084–4092.
- Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, 17(6), 1003–1011.
- 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*, 32(19), 1213–1220.

- Cirelli, L. K., & Trehub, S. E. (2018). Infants help singers of familiar songs. *Music & Science*, 1, 205920431876162. <https://doi.org/10.1177/2059204318761622>.
- Cirelli, L. K., & Trehub, S. E. (2020). Familiar songs reduce infant distress. *Developmental Psychology*, 56(5), 861–868.
- Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences*, 1423(1), 66–72.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–253.
- Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, 460(7256), 737–740.
- Corbeil, M., Trehub, S. E., & Peretz, I. (2016). Singing delays the onset of infant distress. *Infancy*, 21(3), 373–391.
- Cross, E. S., Kraemer, D. J., Hamilton, A. F. D. C., Kelley, W. M., & Grafton, S. T. (2008). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex*, 19(2), 315–326.
- Cross, I., & Morley, I. (2009). The evolution of music: Theories, definitions, and the nature of the evidence. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 61–82). Oxford University Press.
- Darwin, C. (1871). *The descent of man, and selection in relation to sex*. John Murray.
- Davila Ross, M., Owren, M. J., & Zimmerman, E. (2009). Reconstructing the evolution of laughter in great apes and humans. *Current Biology*, 19, 1106–1111.
- Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, 23(2), 744–751.
- Declerck, C. H., Boone, C., Pauwels, L., Vogt, B., & Fehr, E. (2020). A registered replication study on oxytocin and trust. *Nature Human Behaviour*, 4, 646–55. <https://doi.org/10.1038/s41562-020-0878-x>.
- De Dreu, C. K. W., & Kret, M. E. (2016). Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense. *Biological Psychiatry*, 79(3), 165–173.
- Dengler-Crish, C. M., & Catania, K. C. (2007). Phenotypic plasticity in female naked mole-rats after removal from reproductive suppression. *Journal of Experimental Biology*, 210, 4351–4358.
- DeNora, T. (2000). *Music in everyday life*. Cambridge University Press.
- Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *Journal of the Acoustical Society of America*, 129, 2245–2252.
- Dissanayake, E. (2000). Antecedents of the temporal arts in early mother–infant interaction. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 389–410). MIT Press.
- Dissanayake, E. (2009). Root, leaf, blossom, or bole: Concerning the origin and adaptive function of music. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 17–30). Oxford University Press.
- Dölen, G., Darvishzadeh, A., Huang, K. W., & Malenka, R. C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. *Nature*, 501(7466), 179–184.
- Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves “mind-reading” in humans. *Biological Psychiatry*, 61(6), 731–733.
- Dunbar, R. I. M. (1991). Functional significance of social grooming in primates. *Folia Primatologica*, 57, 121–131.
- Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16, 681–735.
- Dunbar, R. I. M. (2012a). On the evolutionary function of song and dance. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 201–214). Oxford University Press.
- Dunbar, R. I. M. (2012b). Bridging the bonding gap: The transition from primates to humans. *Philosophical Transactions of the Royal Society B: Biological Science*, 367 (1597), 1837–1846.
- Dunbar, R. I. M., & Shultz, S. (2010). Bondedness and sociality. *Behaviour*, 147(7), 775–803.
- Durham, W. H. (1991). *Coevolution: Genes, culture, and human diversity*. University of California Press.
- Erber, R., Wegner, D. M., & Theriault, N. (1996). On being cool and collected: Mood regulation in anticipation of social interaction. *Journal of Personality and Social Psychology*, 70(4), 757–766.
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(5), 429–492.
- Ewell, P. A. (2020). Music theory and the white racial frame. *Music Theory Online*, 26(2), 1–29. <https://doi.org/10.30535/mt0.26.2.4>.
- Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, 27(4), 491–541.
- Farabaugh, S. M. (1982). The ecological and social significance of duetting. In D. S. Kroodsma & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol. 2, pp. 85–124). Academic Press.
- Fazenda, B., Scarre, C., Till, R., Pasalodos, R. J., Guerra, M. R., Tejedor, C., ... Foulds, F. (2017). Cave acoustics in prehistory: Exploring the association of Palaeolithic visual motifs and acoustic response. *The Journal of the Acoustic Society of America*, 142, 1332–1349.
- Feeles, F. (1982). Song mimesis within colonies of *Cacicus C. cela* (Icteridae: Aves): A colonial password? *Zeitschrift Tierpsychologie*, 58, 119–152.
- Finlay, B. L., Darlington, R. B., & Nicastro, N. (2001). Developmental structure in brain evolution. *Behavioral and Brain Sciences*, 24, 263–308.
- Fitch, W. T. (2004). Evolving honest communication systems: Kin selection and “mother tongues.” In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 275–296). MIT Press.
- Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100(1), 173–215.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge University Press.
- Fitch, W. T. (2015a). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1664), e2014.0091. <http://doi.org/10.1098/rstb.2014.0091>.
- Fitch, W. T. (2015b). The biology and evolution of musical rhythm: An update. In I. Toivonen, P. Csúri & E. van der Zee (Eds.), *Structures in the mind: Essays on language, music, and cognition in honor of Ray Jackendoff* (pp. 293–324). MIT Press.
- Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*, 24, 3–33.
- Fitch, W. T., & Jarvis, E. D. (2013). Birdsong and other animal models for human speech, song, and vocal learning. In M. A. Arbib (Ed.), *Language, music, and the brain: A mysterious relationship* (pp. 499–539). MIT Press.
- Fitch, W. T., von Graevenitz, A., & Nicolas, E. (2009). Bio-aesthetics and the aesthetic trajectory: A dynamic cognitive and cultural perspective. In M. Skov & O. Vartanian (Eds.), *Neuroaesthetics* (pp. 59–102). Baywood.
- Fletcher, P. T., Whitaker, R. T., Tao, R., DuBray, M. B., Froehlich, A., Ravichandran, C., ... Lainhart, J. E. (2010). Microstructural connectivity of the arcuate fasciculus in adolescents with high-functioning autism. *NeuroImage*, 51(3), 1117–1125.
- Freeman, W. (2000). A neurobiological role for music in social bonding. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 411–424). MIT Press.
- Frigyesi, J. (1993). Preliminary thoughts toward the study of music without clear beat: The example of “Flowing Rhythm” in Jewish “Nusah.” *Asian Music*, 24, 59–88.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127–137.
- Fujii, T., Schug, J., Nishina, K., Takahashi, T., & Okada, H. (2016). Relationship between salivary oxytocin levels and generosity in preschoolers. *Scientific Reports*, 6 (38662), 1–7.
- Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, 35(45), 15187–15198.
- Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., & Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science* (New York, N.Y.), 354(6317), 1278–1282.
- Gamble, C. (2010). Technologies of separation and the evolution of social extension. In R. Dunbar, C. Gamble & J. Gowlett (Eds.), *Social brain, distributed mind* (pp. 17–42). Oxford University Press.
- Gebauer, L., Witek, M. A. G., Hansen, N. C., Thomas, J., Konvalinka, I., & Vuust, P. (2016). Oxytocin improves synchronisation in leader-follower interaction. *Scientific Reports*, 6, 38416. <https://doi.org/10.1038/srep38416>.
- Geissmann, T. (1999). Duet songs of the siamang, *Hylobates syndactylus*: II. Testing the pair-bonding hypothesis during a partner exchange. *Behaviour*, 136(8), 1005–1039.
- Gelfand, M. J., Caluori, N., Jackson, J. C., & Taylor, M. K. (2020). The cultural evolutionary trade-off of ritualistic synchrony. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375, 20190432. <https://doi.org/10.1098/rstb.2019.0432>.
- Ghazban, N. (2013). *Emotion regulation in infants using maternal singing and speech*. PhD dissertation, Ryerson University, Toronto.
- Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PLoS ONE*, 4 (12), e8144. <https://doi.org/10.1371/journal.pone.0008144>.
- Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2018). Defining the biological bases of individual differences in musicality. In H. Honing (Ed.), *The origins of musicality* (pp. 221–250). MIT Press.
- Gioia, T. (2019). *Music: A subversive history*. Basic Books.
- Gold, B. P., Mas-Herrero, E., Zeighami, Y., Benovoy, M., Dagher, A., & Zatorre, R. J. (2019). Musical reward prediction errors engage the nucleus accumbens and motivate learning. *Proceedings of the National Academy of Sciences of the USA*, 116 (8), 3310–3315.
- Good, A., & Russo, F. A. (2016). Singing promotes cooperation in a diverse group of children. *Social Psychology*, 47(6), 340–344.
- Grahn, J., Bauer, A.-K. R., & Zamm, A. (2020). Music-making brings us together during the coronavirus pandemic. *The Conversation*. <https://theconversation.com/music-making-brings-us-together-during-the-coronavirus-pandemic-137147>.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19(5), 893–906.
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, 29(23), 7540–7548.

- Grape, C., Sandgren, M., Hansson, L.-O., Ericson, M., & Theorell, T. (2003). Does singing promote well-being?: An empirical study of professional and amateur singers during a singing lesson. *Integrative Physiological & Behavioral Science*, 38(1), 65–74.
- Grawunder, S., Crockford, C., Clay, Z., Kalan, A. K., Stevens, J. M. G., Stoessel, A., & Hohmann, G. (2018). Higher fundamental frequency in bonobos is explained by larynx morphology. *Current Biology*, 28(20), R1188–R1189.
- Greenwood, D. N., & Long, C. R. (2009). Mood specific media use and emotion regulation: Patterns and individual differences. *Personality and Individual Differences*, 46(5–6), 616–621.
- Griffiths, P. E. (2003). Beyond the Baldwin effect: James Mark Baldwin's "social heredity", epigenetic inheritance, and niche construction. In B. Weber & D. J. Depew (Eds.), *Learning, meaning and emergence: Possible Baldwinian mechanisms in the co-evolution of mind and language* (pp. 193–215). MIT Press.
- Gustison, M. L., Aliza, L. R., & Bergman, T. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1847–1859.
- Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, 14(1), 21–51.
- Haimoff, E. H. (1986). Convergence in the duetting of monogamous Old World primates. *Journal of Human Evolution*, 15, 51–59.
- Hall, M. L. (2004). A review of hypotheses for the functions of avian duetting. *Behavioral Ecology and Sociobiology*, 55, 415–430.
- Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, 2, article 156. <https://doi.org/10.3389/fpsyg.2011.00156>.
- Hanslick, E. (1858). *Vom Musikalisch-Schönen*. Rudolf Weigel.
- Harrison, P. M. C., & Pearce, M. T. (2020). Simultaneous consonance in music perception and composition. *Psychological Review*, 127(2), 216–244.
- Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., & Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28, 795–855.
- Hoeschele, M., Merchant, H., Kikuchi, Y., Hattori, Y., & ten Cate, C. (2018). Searching for the origins of musicality across species. In H. Honing (Ed.), *The origins of musicality* (pp. 149–170). MIT Press.
- Honing, H. (Ed.). (2018). *The origins of musicality*. MIT Press.
- Honing, H., Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20140088. <https://doi.org/10.1098/rstb.2014.0088>.
- Horowitz, J. (2020). Italians find "a moment of joy in this moment of anxiety." *New York Times*. <https://www.nytimes.com/2020/03/14/world/europe/italians-find-a-moment-of-joy-in-this-moment-of-anxiety.html>.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27(6), 949–960.
- Hrdy, S. B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. Harvard University Press.
- Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., ... Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience*, 30, 4999–5007.
- Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, 930, 43–61.
- Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.
- Inagaki, T. K. (2018). Opioids and social connection. *Current Directions in Psychological Science*, 27, 85–90.
- Inagaki, T. K., Ray, L. A., Irwin, M. R., Way, B. M., & Eisenberger, N. I. (2016). Opioids and social bonding: Naltrexone reduces feelings of social connection. *Social Cognitive and Affective Neuroscience*, 11, 728–735.
- Iyer, V., & Born, G. (2020). Of musicalities and musical experience: Vijay Iyer and Georgina Born in conversation. Wigmor Hall Podcasts. <https://wigmor-hall.org.uk/podcasts/of-musicalities-and-musical-experience-vijay-iyer-and-georgina-born-in-conversation>.
- Jablonka, E., Ginsburg, S., & Dor, D. (2012). The co-evolution of language and emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599), 2152–2159.
- Jablonka, E., & Lamb, M. J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT Press.
- Jackson, S. J., & Hokowhitu, B. (2002). Sport, tribes, and technology: The New Zealand All Blacks haka and the politics of identity. *Journal of Sport & Social Issues*, 26(2), 125–139.
- Jacoby, N., Margulis, E., Clayton, M., Hannon, E., Honing, H., Iversen, J., ... Wald-Fuhrmann, M. (2020). Cross-cultural work in music cognition: Methodologies, pitfalls, and practices. *Music Perception*, 37(3), 185–195.
- Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, 27, 359–370.
- Jacoby, N., Polak, R., Grahn, J., Cameron, D., Lee, K. M., Godoy, R., ... McDermott, J. H. (Preprint). Universality and cross-cultural variation in mental representations of music revealed by global comparison of rhythm priors. *PsyArXiv*, <https://doi.org/10.31234/osf.io/b879v>.
- 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, 3229–3243.
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology*, 141(1), 54–75.
- Janik, V. M., & Slater, P. J. B. (1999). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99.
- Järvinen-Pasley, A., Vines, B. W., Hill, K. J., Yam, A., Grichanik, M., Mills, D., & ... Bellugi, U. (2010). Cross-modal influences of affect across social and non-social domains in individuals with Williams syndrome. *Neuropsychologia*, 48(2): 456–466.
- Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science (New York, N.Y.)*, 366, 50–54.
- Jones, M. R. (2018). *Time will tell: A theory of dynamic attending*. Oxford University Press.
- Josef, L., Goldstein, P., Maysless, N., Ayalon, L., & Shamay-tsoory, S. G. (2019). The oxytocinergic system mediates synchronized interpersonal movement during dance. *Scientific Reports*, 9(1894), 1–8.
- Keeler, J. R., Roth, E. A., Neuser, B. L., Spitsbergen, J. M., Waters, D. J. M., & Vianney, J.-M. (2015). The neurochemistry and social flow of singing: Bonding and oxytocin. *Frontiers in Human Neuroscience*, 9, article 518. <https://doi.org/10.3389/fnhum.2015.00518>.
- Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology*, 14, 155–161.
- Kim, M., & Schachner, A. (2020). *The origins of dance: Characterizing the development of infants' earliest dance behavior*. Poster presented at the Brain Cognition Emotion Music Conference. <https://osf.io/meetings/BCEM/>.
- Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin & Review*, 24, 118–137.
- Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, 31(5), 354–364.
- Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). fMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage*, 12(1), 20–27.
- Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, 23(1), 63–77.
- Kokal, I., Engel, A., Kirschner, S., & Keyzers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment – if the rhythm comes easily. *PLoS ONE*, 6(11), e27272. <https://doi.org/10.1371/journal.pone.0027272>.
- Kornhaber, S. (2020). The coronavirus has forced a repurposing of music. *The Atlantic*. <https://www.theatlantic.com/culture/archive/2020/04/coronavirus-has-forced-repurposing-music/609601/>.
- Kosfeld, M., Heinrichs, M., Zak, P., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042), 673–676.
- Kroodmas, D. E. (1978). Continuity and versatility in bird song: Support for the monotony-threshold hypothesis. *Nature*, 274(5672), 681–683.
- Kuroyanagi, J., Sato, S., Ho, M.-J., Chiba, G., Six, J., Pföndresher, P., ... Savage, P. E. (2019). Automatic comparison of human music, speech, and bird song suggests uniqueness of human scales. Proceedings of the 9th International Workshop on Folk Music Analysis (FMA2019), pp. 35–40. <https://doi.org/10.31234/osf.io/zpv5w>.
- Lagacé, R. O. (1979). The HRAF probability sample: Retrospect and prospect. *Cross-Cultural Research*, 14(3), 211–229.
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320(5872), 110–113.
- Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Social Cognition*, 29(1), 1–14.
- Laland, K. N., & Brown, G. R. (2011). *Sense and nonsense: Evolutionary perspectives on human behaviour* (2nd ed.). Oxford University Press.
- Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2000). Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences*, 23, 131–175.
- Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: Bringing genetics and the human sciences together. *Nature Reviews Genetics*, 11, 137–148.
- Laland, K., Wilkins, C., & Clayton, N. (2016). The evolution of dance. *Current Biology*, 26(1), R5–R9.



- Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology*, 127, 191–197.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159.
- Large, E. W., Kim, J. C., Flaig, N. K., Bharucha, J., & Krumhansl, C. L. (2016). A neurodynamic account of musical tonality. *Music Perception*, 33(3), 319–331.
- Launay, J., Dean, R. T., & Bailes, F. (2013). Synchronization can influence trust following virtual interaction. *Experimental Psychology*, 60(1), 53–63.
- Launay, J., Tarr, B., & Dunbar, R. I. M. (2016). Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology*, 122, 779–789.
- Le Merrer, J., Becker, J. A., Befort, K., & Kieffer, B. L. (2009). Reward processing by the opioid system in the brain. *Physiological Reviews*, 89(4), 1379–1412.
- Lense, M. D., Gordon, R. L., Key, A. P. F., & Dykens, E. M. (2014). Neural correlates of cross-modal affective priming by music in Williams syndrome. *Social Cognitive and Affective Neuroscience*, 9(4), 529–537.
- Lepage, C., Drolet, P., Girard, M., Grenier, Y., & DeGagné, R. (2001). Music decreases sedative requirements during spinal anesthesia. *Anesthesia & Analgesia*, 93(4), 912–916.
- Loersch, C., & Arbuckle, N. L. (2013). Unraveling the mystery of music: Music as an evolved group process. *Journal of Personality and Social Psychology*, 105(5), 777–798.
- Lomax, A. (Ed.). (1968). *Folk song style and culture*. American Association for the Advancement of Science.
- London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.
- Longfellow, H. W. (1835). *A pilgrimage beyond the sea*. Harper & Bros.
- Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *The Journal of Neuroscience*, 29(33), 10215–10220.
- Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage*, 55(2), 500–507.
- Loui, P., Patterson, S., Sachs, M. E., Leung, Y., Zeng, T., & Przyssinda, E. (2017). White matter correlates of musical anhedonia: Implications for evolution of music. *Frontiers in Psychology*, 8, article 1664. <http://doi.org/10.3389/fpsyg.2017.01664>.
- Love, T. M. (2014). Oxytocin, motivation and the role of dopamine. *Pharmacology Biochemistry and Behavior*, 119, 49–60.
- Maestripietri, D. (2010). Neurobiology of social behavior. In M. Platt & A. Ghazanfar (Eds.), *Primate neuroethology* (pp. 359–384). Oxford University Press.
- Mann, N. I., Dingess, K. A., Barker, K., Graves, J. A., & Slater, P. J. B. (2009). A comparative study of song form and duetting in neotropical *Thryothorus* wrens. *Behaviour*, 146, 1–43.
- Mann, N. I., Dingess, K. A., & Slater, P. J. B. (2006). Antiphonal four-part synchronized chorusing in a neotropical wren. *Biology Letters*, 2, 1–4.
- Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford University Press.
- Martínez-Molina, N., Mas-Herrero, E., Rodríguez-Fornells, A., Zatorre, R. J., & Marco-Pallarés, J. (2016). Neural correlates of specific musical anhedonia. *Proceedings of the National Academy of Sciences of the USA*, 113(46), E7337–E7345.
- Mas-Herrero, E., Zatorre, R. J., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia. *Current Biology*, 24(6), 699–704.
- McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, 135(3), 348–367.
- McBride, J. M., & Tlustý, T. (2020). Cross-cultural data shows musical scales evolved to maximise imperfect fifths. *ArXiv Preprint*. <http://arxiv.org/abs/1906.06171>.
- McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, 20(11), 1035–1041.
- McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, 535, 547–550.
- McNeil, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Harvard University Press.
- Mehr, S. A., & Krasnow, M. M. (2017). Parent-offspring conflict and the evolution of infant-directed song. *Evolution and Human Behavior*, 38, 674–684.
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., ... Glowacki, L. (2019). Universality and diversity in human song. *Science (New York, N.Y.)*, 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, 356–368.
- Mehr, S. A., Song, L. A., & Spelke, E. S. (2016). For 5-month-old infants, melodies are social. *Psychological Science*, 27(4), 486–501.
- Mercado, E., Mantell, J. T., & Pfordresher, P. Q. (2014). Imitating sounds: A cognitive approach to understanding vocal imitation. *Comparative Cognition & Behavior Reviews*, 9, 17–74.
- Merchant, H., Grah, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20140093. <http://doi.org/10.1098/rstb.2014.0093>.
- Merker, B. (1999). Synchronous chorusing and the origins of music. *Musicae Scientiae*, 3, 59–73.
- Merker, B. (2000). Synchronous chorusing and human origins. *The origins of music* (pp. 315–328). MIT Press.
- Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae*, 6(1), 3–21.
- Merker, B., Morley, I., & Zuidema, W. (2018). Five fundamental constraints on theories of the origins of music. In H. Honing (Ed.), *The origins of musicality* (pp. 70–104). MIT Press.
- Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45(1), 4–17.
- Merriam, A. P. (1964). *The anthropology of music*. Northwestern University Press.
- Miles, L. K., Nind, L. K., & Macrae, C. N. (2009). The rhythm of rapport: Interpersonal synchrony and social perception. *Journal of Experimental Social Psychology*, 45(3), 585–589.
- Miller, G. F. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 329–360). MIT Press.
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, 17(8), 1306–1315.
- Mithen, S. J. (2005). *The singing Neanderthals: The origins of music, language, mind, and body*. Weldenfeld & Nicholson.
- Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, 72, 13–20.
- Moore, E., Schaefer, R. S., Bastin, M. E., Roberts, N., & Overy, K. (2017). Diffusion tensor MRI tractography reveals increased fractional anisotropy (FA) in arcuate fasciculus following music-cued motor training. *Brain and Cognition*, 116, 40–46.
- Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences of the USA*, 114(42), E8913–E8921.
- Morley, I. (2013). *The prehistory of music: Human evolution, archaeology, and the origins of musicality*. Oxford University Press.
- Mueller, C., Klega, A., Buchholz, H.-G., Rolke, R., Magerl, W., Schirmacher, R., ... Schreckenberger, M. (2010). Basal opioid receptor binding is associated with differences in sensory perception in healthy human subjects: A [18F]diprenorphine PET study. *NeuroImage*, 49(1), 731–737.
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, 27(4), 455–464.
- Nave, G., Camerer, C., & McCullough, M. (2015). Does oxytocin increase trust in humans? A critical review of research. *Perspectives on Psychological Science*, 10(6), 772–789.
- Nettl, B. (2015). *The study of ethnomusicology: Thirty-three discussions* (3rd ed.). University of Illinois Press.
- Nettl, B., Stone, R., Porter, J., & Rice, T. (Eds.). (1998–2002). *The Garland encyclopedia of world music*. Garland Pub.
- Niarchou, M., Sathirapongsasuti, J. F., Jacoby, N., Bell, E., McArthur, E., Straub, P., ... Gordon, R. L. (2019). Unravelling the genetic architecture of rhythm. *bioRxiv preprint*. <https://doi.org/10.1101/836197>.
- Nooshin, L. (2011). Introduction to the special issue: The ethnomusicology of Western art music. *Ethnomusicology Forum*, 20(3), 285–300.
- North, A. C., Hargreaves, D. J., & Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Perception*, 22(1), 41–77.
- Notroff, J., Dietrich, O., & Schmidt, K. (2015). Gathering of the dead? The early Neolithic sanctuaries of Gobekli Tepe, southeastern Turkey. In S. C. Renfrew, M. J. Boyd & I. Morely (Eds.), *Death rituals, social order and the archaeology of immortality in the ancient world* (pp. 65–18). Cambridge University Press.
- Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, 28, 48–53.
- Oesch, N. (2019). Music and language in social interaction: Synchrony, antiphony, and functional origins. *Frontiers in Psychology*, 10, 1514. <https://doi.org/10.3389/fpsyg.2019.01514>.
- Parkinson, C., & Wheatley, T. (2014). Relating anatomical and social connectivity: White matter microstructure predicts emotional empathy. *Cerebral Cortex*, 24(3), 614–625.
- Patel, A. D. (2008). *Music, language and the brain*. Oxford University Press.
- Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was Darwin wrong? *PLoS Biology*, 12(3), e1001821. <http://doi.org/10.1371/journal.pbio.1001821>.
- Patel, A. D. (2018). Music as a transformative technology of the mind: An update. In H. Honing (Ed.), *The origins of musicality* (pp. 113–126). MIT Press.
- Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, 19, 827–830.
- Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science*, 2, 150221. <http://doi.org/10.1098/rsos.150221>.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6(7), 688–691.



- Peretz, I., Vuvan, D. T., Lagrois, M.-É., & Armony, J. (2018). Neural overlap in processing music and speech. In H. Honing (Ed.), *The origins of musicality* (pp. 205–220). MIT Press.
- Pfordresher, P. Q., & Brown, S. (2017). Vocal mistuning reveals the origin of musical scales. *Journal of Cognitive Psychology*, 29(1), 35–52.
- Pinker, S. (1997). *How the mind works*. Norton.
- Pinker, S. (2007). Toward a consilient study of literature. *Philosophy and Literature*, 31(1), 162–178.
- Pinker, S. (2012). The false allure of group selection. *Edge*. Available at [https://www.edge.org/conversation/steven\\_pinker-the-false-allure-of-group-selection](https://www.edge.org/conversation/steven_pinker-the-false-allure-of-group-selection).
- Podlipniak, P. (2017). The role of the Baldwin effect in the evolution of human musicality. *Frontiers in Neuroscience*, 11(542), 1–12. <https://doi.org/10.3389/fnins.2017.00542>.
- Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronization. *Music Perception*, 36(1), 1–23.
- Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance*, 35(5), 1598–1617.
- Rabinowitch, T.-C., Cross, I., & Burnard, P. (2013). Long-term musical group interaction has a positive influence on empathy in children. *Psychology of Music*, 41(4), 484–498.
- Rabinowitch, T. C., & Meltzoff, A. N. (2017). Synchronized movement experience enhances peer cooperation in preschool children. *Journal of Experimental Child Psychology*, 160, 21–32.
- Ragen, B. J., Maninger, N., Mendoza, S. P., Jarcho, M. R., & Bales, K. L. (2013). Presence of a pair-mate regulates the behavioral and physiological effects of opioid manipulation in the monogamous titi monkey (*Callicebus cupreus*). *Psychoneuroendocrinology*, 38(11), 2448–2461.
- Rainio, R., Lahelma, A., Aikas, T., Lassfolk, K., & Okkonen, J. (2018). Acoustic measurements and digital image processing suggest a link between sound rituals and sacred sites in northern Finland. *Journal of Archaeological Method and Theory*, 25, 453–474.
- Ravignani, A., Delgado, T., & Kirby, S. (2017). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, 1, article 7. <http://doi.org/10.1038/s41562-016-0007>.
- Read, D., & Van der Feeuw, S. (2015). The extension of social relations in time and space during the Palaeolithic and beyond. In F. Coward, R. Hosfield, M. Pope & F. Wenban-Smith (Eds.), *Settlement, society and cognition in human evolution* (pp. 31–53). Cambridge University Press.
- Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized prosociality? *Religion, Brain & Behavior*, 4(1), 3–19.
- Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: Synchrony, shared intentionality and cooperation. *PLoS ONE*, 8(8), e71182. <http://doi.org/10.1371/journal.pone.0071182>.
- Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A meta-analysis. *Zeitschrift für Psychologie*, 224(3), 168–189.
- Rentfrow, P. J., & Gosling, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science*, 17(3), 236–242.
- Richerson, P., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., ... Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, 39, e30. <http://doi.org/10.1017/S0140525X1400106X>.
- Richerson, P. J., Boyd, R., & Henrich, J. (2010). Gene-culture coevolution in the age of genomics. *Proceedings of the National Academy of Sciences of the USA*, 107, 8985–8992.
- Richman, B. (1978). The synchronisation of voices by gelada monkeys. *Primates*, 19, 569–581.
- Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates*, 28, 199–223.
- Rock, A. M. L., Trainor, L. J. D., & Addison, T. L. (1999). Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychology*, 35(2), 527–534.
- Roederer, J. G. (1984). The search for a survival value of music. *Music Perception*, 1(3), 350–356.
- Sachs, M. E., Ellis, R. J., Schlaug, G., & Loui, P. (2016). Brain connectivity reflects human aesthetic responses to music. *Social Cognitive and Affective Neuroscience*, 11(6), 884–891.
- Sakata, J. T., & Brainard, M. S. (2008). Online contributions of auditory feedback to neural activity in avian song control circuitry. *Journal of Neuroscience*, 28(44), 11378–11390.
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, 14(2), 257–264.
- Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science (New York, N.Y.)*, 340(6129), 216–219.
- Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: How musical sounds become rewarding. *Trends in Cognitive Sciences*, 19(2), 86–91.
- Sammler, D., Grosbras, M. H., Anwander, A., Bestelmeyer, P. E. G., & Belin, P. (2015). Dorsal and ventral pathways for prosody. *Current Biology*, 25(23), 3079–3085.
- Sanchez, J. L. J. (2007). Procession and performance: Recreating ritual soundscapes among the Ancient Maya. *The World of Music*, 49, 35–44.
- Savage, P. E. (2018). Alan Lomax's Cantometrics Project: A comprehensive review. *Music & Science*, 1, 1–19. <http://doi.org/10.1177/2059204318786084>.
- Savage, P. E. (2019a). Cultural evolution of music. *Palgrave Communications*, 5, article 16. <http://doi.org/10.1057/s41599-019-0221-1>.
- Savage, P. E. (2019b). Universals. In J. L. Sturman (Ed.), *The Sage international encyclopedia of music and culture* (pp. 2282–2285). Sage Publications. <http://doi.org/10.4135/9781483317731.n759>.
- Savage, P. E. (2019c). An overview of cross-cultural music corpus studies. In D. Shanahan, A. Burgoyne & I. Quinn (Eds.), *Oxford handbook of music and corpus studies* (pp. 1–15). Oxford University Press. <http://doi.org/10.31235/osf.io/nextbg>.
- Savage, P. E., & Brown, S. (2013). Toward a new comparative musicology. *Analytical Approaches to World Music*, 2(2), 148–197.
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences of the USA*, 112(29), 8987–8992.
- Savage, P. E., Merritt, E., Rzesutek, T., & Brown, S. (2012). Cantocore: A new cross-cultural song classification scheme. *Analytical Approaches to World Music*, 2(1), 87–137.
- Savage, P. E., Yamauchi, M., Hamaguchi, M., Tarr, B., Kitayama, Y., & Fujii, S. (2020). Rhythm, synchrony, and cooperation [Stage 1 Registered Report]. *PsyArXiv*. <https://doi.org/10.31234/osf.io/46bd9>.
- Scarl, J. C., & Bradbury, J. W. (2009). Rapid vocal convergence in an Australian cockatoo, the galah. *Animal Behaviour*, 77, 1019–1026.
- Schachner, A., Brady, T. F., Oro, K., & Lee, M. (Preprint). Intuitive archeology: Detecting social transmission in the design of artifacts. *PsyArXiv*. <https://doi.org/10.31234/osf.io/g6qxx>.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19, 831–836.
- Schino, G., & Troisi, A. (1992). Opiate receptor blockade in juvenile macaques: Effect on affiliative interactions with their mothers and group companions. *Brain Research*, 576, 125–130.
- Schruth, D., Templeton, C. N., & Holman, D. J. (Preprint). A definition of song from human music universals observed in primate calls. *BioRxiv preprint*. <https://doi.org/10.1101/649459>.
- Schulkin, J., & Raglan, G. B. (2014). The evolution of music and human social capability. *Frontiers in Neuroscience*, 8, article 292. <https://doi.org/10.3389/fnins.2014.00292>.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, 275(5306), 1593–1599.
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23, 473–500.
- Sethares, W. (2004). *Tuning, timbre, spectrum, scale*. Springer-Verlag.
- Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The social salience hypothesis of oxytocin. *Biological Psychiatry*, 79(3), 194–202.
- Shilton, D., Breski, M., Dor, D., & Jablonka, E. (2020). Human social evolution: Self-domestication or self-control? *Frontiers in Psychology*, 11(134), 1–22. <https://doi.org/10.3389/fpsyg.2020.00134>.
- Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature*, 479(7372), 219–222.
- Silk, J. B. (2007). Social components of fitness in primate groups. *Science (New York, N.Y.)*, 317, 1347–1351.
- Sloboda, J. A., O'Neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the experience sampling method. *Musicae Scientiae*, 5(1), 9–32.
- Small, C. (1998). *Musicking: The meanings of performing and listening*. University Press of New England.
- Soley, G., & Spelke, E. S. (2016). Shared cultural knowledge: Effects of music on young children's social preferences. *Cognition*, 148, 106–116.
- Stefano, G. B., Zhu, W., Cadet, P., Salamon, E., & Mantione, K. J. (2004). Music alters constitutively expressed opiate and cytokine processes in listeners. *Medical Science Monitor*, 10(6), MS18–27.
- Stevens, C., & Byron, T. (2016). Universals in music processing. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 19–31). Oxford University Press.
- Stokes, M. (Ed.). (1994). *Ethnicity, identity, and music: The musical construction of place*. Berg.
- Stupacher, J., Hove, M. J., Novembre, G., Schutz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, 82(2), 127–136.

- Syal, S., & Finlay, B. L. (2011). Thinking outside the cortex: Social motivation in the evolution and development of language. *Developmental Science*, 14(2), 417–430.
- Tamir, D. I., & Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences of the USA*, 109(21), 8038–8043.
- Tarr, B. (2017). Social bonding through dance and “musiking.” In N. L. Enfield & P. Kockelman (Eds.), *Distributed agency* (pp. 151–158). Oxford University Press.
- Tarr, B., Launay, J., Benson, C., & Dunbar, R. I. M. (2017). Naltrexone blocks endorphins released when dancing in synchrony. *Adaptive Human Behavior and Physiology*, 3(3), 241–254.
- Tarr, B., Launay, J., Cohen, E., & Dunbar, R. I. M. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, 11, 20150767. <http://doi.org/http://dx.doi.org/10.1098/rsbl.2015.0767>.
- Tarr, B., Launay, J., & Dunbar, R. I. M. (2014). Music and social bonding: “self-other” merging and neurohormonal mechanisms. *Frontiers in Psychology*, 5, article 1096. <https://doi.org/10.3389/fpsyg.2014.01096>.
- Tarr, B., Launay, J., & Dunbar, R. I. M. (2016). Silent disco: Dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior*, 37(5), 343–349.
- Tarr, B., Slater, M., & Cohen, E. (2018). Synchrony and social connection in immersive virtual reality. *Scientific Reports*, 8, article 3693. <http://doi.org/10.1038/s41598-018-21765-4>.
- Tchaloova, K., & Macdonald, G. (2020). Opioid receptor blockade inhibits self-disclosure during a closeness-building social interaction. *Psychoneuroendocrinology*, 113, 104559.
- Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception*, 1(3), 276–295.
- Thornton, S. (1995). *Club cultures: Music, media, and subcultural capital*. Wesleyan University Press.
- Thorpe, W. H. (1972). Duetting and antiphonal song in birds: Its extent and significance. *Behaviour*, Supplement, 18, 1–197.
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Für Tierpsychologie*, 20, 410–433.
- Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., ... Deloukas, P. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, 39(1), 31–40.
- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America*, 85(4), 1699–1707.
- Tolbert, E. (1990). Women cry with words: Symbolization of affect in the Karelian lament. *Yearbook for Traditional Music*, 22, 80–105.
- Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, 64(1), 231–255.
- Tomlinson, G. (2018). *Culture and the course of human evolution*. University of Chicago Press.
- Trainor, L. J. (2018). The origins of music: Auditory scene analysis, evolution, and culture in music creation. In H. Honing (Ed.), *The origins of musicality* (pp. 81–112). MIT press.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11(3), 188–195.
- Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, 20(3), 383–396.
- Traulsen, A., & Nowak, M. A. (2006). Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the USA*, 103(29), 10952–10955.
- Trehub, S. E. (2016). Infant musicality. In S. Hallam, I. Cross & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 387–397). Oxford University Press.
- Trehub, S. E., Becker, J., & Morley, I. (2018). Cross-cultural perspectives on music and musicality. In H. Honing (Ed.), *The origins of musicality* (pp. 129–148). MIT Press.
- Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, 16, 193–211.
- Trehub, S. E., Unyk, A. A., Kamenetsky, S., Hill, D., Trainor, L. J. D., Henderson, J. L., & Saraza, M. (1997). Mothers’ and fathers’ singing to infants. *Developmental Psychology*, 33(3), 500–507.
- Turchin, P., Currie, T. E., Whitehouse, H., François, P., Feeney, K., Mullins, D., ... Spencer, C. (2018). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences of the USA*, 115(2), E144–E151.
- Turino, T. (2008). *Music as social life: The politics of participation*. University of Chicago Press.
- Valdesolo, P., & Desteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion (Washington, D.C.)*, 11(2), 262–266.
- Vanden Bosch der Nederlanden, C. M., Joannisse, M. F., & Grah, J. A. (2020). Music as a scaffold for listening to speech: Better neural phase-locking to song than speech. *NeuroImage*, 214, 116767. <https://doi.org/10.1016/j.neuroimage.2020.116767>.
- Vlismas, W., Malloch, S., & Burnham, D. (2013). The effects of music and movement on mother–infant interactions. *Early Child Development and Care*, 183(11), 1669–1688.
- Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2000). *The origins of music*. MIT Press.
- Walum, H., Waldman, I. D., & Young, L. J. (2016). Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. *Biological Psychiatry*, 79(3), 251–257.
- Wan, C. Y., Demaine, K., Zipse, L., Norton, A., & Schlaug, G. (2010). From music making to speaking: Engaging the mirror neuron system in autism. *Brain Research Bulletin*, 82 (3–4), 161–168.
- Wanker, R., Sugama, Y., & Prinage, S. (2005). Vocal labelling of family members in spectated parrotlets, *Forpus conspiciatus*. *Animal Behaviour*, 70, 111–118.
- Watson, A., & Keating, D. (1999). Architecture and sound: An acoustic analysis of megalithic monuments in prehistoric Britain. *Antiquity*, 73, 325–336.
- Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Group music performance causes elevated pain thresholds and social bonding in small and large groups of singers. *Evolution and Human Behavior*, 37(2), 152–158.
- Welch, S. P., & Eads, M. (1999). Synergistic interactions of endogenous opioids and cannabinoid systems. *Brain Research*, 848, 183–190.
- Whitehouse, H. (2004). *Modes of religiosity: A cognitive theory of religious transmission*. AltaMira Press.
- Whitehouse, H. (2018). Dying for the group: Towards a general theory of extreme self-sacrifice. *Behavioral and Brain Sciences*, 41, e192. <http://doi.org/10.1017/S0140525X18000249>.
- Whitehouse, H., François, P., Savage, P. E., Hoyer, D., Feeney, K. C., Cioni, E., ... Turchin, P. (Preprint). Big Gods did not drive the rise of big societies throughout world history. *SocArXiv preprint*. <https://doi.org/10.31219/osf.io/mbnvg>
- Wickler, W. (1980). Vocal Duetting and the pair bond: 1. Coyness and partner commitment: A hypothesis. *Zeitschrift für Tierpsychologie*, 52, 201–209.
- Wiggins, G. A., Tyack, P., Scharff, C., & Rohrmeier, M. (2018). The evolutionary roots of creativity: Mechanisms and motivations. In H. Honing (Ed.), *The origins of musicality* (pp. 286–308). MIT Press.
- Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton University Press.
- Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *The Quarterly Review of Biology*, 82(4), 327–348.
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, 20(1), 1–5.
- Winkler, I., Håden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the USA*, 106(7), 2468–2471.
- 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), e94446. <https://doi.org/10.1371/journal.pone.0094446>
- Wood, A. L. C., Kirby, K. R., Ember, C. R., Silbert, S., Daikoku, H., McBride, J., ... Savage, P. E. (Preprint). The Global Jukebox: A public database of performing arts and culture. *PsyArXiv*. <https://doi.org/10.31234/osf.io/4z97j>
- Wrangham, R. (2009). *Catching fire: How cooking made us human*. Basic Books.
- Zak, P., Kurzban, R., & Matzner, W. T. (2005). Oxytocin is associated with human trustworthiness. *Hormones and Behavior*, 48(5), 522–527.
- Zak, P., Stanton, A. A., & Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS ONE*, 2(11), e1128. <http://doi.org/10.1371/journal.pone.0001128>.
- Zaki, J., & Mitchell, J. P. (2013). Intuitive prosociality. *Current Directions in Psychological Science*, 22(6), 466–470.
- Zatorre, R. J. (2018). Why do we love music? *Cerebrum*. [http://www.dana.org/Cerebrum/2018/Why\\_Do\\_We\\_Love\\_Music/](http://www.dana.org/Cerebrum/2018/Why_Do_We_Love_Music/).
- Zatorre, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences of the USA*, 110 (Suppl. 2), 10430–10437.
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the USA*, 107(13), 5768–5773.

both humans and animals that typically expresses sadness or frustration (Li & Yackle, 2017; Teigen, 2008). In music, sighs are mimicked by a falling narrow melodic motion with a decreasing loudness, a standard device in Western classical music used to signal grief to the listener (Monelle, 2000). Music has also been found to imitate the staccato acoustic profile of laughter, a credible signal found in both humans and many animal species (Bryant, 2020), when communicating humor (Trevor & Huron, 2018). These instances of credible signals in contemporary music are indicative of the continued presence of biologically rooted credible signals in music today, extending the reach of Mehr and colleagues' theory to present day music.

Although such mimicry of vocal signals exists as predicted by the credible signaling theory, many cross-comparisons between music and voices have shown that affective meaning is signaled and perceived more poorly in music than in voices (Frühholz, Trost, & Grandjean, 2014; Juslin & Laukka, 2003; Paquette, Takerkart, Saget, Peretz, & Belin, 2018; Scherer, 1995). For example, Paquette et al. (2018) report overall lower recognition accuracies for fearful, sad, happy, and neutral emotions expressed in music compared to voices. Furthermore, one of our recent studies showed that vocal screams are perceived as significantly more intense and emotionally negative than horror film music excerpts that mimic human screams even though both use the credible signal roughness (Trevor et al., 2020). Affective meaning seems thus less well signaled and recognized in music compared to voices, a difference that is not accounted for in the credibility hypothesis and therefore could be a downside to this theory.

To address these perceptual differences, we propose that the credibility hypothesis could be extended to include a component regarding culturally de-contextualized biological signals. A similar functional de-contextualization component has been described for the evolution of human reasoning (Stanovich & West, 2000). Vocal signals have biological significance, are largely triggered by situational cues, and have direct contextual meanings to listeners (Frühholz & Schweinberger, 2020; Frühholz et al., 2016). On the contrary, musical imitations of these vocal signals are of a more "symbolic" and "fictional" nature, are voluntarily produced along musical principles and cultural rules, and are meant to capture the attention and emotional sway of the listener. The weaker credibility of musically signaled affective meaning could be because of this difference in signal goals and the de-contextualization of the signal. What then is the evolutionary value of these musical signals? The de-contextualized nature of these signals results in the communication of two pieces of information: "emotional fiction" and the biological meaning of the natural signal being imitated. Music-induced emotions are sometimes regarded as "make-believe" emotions, as fictional tools in de-contextualized settings (Walton, 1990). In communicating "emotional fiction," the musical signal tells the listener that the situation is not real, it is a *simulation*. That information might weaken the second part of the signal, the affective impression of the imitated vocal expression. Given this "emotional fiction" component, perhaps the creation of biologically rooted affective musical signals was motivated by an adaptive need for simulating emotional situations.

What evolutionary role do simulations of emotional situations serve? There is a theory that nightmares may have evolved to simulate threatening situations to increase threat preparedness and survival chances in early humans (Revonsuo, 2000). Part of such threat preparedness would include emotional preparedness,

or resilience and emotion regulation skills, because nightmares induce fearful emotions. Some research on other threat simulating activities (horror films and violent videogames) supports this theory. People who enjoy horror movies have been found to be more resilient in the face of real-life dangers, such as the COVID pandemic (Scrivner, Johnson, Kjeldgaard-Christiansen, & Clasen, 2020). Similarly, people who play violent video games have fewer nightmares, suggesting that videogame simulations actually fill that adaptive need for threat simulation (Bown & Gackenbach, 2016). In ancient human cultures, threat simulations were conveyed through *storytelling*. Storytelling is a universal human practice with ancient roots (Smith et al., 2017) and it often involved musical instruments (Pellowski, 1990). Perhaps storytellers were motivated to create sounds that would be similar to real-life signals but also clearly fictional, increasing the emotional impression of the stories and enabling listeners to rehearse the emotions of the tale in a safe, imaginary, and cooperative space.

**Financial support.** C.T. received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement (No. 835682). S.F. received funding from Swiss National Science Foundation (Grants Nos. SNSF PP00P1\_157409/1 and PP00P1\_183711/1).

**Conflict of interest.** None.


## References

- Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A. L., & Poeppel, D. (2015). Human screams occupy a privileged niche in the communication soundscape. *Current Biology*, 25(15), 2051–2056. <https://doi.org/10.1016/j.cub.2015.06.043>.
- Bown, J., & Gackenbach, J. (2016). Video games, nightmares, and emotional processing. In S. Y. Tettegah & W. D. Huang (Eds.), *Emotions, technology, and digital games* (pp. 3–14). Academic Press.
- Bryant, G. A. (2020). Evolution, structure, and functions of human laughter. In K. Floyd & R. Weber (Eds.), *The handbook of communication science and biology* (pp. 63–77). Routledge.
- Engelberg, J. W. M., & Gouzoules, H. (2019). The credibility of acted screams: Implications for emotional communication research. *Quarterly Journal of Experimental Psychology*, 72(8), 1889–1902. <https://doi.org/10.1177/1747021818816307>.
- Frühholz, S., & Schweinberger, S. (2020). Nonverbal auditory communication – evidence for integrated neural systems for voice signal production and perception. *Progress in Neurobiology*, 199, 101948.
- Frühholz, S., Trost, W., & Grandjean, D. (2014). The role of the medial temporal limbic system in processing emotions in voice and music. *Progress in Neurobiology*, 123, 1–17. <https://doi.org/10.1016/j.pneurobio.2014.09.003>.
- Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions – towards a unifying neural network perspective of affective sound processing. *Neuroscience and Biobehavioral Reviews*, 68, 1–15. doi:10.1016/j.neubiorev.2016.05.002.
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129(5), 770–814. <https://doi.org/10.1037/0033-2909.129.5.770>.
- Li, P., & Yackle, K. (2017). Sighing. *Current Biology*, 27(3), R88–R89.
- Monelle, R. (2000). *The sense of music: Semiotic essays*. Princeton University Press.
- Paquette, S., Takerkart, S., Saget, S., Peretz, I., & Belin, P. (2018). Cross-classification of musical and vocal emotions in the auditory cortex. *Annals of the New York Academy of Sciences*, 1423(1), 329–337. <https://doi.org/10.1111/nyas.13666>.
- Pellowski, A. (1990). *The world of storytelling*. H.W. Wilson.
- Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences*, 23(6), 877–901.
- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, 9(3), 235–248. [https://doi.org/10.1016/S0892-1997\(05\)80231-0](https://doi.org/10.1016/S0892-1997(05)80231-0).
- Schwartz, J. W., Engelberg, J. W., & Gouzoules, H. (2019). What is a scream? Acoustic characteristics of a human call type. *The Journal of the Acoustical Society of America*, 145(3), 1776–1776. <https://doi.org/10.1121/1.5101500>.
- Scrivner, C., Johnson, J. A., Kjeldgaard-Christiansen, J., & Clasen, M. (2020). Pandemic practice: Horror fans and morbidly curious individuals are more psychologically resilient during the COVID-19 pandemic. *Personality and Individual Differences*, 168, 110397.



## Authors' Responses

### Toward inclusive theories of the evolution of musicality

Patrick E. Savage<sup>a</sup> , Psyche Loui<sup>b</sup>, Bronwyn Tarr<sup>c</sup>, Adena Schachner<sup>d</sup>, Luke Glowacki<sup>e</sup>, Steven Mithen<sup>f</sup> and W. Tecumseh Fitch<sup>g</sup>

<sup>a</sup>Faculty of Environment and Information Studies, Keio University, Fujisawa 252-0882, Japan; <sup>b</sup>College of Arts, Media and Design, Northeastern University, Boston, MA 02115, USA; <sup>c</sup>Department of Experimental Psychology, Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford OX2 6PN, UK; <sup>d</sup>Department of Psychology, University of California San Diego, La Jolla, CA 92093, USA; <sup>e</sup>Department of Anthropology, Boston University, Boston, MA 02215, USA; <sup>f</sup>Department of Archaeology, University of Reading, Reading RG6 6AB, UK and <sup>g</sup>Department of Behavioral and Cognitive Biology, University of Vienna, Vienna 1090, Austria.

psavage@sfc.keio.ac.jp, <http://PatrickESavage.com>  
p.loui@northeastern.edu, <http://www.psycheloui.com>  
bronwyn.tarr@anthro.ox.ac.uk, [bronwyn.tarr@gmail.com](mailto:bronwyn.tarr@gmail.com),  
<https://www.anthro.ox.ac.uk/people/dr-bronwyn-tarr>  
schachner@ucsd.edu, <https://madlab.ucsd.edu>  
glowacki@fas.harvard.edu, <https://www.hsb-lab.org/>  
s.j.mithen@reading.ac.uk,  
<http://www.reading.ac.uk/archaeology/about/staff/s-j-mithen.aspx>  
tecumseh.fitch@univie.ac.at, <https://homepage.univie.ac.at/tecumseh.fitch/>

doi:10.1017/S0140525X21000042, e121

#### Abstract

We compare and contrast the 60 commentaries by 109 authors on the pair of target articles by Mehr et al. and ourselves. The commentators largely reject Mehr et al.'s fundamental definition of music and their attempts to refute (1) our social bonding hypothesis, (2) byproduct hypotheses, and (3) sexual selection hypotheses for the evolution of musicality. Instead, the commentators generally support our more inclusive proposal that social bonding and credible signaling mechanisms complement one another in explaining cooperation within and competition between groups in a coevolutionary framework (albeit with some confusion regarding terminologies such as “byproduct” and “exaptation”). We discuss the proposed criticisms and extensions, with a focus on moving beyond adaptation/byproduct dichotomies and toward testing of cross-species, cross-cultural, and other empirical predictions.

“Music is the most powerful form of communication in the world. It brings us all together.”

— Sean Combs aka Puff Daddy (Poggi, 2013)

“Nirgends können zwei Menschen leichter Freunde werden als beim Musizieren.”

(There is no easier way for two people to become friends than by making music together)

— Hermann Hesse, *Das Glasperlenspiel* (1943, p. 51)

“who hears music, feels his solitude Peopled at once.”

— Robert Browning, *Balaustion's adventure* (1871, lines 323–324)

#### R1. Introduction

The joint publication of our target article, the companion target article by Mehr et al., and 60 commentaries on these target articles by 109 experts represents a chance to synthesize in a single discussion the complex debate about the origins of music. Such debates date back at least to Rousseau (1760/1998), were developed by Darwin (1871), and have expanded dramatically in the past few decades – notably with the publication of edited volumes and special issues published by MIT Press and *Philosophical Transactions of the Royal Society B* (Honing, 2018; Honing et al., 2015; Wallin et al., 2000).

Although the *Behavioral and Brain Sciences* editors only required us to respond to the commentaries specifically addressing our own target article, they provided us with all accepted responses, including those addressing Mehr et al.'s target article. It became clear when reading these responses that doing justice to the debate would require us to simultaneously address responses to both target articles. This is especially true because Mehr et al. not only describe their own “credible signaling” hypothesis, but also devote substantial space to critiquing three of the most prominent alternative hypotheses: (1) the *social bonding* hypothesis detailed in our target article; (2) the hypothesis originally proposed by Darwin (1871) and championed most notably by Miller (2000) that musicality evolved through *sexual selection*; and (3) the hypothesis popularized by Pinker (1997) that musicality is a *byproduct* of the evolution of language or other adaptations (memorably captured by Pinker's description of music as “auditory cheesecake”).

The combined 60 responses analyze all four hypotheses (credible signaling, social bonding, sexual selection, and byproduct). Because all commentaries focus on one or both target articles, we have created [Figure R1](#) and [Table R1](#) to visualize the degree to which – in our subjective evaluation – each commentary is supportive or critical of the ideas proposed in each of the two target articles. This allows us to easily visualize the broad space of agreement/disagreement among the responses and highlight the relationships between particularly notable commentaries.

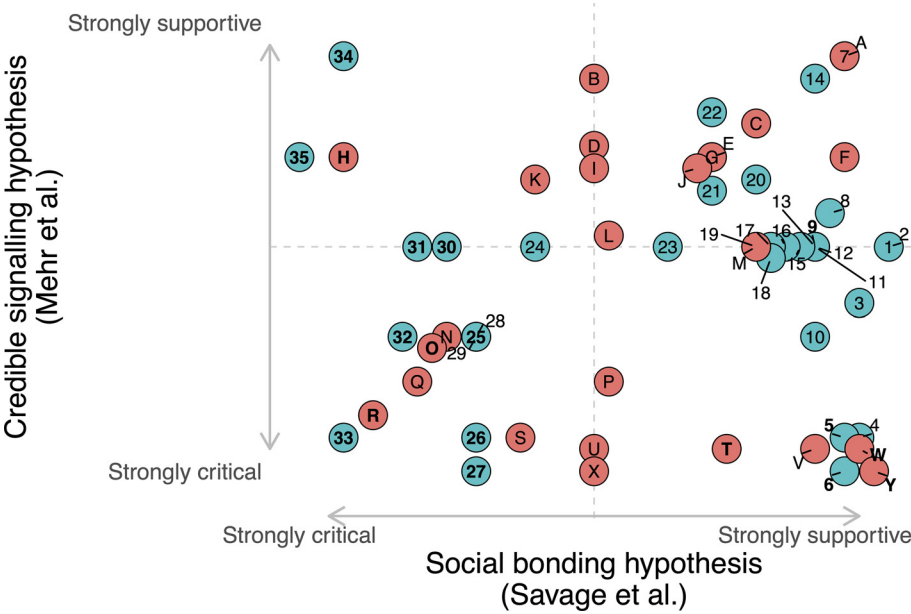
Across all commentaries, four key themes repeatedly emerge: (1) defining “music” and “musicality”; (2) relationships between the social bonding and credible signaling hypotheses, (3) distinguishing between adaptations and byproducts; and (4) extensions/applications/tests of the hypotheses. We have highlighted in bold the 16 commentaries that we believe most comprehensively capture the full spectrum of debate. In the following sections, we will address each of these key themes in detail, with a particular emphasis on these 16 commentaries.

#### R2. Defining “music” and “musicality”

The definitions of the fundamental terms “music” and “musicality” were critiqued by a number of commentators. We avoided providing a precise definition of “music,” citing long-standing debates regarding “practical and ethical challenges involved in defining and comparing ‘music’ and ‘musicality’ in cross-culturally valid ways.” Mehr et al. offered the following definition: “Music is an auditory display built from melodies and rhythms.”

Cross; Iyer; Margulis; and Wald-Fuhrmann, Pearson, Roeske, Grüny, & Polak (Wald-Fuhrmann et al.) all noted the dangers of ethnocentrism in defining music as a purely auditory phenomenon in terms derived from European heritage. Dissanayake; Sievers & Wheatley; and Trehub also pointed out





**Figure R1.** Visual comparison of the 60 commentaries responding to the pair of target articles, based on our subjective evaluation of the degree to which they are supportive or critical of each target article. Figure R1 plots the average of subjective ratings by PES and PL on a scale from –10 (“strongly critical”) to 10 (“strongly supportive”). Agreement between the two raters was high (intraclass correlation coefficient (ICC) = 0.89). See [github.com/comp-music-lab/social-bonding](https://github.com/comp-music-lab/social-bonding) for full data and code. Responses published with our target article are ordered using numbers (1–35; colored blue), whereas those published with Mehr et al. are ordered using letters (A–Y; colored red). Key commentaries discussed in detail in our response are highlighted in bold.

**Table R1 (Savage et al.).** List of the 60 commentaries accompanying target articles by Mehr et al. and ourselves

| Savage et al. responses                                       | Mehr et al. responses                             |
|---|---|
| 1. del Mastro, Strollo, & El Maj                              | A. Akkermann, Akkaya, Dermiel, Pflüger, & Dresler |
| 2. Wang & Zou   | B. Lumaca, Brattico, & Baggio                     |
| 3. Margulis   | C. Dubourg, André, & Baumard                      |
| 4. Gabriel & Paravati   | D. Hilton, Asano, & Boeckx                        |
| 5. <b>Popescu, Oesch, &amp; Buck</b>                          | E. Snyder & Creanza                               |
| 6. <b>Rendell, Doolittle, Garland, &amp; South</b>            | F. Kasdan, Gordon, & Lense                        |
| 7. Morrison   | G. Levitin  |
| 8. Ravignani  | H. <b>Kennedy &amp; Radford</b>                   |
| 9. <b>Eirdosh &amp; Hanisch</b>                               | I. Trevor & Frühholz                              |
| 10. Hattori   | J. Sievers & Wheatley                             |
| 11. Belfi   | K. Zentner  |
| 12. Juslin  | L. Moser, Ackerman, Dayer, Proksch, & Smaldino    |
| 13. Grahn, Bauer, & Zamm                                      | M. Gardiner                                       |
| 14. Benítez-Burraco   | N. Pinker   |
| 15. Sachs, FeldmanHall, & Tamir                               | O. <b>Trainor</b>                                 |
| 16. Ashley  | P. Scott-Phillips, Tominaga, & Miton              |
| 17. Brown   | Q. Számadó  |
| 18. Hansen & Keller   | R. <b>Stewart-Williams</b>                        |
| 19. Pfordresher   | S. Tichko, Bird, & Parker                         |
| 20. Dissanayake   | T. <b>Bowling, Hoeschele, &amp; Dunn</b>          |
| 21. van Mulukom   | U. Levaïda  |
| 22. Gingras   | V. Trehub   |
| 23. Fritz   | W. <b>Wood</b>                                    |
| 24. Zhang & Shi   | X. Lieberman & Billingsley                        |
| 25. <b>Honing</b>   | Y. <b>Killin, Brusse, Currie, &amp; Planer</b>    |
| 26. <b>Wald-Fuhrmann, Pearson, Roeske, Grüný, &amp; Polak</b> |   |
| 27. <b>Cross</b>  |   |
| 28. Hannon, Crittenden, Snyder, & Nave                        |   |
| 29. Iyer  |   |
| 30. Kraus & Hesselmann  |   |
| 31. Atzil & Abramson  |   |
| 32. <b>Patel &amp; von Rueden</b>                             |   |
| 33. <b>Harrison &amp; Seale</b>                               |   |
| 34. <b>Merker</b>   |   |
| 35. <b>Verpooten &amp; Eens</b>                               |   |

the need for a multimodal treatment including movement (e.g., dance), touch, and so on in addition to sound.

We have previously explained why cross-culturally universal definitions of “music” are not possible, particularly when it comes to delineating speech from song or music from dance (Savage, 2019b; Savage et al., 2015). Instead, a more useful definition cited in our target article is Honing’s (2018) distinction between “music” as cultural products (songs, instruments, dance styles, etc.) and “musicality” as the set of biological capacities underlying the creation of those products. Although this circular definition leaves unanswered the unanswerable question of defining music itself, it does allow us to focus on the ways that cultural and biological evolution can work in tandem and in parallel to produce the diverse products around the world that many recognize as “music” or “music-like.”

Honing and Wald-Fuhrmann et al. accurately note that at times both target articles fail to carefully distinguish between “music” and “musicality,” and that in some cases we might have more appropriately focused on musicality, not on music. Indeed, in retrospect a more accurate title for our target article might have been “*Musicality as a coevolved system for social bonding*” (just as a more accurate title for Mehr et al. might have been “*Origins of musicality in credible signaling*”).

**R3. Social bonding versus credible signaling**

We start by noting significant areas of agreement and/or synergy between the two target articles. First, both articles agree that music’s social aspects are the strongest candidate for adaptive functions. Second, our focus on gene-culture coevolution in our paper is endorsed by Mehr et al., although they do not pursue this idea in depth. Third, we agree that musicality has deep roots in nonhuman animal vocalizations.

One primary difference between the two target articles is that Mehr et al. spent the bulk of their article refuting alternative theories, whereas we attempted to synthesize several existing theories into a broader, more inclusive framework. Based on our reading, as well as the commentaries, we argue below that Mehr et al.’s

critiques do not succeed in showing that credible signaling is the sole or primary cause of the evolution of musicality. Instead, we believe that the credible signaling hypothesis can be incorporated as one sub-component of our broader, more inclusive framework.

### ***R3.1 Social bonding and credible signaling are complementary, not mutually exclusive***

The two target articles have an asymmetrical relationship. Although we did attempt to describe ways in which the social bonding and credible signaling hypotheses might produce contrasting predictions that could be tested experimentally (cf. sect. 6.5), ultimately we stated that “Bonding and signaling hypotheses are not mutually exclusive, but rather complementary.” In contrast, Mehr et al. devote over 2,000 words to categorically rejecting the social bonding hypothesis, arguing that “music does not *directly* cause social cohesion: rather, it signals existing social cohesion that was obtained by other means” (target article, sect. 4.2.1, para. 14).

Overall, the most consistent point unifying multiple commentaries was a consensus in favor of our argument of complementarity (e.g., **Benítez-Burraco; Gingras; Honing; Juslin; Morrison; Trainor**), and against Mehr et al.’s of mutual exclusivity. Only three commentaries (**Kennedy & Radford; Pinker; and Zentner**) appeared convinced by Mehr et al.’s arguments against social bonding.<sup>1</sup> In contrast, many commentators rejected these arguments, for a variety of reasons, including: (1) they turn the origin of the social cohesion being signaled into “somebody else’s problem” (**Rendell, Doolittle, Garland, & South [Rendell et al.]**); (2) they are inconsistent with substantial experimental evidence showing causal effects of synchrony on cooperation (**Gabriel & Paravati; Wood**); (3) they incorrectly assume that music-making is a purely altruistic sacrifice that does not benefit the performer (**Harrison & Seale**); (4) their criticisms of social bonding apply equally to their own favored hypothesis (**Bowling, Hoeschele, & Dunn [Bowling et al.]**); and (5) they rely on a “misguided” adaptation-byproduct dichotomy (**Killin, Brusse, Currie, & Planer [Killin et al.]**) that “do[es] not reflect the nuance of current evolutionary thinking” (**Rendell et al.**). We will return to this adaptation-byproduct dichotomy in detail in section R4, as it is a primary source of confusion and disagreement.

Our social bonding account incorporates some discussion of ways music may function as an honest social signal (e.g., of social or cultural background), and how this likely contributes to social bonding, rather than simply reflecting pre-existing bonds (sects 2.2.4 and 3.3). **Dubourg, André, & Baumard; Harrison & Seale; Kennedy & Radford; and Killin et al.** further argue that the two hypotheses are even more complementary than we had implied, suggesting that the social bonding hypothesis would be enhanced by more explicitly integrating the role of signaling. However, as **Rendell et al.** put it: “one surely has to have a social bond before one can credibly signal about it,” a sequence also endorsed by **Benítez-Burraco** and **Hattori**. **Popescu, Oesch & Buck** go even further to characterize credible signaling as “a special case of [social bonding], albeit that signaling focuses on between-groups and social-bonding focuses on within-group relations,” a distinction also echoed by **Hansen & Keller**.

After reading the commentaries, we agree that credible signaling should be integrated into our hypothesis to more explicitly

account for interactions between groups. Such integration follows naturally from our discussion in section 6.4 of our target article on “Parochial altruism and outgroup exclusion,” and from **Figure 1** in our target article, which showed that we see the “war songs and lullabies” (Washington State University, 2020) championed by Mehr et al. as “sub-components of a broader social bonding function.” This also is consistent with **Pinker’s** critique that Mehr et al.’s (2019) own study found that war songs and lullabies were not more widespread than any of the other 18 genres that they analyzed, all 20 of which we argued represent different expressions of social bonding.

Importantly, Mehr and colleagues’ critiques are directed at an omnibus “social bonding hypothesis” for which they list 33 references, not including our own (Mehr et al., sect. 3.2, para. 1). This means that many of their critiques do not apply to our current hypothesis (which was intended to extend and clarify previous study). For example, their argument that “the” social bonding hypothesis conflates proximate- and ultimate-level reasoning does not bear on our proposal: We explicitly distinguish between functional and mechanistic levels of explanation, and add phylogenetic and ontogenetic levels (cf. **Fig. 2** in our target article). The same applies regarding their requirement for genetic group selection in the evolution of musicality: this is Steven Brown’s hypothesis, not ours (we explicitly eschew any such requirement, see section 6.2 in our target article and section R3.2).

There are three major specific differences between our and Mehr and colleagues’ arguments: (1) We posit a broad and inclusive hypothesis about the adaptive functions of musicality (which includes both the infant-directed songs and coalition signaling proposed by Mehr et al. as special cases; cf. **Fig. 1** in our target article). (2) We argue that the design features of music make it better suited to social bonding than other ancestral bonding mechanisms (ABMs) such as grooming, or than language. Mehr et al. assert that “language adequately provides whatever social functions grooming may have” and that “music thus appears to have no advantages over language and many disadvantages” (sect. 3.2.2, para. 5). We disagree, and our target article specifies how multiple specific features of musicality outperform the functions of group coordination and bonding relative to language or ABMs (cf. sects 2, 5.1, and 6.1 in our target article, and cf. **Bowling et al.**). (3) Mehr et al. see group music-making as broadcasting an honest signal of social bonds, but crucially argue that these bonds are formed through some other unspecified means. In contrast, we see music as providing a medium or domain in which such bonds can be developed and strengthened, and see this as parsimoniously related to the idea that music also serves as a signal of these bonds.

By Mehr et al.’s hypothesis, group singing is a simple, direct signal of coalitionary strength, directed outside the group, that indexes past practice: “a high level of synchronous coordination among signalers requires considerable effort to achieve” (sect. 4.2.1, para. 4) If so, why does group singing have features, such as steady rhythm, that make it easy for an outsider to join in (cf. **Wood**)? Why isn’t maximization of raw acoustic energy – an honest signal of group size and coordination, achieved by simultaneous calling in many insects and frogs (**Greenfield, 2005**) – the norm in group performances? By our account, rhythm provides a rich domain enabling multiple types of meaningful social interactions, including “crutches” allowing easy engagement (e.g., isochronicity), AND space for individual embellishment and showing high levels of skill (e.g., meter),

AND the potential for cultural embellishments that could serve as shibboleths for group membership. For example, Balkan additive meters can be easily parsed by infants but are difficult to process for North American adults (Hannon & Trehub, 2005) – just the developmental characteristics expected for a shibboleth. Each of these expressive channels can serve both inter- and intra-group signaling, and it seems procrustean to single out one as *the* “proper function” (cf. Gingras) – particularly once cultural evolutionary processes are overlain on ordinary biological evolution by genetic change. Finally, Mehr et al.’s argument that stress-reduction is “superfluous” because “the *net* fitness benefits of sociality exceed those of solitary life” ignores the fact that once group living is established in a species (as it is for most terrestrial primates; cf. Shultz et al., 2011) any additional adaptations that further reduce the costs of group living and/or increase its benefits will be selected (e.g., better cooperation for group defense or hunting; cf. Bowling et al.).

In summary, we do not see our hypothesis as “diametrically opposed” to that of Mehr et al. (contra Kennedy & Radford), but rather see ours as a broader and more inclusive superset, encompassing aspects of the hypotheses of Mehr et al. and many others.

### R3.2 Multilevel selection

The idea that social bonding and credible signaling may be working in parallel at within- and between-group levels provides a potential solution to the issue of multilevel selection raised by Brown; Eirdosh & Hanisch; and Moser, Ackerman, Dayer, Proksch, & Smaldino. These authors were not convinced by our brief attempt in section 6.2 to side-step long-standing debates about group selection by arguing it is not required for our hypothesis. Eirdosh & Hanisch, in particular, argue that the social bonding hypothesis logically requires us to embrace group selection, because “one would be hard pressed to argue that [social bonding] functions of musicality increase the relative fitness of individuals compared to their (presumably equally socially bonding) group members.” We disagree: This statement assumes that musical performance bonds all group members identically. In contrast, (1) within any group individual variation exists, and (2) individuals can and do form sub-groups who share stronger bonds than with others in the group. Individual selection at a local level, because of some group members accruing more or stronger bonds than others, can drive the genetic evolution of musicality without the need for genetic group selection at a global level (although it obviously does not preclude additional between-group selection).

Although we embrace cultural group selection (Boyd & Richerson, 1985; Richerson et al., 2016), we think it is crucial to distinguish this from the genetic group selection endorsed by David Sloan Wilson and colleagues (Eldakar & Wilson, 2011; Sober & Wilson, 1998), particularly when gene-culture coevolution is under discussion (Brown & Richerson, 2014; West, Griffin, & Gardner, 2008). Despite some differences among us (the target authors) regarding our enthusiasm for multi-level selection theory, we agree in rejecting Eirdosh & Hanisch’s claim that it is logically necessary for our hypothesis to work.

### R3.3 Signaling theory

Contra Kennedy & Radford, we neither reject signaling theory, nor dispute the idea that music conveys information. At issue here is what *type* of information music conveys, and *to whom*.

We find Mehr et al.’s claim that we focus “on the neurobiology of the performers, rather than...information encoded in music” a false dichotomy: both domains are important and interact, as shown in Figure 2 in our target article. Indeed, as noted by Margulis, we specifically gathered a team of authors with expertise spanning neuroscience, musicology, psychology, anthropology, evolutionary biology, and other fields in order to synthesize these domains and avoid such dichotomies. We see no compelling reason to choose between neuroscience and signaling (cf. Killin et al.; Rendell et al.).

By our hypothesis, information concerning rhythm (e.g., tempo and meter) and melody/harmony (e.g., pitch range and key) is crucial to achieve synchronization and coordination, and thus to achieve optimal social coordination and bonding within a group. This is echoed by Grahm, Bauer, & Zamm (Grahm et al.), with the amendment that although entrainment of bodies and minds may be a key mechanism by which music confers its effects on social bonding, accurate entrainment ability may not be required for such effects. We see musical information as directly serving social bonding functions, rather than solely signaling extra-musical information (e.g., group size or coalition strength) as Mehr et al. hypothesize. However, this does not prevent other listeners from extracting extra-musical information from a performance (e.g., about the sex of performers or group size). Instead, we suggest that such extraction is not necessary for music to have adaptive value.

Turning to the costs of musical signals, we disagree with Kennedy & Radford that high costs are required to “maintain the credibility of diverse signals across the natural world.” Despite its remarkable persistence, Zahavi’s “handicap principle” that high costs are required to maintain honesty is argued by some to be a fallacy (Maynard Smith, 1976; Penn & Számadó, 2020; Számadó, 2011). Low-cost signaling can be evolutionarily stable whenever interests are aligned (e.g., among relatives because of inclusive fitness benefits; Bergstrom & Lachmann, 1998), and in so-called “indices,” physical or anatomical constraints that can enforce honesty with zero handicap or “strategic” costs (Fitch & Hauser, 2002; Maynard Smith & Harper, 2003).

We certainly agree that evolutionary models for musicality should take the costs of signaling into account. Unfortunately, there is very little empirical data upon which to base such theorizing. Human vocalization is in general low cost; for quiet speech this cost is almost unmeasurable (Moon & Lindblom, 2003). Based on physiological principles (Titze, 1994) and animal research, loud singing is somewhat more metabolically costly than normal speech (Oberweger & Goller, 2001; Ward, Speakman, & Slater, 2003), and vigorous dancing is probably an order of magnitude more metabolically costly than song. Accepting this presumed ranking, we might hypothesize that high-cost dance can serve as a more honest signal of current energy and investment than lower cost song. Song may instead signal past practice, knowledge, cultural embeddedness, or other social information. Further empirical data are required to ground and test this or similar hypotheses.

Finally, the apparent disagreement between us and Mehr et al. on the intended recipient of the musical signal may reflect a false dichotomy. By our argument, the musical signal is primarily directed within the group, and for Mehr et al. it is directed to other, competing groups. But, even a signal “intended” by its emitter for a particular listener can be intercepted by an eavesdropper (McGregor & Dabelsteen, 1996), and the resulting effects



(positive or negative) can in turn lead to selection on the original signal (Ryan, 1985). Thus it seems reasonable to accept that music plays both intra- and inter-group signaling roles.

### R3.4 Sexual selection mechanisms cannot be ruled out

Several commentators were unconvinced by Mehr et al.'s argument that the sexual selection hypothesis is refuted by a lack of musical sex differences in humans. **Merker** and **Verpooten & Eens** noted that sex differences are not necessarily required for sexual selection, whereas **Bowling et al.** note that the human voice is in fact unexpectedly sexually dimorphic relative to other primates. Although Mehr et al. argue that "A lone report of sex differences in the frequency of music performance across human societies (Savage et al. 2015) is likely the result of sampling bias," we note that the predominance of male performers is replicated in other studies by Mehr and colleagues involving a "representative sample of human music" (Mehr & Singh et al., 2018, 2019).<sup>2</sup>

We emphasize that cross-cultural sex differences in the frequency of music performance among humans are more likely because of the cultural evolution of patriarchal restrictions on female performance than to biology (Savage et al., 2015).<sup>3</sup> However, as we have described, such cultural evolution can have feedback effects on the biological evolution of musicality. We restate our position from section 6.5 of our target article that we do *not* reject the sexual selection hypothesis and that we encourage cross-species and other comparative analyses that might enable quantification and testing of the relative effects of sexual selection, social bonding, and other factors on the evolution of musicality.

We found **Merker's** statement that we believe "not one of these [mechanisms of musicality] evolved by ordinary natural or sexual selection" puzzling. Our hypothesis is not a blanket appeal to the Baldwin effect for all aspects of the evolutionary process. We fully agree that "ordinary" natural and/or sexual selection must have played a role during certain stages in the protracted evolution of musicality. For example, we agree that vocal learning is a central capacity for musicality, and that the underlying neural circuitry had to evolve biologically (both in humans and other species). We simply observe that, once vocal learning is in place, cultural evolution becomes almost inevitable, and posit that in some cases this could modify selective regimes ("niche construction"), leading to gene-culture coevolution.

### R3.5 The evolutionary age of musicality

A surprising number of commentators accepted Mehr et al.'s mischaracterization of our hypothesis as proposing that "musicality arose fairly recently" on the order of "tens of thousands of years." We made no such claim. Instead, given its universality across the world's cultures, the evolution of human musicality must have been largely completed by the time modern humans expanded out of Africa about 100,000 years ago. The sophistication of 40,000 year old bone flutes (cf. sect. 3.2 in our target article) suggests that the evolution of musicality was already far progressed at that date, and our coevolutionary model posits cycles of gene-culture coevolution preceding these dates considerably. Although hard evidence is absent, this leads us to suspect that musicality had its beginnings considerably before modern

*Homo sapiens*, probably in *Homo erectus* or even earlier (Mithen, 2005). Both fossil and comparative evidence suggests that early *Homo* would have had the ability to make a wide range of vocalizations, body movements, and gestures, especially after the appearance of full bipedalism at c. 1.8 mya, suggesting that some initial form of proto-musicality dates back to that time. We further speculate that our extinct Neanderthal and Denisovan cousins may well have used musicality for social bonding (although a pierced bone claimed to be a Neanderthal flute from Divje Babe cave in Slovenia may simply be a carnivore-chewed bone; cf. D'Errico, Villa, Llonca, & Idarraga, 1998; Kunej & Turk, 2000). A rough time period for the evolution of musicality spans over 1 million years (Tomlinson, 2015).

## R4. Adaptation, byproducts, and exaptation

The point of most disagreement among commentators revolved around the venerable question of whether musicality is an adaptation or a byproduct of some other adaptation. **Harrison & Seale**; **Leivada**; **Lieberman & Billingsley**; **Pinker**; and **Stewart-Williams Zhang & Shi** appear to support a version of **Pinker's** (1997) hypothesis that musicality is primarily a byproduct of language evolution (or at least felt there was not enough evidence to reject this hypothesis). Others pointed to domains other than language as the adaptive source of musicality, such as auditory scene analysis (**Trainor**), prediction reward (**Atzil & Abramson**; **Kraus & Hesselmann**), pre-hunt charade (**Szamadó**), artistic symbolism (**van Mulukom**), hierarchical processing (**Hilton, Asano, & Boeckx**), and mother-infant mutuality (**Dissanayake**).

Mehr et al.'s arguments against byproduct explanations were largely rejected by these commentators. But, although some commentators (e.g., **Harrison & Seale**; **Trainor**) also believed that we too were trying to overturn byproduct explanations, we stated in our target article that adaptation-byproduct relationships between music, language, and other social behaviors remain "open to debate." Rather, our goal was to move beyond the "misguided" (**Killin et al.**), "over-simplistic" (**Rendell et al.**) adaptation-byproduct dichotomy underlying earlier debates, toward a more nuanced continuum incorporating concepts such as exaptation and gene-culture coevolution. Our argument explicitly built on the proposal of **Patel**, who was originally one of the strongest supporters of the idea that music was a purely cultural invention (**Patel, 2008, 2010**), but recently modified his view to include exaptation and gene-culture coevolution of musicality (**Patel, 2018**). This coevolutionary approach does not reject byproduct explanations entirely; instead, as **Degen (2020)** noted, it supports "having **Pinker's** cheesecake and eating it too."

We particularly wish to emphasize the important distinction between "byproducts" and "exaptations" discussed by **Bowling et al.**; **Dissanayake**; and **Killin et al.** We distinguish byproducts (which have no function) from exaptations (where a trait is put to new use, and is functional, but not shaped by selection for that purpose). Most of the commentators supporting variants of **Pinker's** byproduct hypothesis appear to miss this distinction (e.g., when **Harrison & Seale** offer spider webs as an example of a "byproduct account," or when **Trainor** uses "byproduct" and "exaptation" interchangeably). As Darwin recognized with his famous example of lungs and swim bladders (Darwin, 1859), and Gould and Vrba stressed when introducing the term exaptation using examples such as feathers, most complex adaptations have gone through

multiple changes in function, and thus started life as exaptations (Gould, 1991; Gould & Vrba, 1982).

Note that hypotheses about common phylogenetic origins do not preclude special adaptation to a new function: the fact that mammalian middle ear bones originated as jaw bones does not make them “byproducts” of chewing (Fitch, 2010). They may have constituted exaptations for audition initially, but once variants were selected for this new function they became bona fide adaptations for hearing. Similarly, if Darwin was correct that music and language share a common origin, the function of this original “protolanguage/protomusic” may remain the same in the “daughter” systems (e.g., social bonding) or have changed (e.g., propositional information transfer for language and bonding via prediction enhancement for music). But, in neither case would music constitute a “byproduct” of language – more an evolutionary fellow traveler.

Asking whether “music is an adaptation” (as Mehr et al. and Stewart-Williams do) oversimplifies these issues, and obscures precisely the sorts of questions that biomusicology should be confronting, by distinguishing “music” from musicality, exaptations from byproducts, and phylogenetic from adaptive functional explanations (Tinbergen, 1963). For example, we agree with Trainor that the complex perceptual processes underlying pitch perception, where many harmonics are fused into a perceived whole indexed by its fundamental frequency, plays an important role in auditory scene analysis and probably evolved in early vertebrates in that context (Trainor, 2018). Their initial use in music was thus an exaptation. But, these mechanisms appear likely to have been further fine-tuned in the human musical context of group singing, as relative pitch perception is typical of most humans but not most other animals (Hoeschele et al., 2018). Further evidence for the fine-tuning of pitch perception for music comes from people with congenital amusia, who have selective impairments in fine-grained pitch perception especially from the lower harmonics (Cousineau, Oxenham, & Peretz, 2015; Peretz et al., 2002), but show no impairments in pitch-based perceptual organization or auditory scene analysis (Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Peretz & Hyde, 2003). Thus, even if human pitch perception started as an exaptation of scene analysis, it seems plausible that later biological evolution could have fine-tuned this mechanism to its new use in musicality and group singing.

## R5. Tests, extensions, and applications

### R5.1 Explaining solo music-making

We agree with Fritz; Patel & von Rueden; Wald-Fuhrmann et al. and Zentner that the role of solo music-making in our hypothesis requires explanation. But, these commentators appear to overlook the crucial point we made in section 6.5 of our target article that music is often performed by a soloist or listened to by an individual in order to bond with others, to practice prior to group music-making, or to remember past social experiences. Most of the counter-examples cited fit this mold. For example, Patel & von Rueden follow their main counter-example that “Tsimané music-making was largely solo” with the explanation that these solo songs “conveyed traditional knowledge, reinforced cultural norms, and propitiated ancestors and the guardian spirits of forest animals.” Cultural evolutionary theories of religion, prosociality, and cultural transmission would treat all of these as crucial social functions facilitated by music (Norenzayan et al., 2016). Similarly, Fritz’s counter-example of people selecting “Desert

Island Discs” they would want to listen to if stranded alone highlights the social power of solo listening. In our qualitative experience listening to this (fantastic!) show, the vast majority of music is selected specifically to cherish the memories of the most important people in the listener’s life – to feel their “solitude Peopled,” in the words of Browning’s epigraph. Indeed, del Mastrao et al. emphasize that musical memories are often among the last connections to others preserved by patients with Alzheimer’s or other forms of dementia. We thus disagree with Zentner’s claim that “if music had a social purpose, this purpose seems to have largely vanished.” This social purpose is alive and well, although it takes new forms, in solo listening.

Clearly, however, cultural evolution can have strong effects on the frequency of group music-making (cf. Scott-Philipps et al.). The recent prevalence of recorded music and headphones (Thompson & Olsen, 2021) is a case in point, as we discussed in section 2.5 of our target article. Although we agree with Wald-Fuhrmann et al.’s observation that solitary musicking is “extremely common,” cross-cultural analyses show that group music-making is much more common once the effects of recent expansion of Western music and culture have been controlled for (Lomax, 1968; Mehr et al., 2019; Savage et al., 2015). However, cross-cultural variation in the relative frequency of virtuosic “presentational” versus communal “participatory” musicking provides useful testing grounds for the mechanisms and predictions we outlined in section 5.2 of our target article. We welcome proposals by Benítez-Burraco; Patel & von Rueden; and others to expand and refine these predictions, including co-relationships between music and language.

We disagree with Wald-Fuhrmann et al. that “solitary musicking” is “not predicted by any of the proposed evolutionary explanations.” For instance, solitary song is typical of songbirds as they acquire and perfect their song, and there is no difficulty explaining at least some solitary human music making in the same way (“practice makes perfect”). Young birds engage in solo “subsong” and young sac-winged bats “babble” as they develop their local group’s song (Knörnschild, Behr, & von Helversen, 2006; Marler & Peters, 1982). Note that a “solo” performance to an audience can also provide a group bonding experience for those attending, particularly if they dance, clap along or are otherwise engaged. Nonetheless, we agree with Patel & von Rueden that the evolution of musicality could have proceeded from originally solo/presentational performance, or that solo music today may be an offshoot of musicality originally evolved in a group/participatory context.

### R5.2 Cross-species testing

A number of the most interesting commentaries suggested ways to extend and test the cross-species predictions we listed in section 5.3 of our target article. Given the facts that music does not itself fossilize (Honing) and that intra-species evidence for genetic variation in humans explicitly linked to musicality are notoriously difficult to identify (Pfordresher; Tichko, Bird, & Parker [Tichko et al.]), cross-species comparisons with extant nonhuman species may be the most promising candidate for testing many of our predictions.

The most forceful empirical challenge came from Verpooten & Eens, who offered a qualitative analysis of avian vocalizations, suggesting that species with complex social systems (e.g., the fission/fusion lifestyle typifying many parrots), tend to feature short “unmusical” calls, whereas subjectively “music-like” songs

are found in many birds with simpler (e.g., monogamous) social systems. We welcome this potential comparative test, but note two distinctions important in evaluating the social bonding hypothesis. First, social complexity is difficult to measure (Bergman & Beehner, 2015; Turchin et al., 2018), and monogamy and joint parental care pose considerable cognitive challenges relative to solitary living (Burley & Kristine, 2002; Lukas & Clutton-Brock, 2013; Shultz et al., 2011). Second, virtually all bird species have calls – typically mostly unlearned – and these are indeed often shorter and simpler than display vocalizations such as song. Calls serve a wide variety of specific functions – food, alarm, and mobbing calls are common – and their brevity and simplicity often reflect these clear adaptive functions (Marler, 1955). Comparing calls with songs requires caution, because they are neither homologous vocalization types, nor analogous in function (cf. Lorenz, 1971; Peters, 2002).

The social bonding hypothesis predicts that learned song should be more complex than unlearned song (e.g., in songbirds and suboscines), and learned calls should be more complex than unlearned calls (cf. Sewall, Young, & Wright, 2016) and acoustic complexity in either case should increase with social complexity. Indeed by Fitch's (2006) definition, *learned* contact calls, such as the signature whistles of dolphins, parrot contact calls, or the rhythmic codas of sperm whales, are "songs," and indeed appear considerably more complex than typical unlearned calls, although their brevity perhaps makes the musical term "riffs" more appropriate than "songs." Finally, comparisons of the same vocal type *within* a species would be valuable; for instance Freeberg (2006) found that chickadees living in larger groups use more complex (learned) calls than those in smaller groups. We strongly agree with Bowling et al.; Hattori; Ravignani; Rendell et al.; Snyder & Creanza; Tichko et al.; and Verpooten & Eens that comparative data are crucial for testing the social bonding hypothesis, but care is required in executing such analyses, as is avoiding human subjective evaluations of how "music-like" a particular vocalization is. We think the qualitative proposals by these commentators are excellent starting points for future quantitative tests of the social bonding hypothesis and alternative hypotheses.

### R5.3 Extending the neurobiological mechanistic model

Several commentators pointed out potential extensions to our proposed neurobiological model regarding the mechanisms underlying musicality's social bonding functions. The multiple neuroanatomical regions highlighted in Figure 3 of our target article were not meant to provide an exhaustive list of brain regions involved in music processing, or of brain regions that relate music to social behavior, and we agree with Fritz that future iterations of this model should add more specific areas and networks. Our neuroanatomical model was meant as a starting list of candidate neurobiological systems and pathways that we know to underlie certain components of social bonding (such as identity fusion or coalition formation; cf. Sachs, FeldmanHall, & Tamir [Sachs et al.]) and the processing of musical features. We agree with Belfi that simultaneous disruption of two cognitive processes from damage to the same region (e.g., vmPFC damage) does not necessarily imply that the processes are related or the same. We also agree with Juslin that a productive way forward would be to reconcile the contributions of discrete components of the BRECHEMA framework of musical emotions (Juslin, 2019) with neurobiological systems such as the perception and action

network, the dopaminergic reward system, and the endogenous opioid system.

Atzil & Abramson and Kraus & Hesselmann noted the importance of prediction, which Figure 2 of our target article emphasized plays a central mechanistic role in our model. We argued that prediction is key for its proximate ties to reward and learning, but agree that it also ties in with allostasis (Atzil & Abramson) and neural entrainment (Grahn et al.). However, we view the ultimate functions (enhanced within-group bonds, improved group coordination, and group membership cues) as a different level of analysis from the proximate mechanisms of prediction and reward, and the neurobiological systems outlined in Figure 3 of our target article. In our view, musicality evolved with and for social bonding via enhanced predictions; there is no need to "question the implied causality" (cf. Kraus & Hesselmann).

### R5.4 Extensions and applications

A large number of commentators expressed general support for the social bonding and/or credible signaling hypotheses, and detailed how these hypotheses could be extended/applied in various ways. Such applications/extensions include: clinical applications in patients with amnesia/Alzheimer's disease (del Mastrao et al.) and neurodevelopmental disorders (Kasdan et al.); applications to music education (Morrison) and sleep research (Akkermann et al.); proposing additional behavioral experiments to explore relationships between specific musical features and specific psychological mechanisms (Sachs et al.); proposing additional cultural transmission experiments to explore mechanisms of cultural evolution (Lumaca et al.; Scott-Philipps et al.); theoretical extension to the evolution of dance (Brown), gesture (Gardiner), play (Ashley), and story-telling (Trevor & Frühholz); exploring coevolution of music and language (Benítez-Burraco); incorporation of the role of knowledge songs (Levitin); cross-cultural extensions to Chinese music (Wang & Zou); capturing variation in musicality at the levels of development (Hannon et al.), vocal production (Pfordresher), and genomes (Tichko et al.); and further details of neurobiological mechanisms including the roles of ventromedial prefrontal cortex (Belfi), the cerebellum (Fritz), oxytocin (Hansen & Keller; cf. Harvey, 2020), entrainment (Grahn et al.), and emotion (Gingras; Juslin). We do not have space to address each of these proposals in detail, but we are delighted our proposals have stimulated such productive extensions and we look forward to seeing the results of their proposals.

### R6. Conclusion: Understanding the value of music

Why has the evolution of musicality elicited such vigorous interdisciplinary debate? Harrison & Seale; Iyer; Margulis; Pfordresher; and Pinker; all mentioned the underlying role that evolutionary theory plays in value judgments about music (and the arts, more generally). Value judgments have dogged music precisely because, as Darwin observed, its practical survival value seems so "mysterious." This leaves funding for teaching and performing music often the first to be cut. It also results in drives by supporters to find evidence for practical, quantifiable values for music, such as benefits of music on individual health or intelligence (Biancolli, 2021). However, such efforts can sometimes be overzealous or counter-productive, as in the infamously debunked "Mozart effect" (Mehr, Schachner, Katz, & Spelke, 2013; Thompson, Schellenberg, & Husain, 2001).



We suggest that the social bonding hypothesis provides a promising framework for scientific investigation of the value of music more in terms of its social benefits, rather than individual ones. As Schellenberg put it, music is “the thing that brings people together and creates social bonding and makes us feel fantastic....If that’s not enough, then I don’t know what is” (Leung, 2019). We are excited by the constructive proposals of commentators to explore these questions, and hope that our hypothesis stimulates collection of additional data to help us better understand why the authors of our epigraphs all agree on the power of music to bring people together.

**Acknowledgments.** We thank Sam Passmore for writing the script to create Figure R1. We thank Sam Passmore, Aniruddh Patel, Peter Harrison, Dor Shilton, Jonathan de Souza, Jessica Grahm, and the other members of the University of Western Ontario music cognition reading group for comments on earlier versions of this manuscript. We also wish to acknowledge the important contributions to this topic by Bruno Nettl (1930–2020) and Iain Morley (1975–2021).

**Financial support.** PES was supported by Grant-in-Aid no. 19KK0064 from the Japan Society for the Promotion of Science and startup grants from Keio University (Keio Global Research Institute, Keio Research Institute at SFC, and Keio Gijuku Academic Development Fund). PL was supported by the National Science Foundation NSF-STTR no. 1720698, NSF-CAREER no. 1945436, NSF-STTR no. 2014870, the Grammy Foundation, and startup funds from Northeastern University. BT was supported by funding from the French Agence Nationale de la Recherche (under the Investissement d’Avenir program, ANR-17-EURE-0010) while on a Visiting Fellowship at the Institute of Advanced Study Toulouse. AS was supported by the National Science Foundation under NSF-BCS no. 1749551. WTF was supported by Austrian Science Fund (FWF) DK Grant “Cognition & Communication” (W1262-B29).

**Conflict of interest.** None.

## Notes

1. Mehr et al.’s primary arguments against the social bonding hypothesis were that: (1) “A ‘stress-reducing’ social bonding mechanism is superfluous,” (2) “The social bonding hypothesis conflates proximate- and ultimate-level reasoning,” and (3) “Music is poorly designed to coordinate groups.”
2. Note that this male predominance (56 songs sung by only males vs. 44 sung by only females in Mehr et al.’s Discography; 1,152 vs. 751, respectively, in their Ethnography) would be even stronger if Mehr et al. included instrumental music in addition to vocal songs (biases toward male performance are much stronger for instrumental performance than for singing; Savage et al., 2015). The male bias would also be stronger if Mehr et al. sampled lullabies (which are predominantly sung by women) for their Discography at rates comparable to the rates they appeared in their Ethnography (i.e., ~7% [89/1,273 song texts coded for function] lullabies found in their Ethnography vs. 25% lullabies sampled in their Discography).
3. Such restrictions may also extend to the process of documenting performance, e.g., male ethnographers may be prevented from documenting music performed by females. However, male biases were also found even for music recorded by female ethnographers (Savage et al., 2015).

## References

- Bergman, T. J., & Beehner, J. C. (2015). Measuring social complexity. *Animal Behaviour*, 103, 203–209.
- Bergstrom, C. T., & Lachmann, M. (1998). Signaling among relatives. III. Talk is cheap. *Proceedings of the National Academy of Sciences*, 95(9), 5100–5105.
- Biancolli, A. (2021). Music aids mental health: Science shows why. *Mad in America*. <https://www.madinamerica.com/2021/01/music-aids-mental-health-science-shows-why/>
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.
- Brown, G. R., & Richerson, P. J. (2014). Applying evolutionary theory to human behaviour: Past differences and current debates. *Journal of Bioeconomics*, 16, 105–128.
- Browning, R. (1871). *Balaustion’s adventure: Including a transcript from Euripides*. Smith, Elder and Co.
- Burley, N. T., & Kristine, J. (2002). The evolution of avian parental care. *Philosophical Transactions of the Royal Society B*, 357, 241–250.
- Cousineau, M., Oxenham, A. J., & Peretz, I. (2015). Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia*, 66, 293–301.
- Darwin, C. (1859). *On the origin of species*. John Murray.
- Darwin, C. (1871). *The descent of man, and selection in relation to sex*. John Murray.
- Degen, R. [DegenRolf]. (2020). Great article. It seems you are having Pinker’s cheesecake and eating it too. [Tweet]. Twitter. <https://twitter.com/DegenRolf/status/1290191006768304130>
- D’Errico, F., Villa, P., Llovera, A. C. P., & Idarraga, R. R. (1998). A middle palaeolithic origin of music? Using cave-bear bone accumulations to assess the Divje Babe I bone “flute.” *Antiquity*, 72, 65–76.
- Eldakar, O. T., & Wilson, D. S. (2011). Eight criticisms not to make about group selection. *Evolution*, 65, 1523–1526.
- Fitch, W. T., & Hauser, M. D. (2002). Unpacking “honesty”: Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, A. N. Popper & R. R. Fay (Eds.), *Acoustic communication* (pp. 65–137). Springer.
- Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100(1), 173–215.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge University Press.
- Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying “tone deafness.” *Brain*, 127(4), 801–810.
- Freeberg, T. M. (2006). Social complexity can drive vocal complexity: Group size influences vocal information in Carolina chickadees. *Psychological Science*, 17, 557–561.
- Gould, S. J. (1991). Exaptation: A crucial tool for evolutionary psychology. *Journal of Social Issues*, 47, 43–65.
- Gould, S. J., & Vrba, E. S. (1982). Exaptation – a missing term in the science of form. *Paleobiology*, 8, 4–15.
- Greenfield, M. D. (2005). Mechanisms and evolution of communal sexual displays in arthropods and anurans. *Advances in the Study of Behavior*, 35, 1–62.
- Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, 16, 48–55.
- Harvey, A. R. (2020). Links between the neurobiology of oxytocin and human musicality. *Frontiers in Human Neuroscience*, 14, 1–19. <https://doi.org/10.3389/fnhum.2020.00350>
- Hesse, H. (1943). *Das Glasperlenspiel*. Suhrkamp.
- Hoeschele, M., Merchant, H., Kikuchi, Y., Hattori, Y., & ten Cate, C. (2018). Searching for the origins of musicality across species. In H. Honing (Ed.), *The origins of musicality* (pp. 149–170). MIT Press.
- Honing, H. (Ed.). (2018). *The origins of musicality*. MIT Press.
- Honing, H., Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20140088. <http://doi.org/10.1098/rstb.2014.0088>
- Juslin, P.N. (2019). *Musical emotions explained: Unlocking the secrets of musical affect*. Oxford University Press.
- Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften*, 93, 451–454.
- Kunej, D., & Turk, I. (2000). New perspectives on the beginnings of music: Archaeological and musicological analysis of a middle Paleolithic bone “flute.” In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 235–268). The MIT Press.
- Leung, W. (2019). B.C. students who took music classes scored higher than peers in math, science and English: study. *The Globe and Mail*. <https://www.theglobeandmail.com/canada/article-bc-students-who-took-music-classes-scored-higher-than-peers-in-math/>
- Lomax, A. (Ed.). (1968). *Folk song style and culture*. American Association for the Advancement of Science.
- Lorenz, K. (1971). Comparative studies on the behaviour of the Anatinae. In K. Lorenz (Ed.), *Studies in animal and human behaviour vol. II* (pp. 14–114). Harvard University Press.
- Lukas, D., & Clutton-Brock, T. H. (2013). The evolution of social monogamy in mammals. *Science (New York, N.Y.)*, 341, 526–530.
- Marler, P. (1955). Characteristics of some animal calls. *Nature*, 176, 6–7.
- Marler, P., & Peters, S. (1982). Subsong and plastic song: Their role in the vocal learning process. In D. E. Kroodsma, E. H. Miller & H. Ouellet (Eds.), *Acoustic communication in birds, Vol. 2: Song learning and its consequences* (pp. 25–50). Academic Press.
- Maynard Smith, J. (1976). Sexual selection and the handicap principle. *Journal of Theoretical Biology*, 57, 239–242.
- Maynard Smith, J., & Harper, D. (2003). *Animal signals*. Oxford University Press.
- McGregor, P. K., & Dabelsteen, T. (1996). Communication networks. In D. E. Kroodsma & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds* (pp. 409–425). Cornell University Press.
- Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE*, 8(12), e82007. <https://doi.org/10.1371/journal.pone.0082007>