Behavioral and Brain Sciences

cambridge.org/bbs

Music as a coevolved system for social bonding

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

[a]Faculty of Environment and Information Studies, Keio University, Fujisawa 252-0882, Japan; [b]College of Arts, Media and Design, Northeastern University, Boston, MA 02115, USA; [c]Institute of Cognitive and Evolutionary Anthropology & Department of Experimental Psychology, University of Oxford, Oxford OX2 6PN, UK; [d]Department of Psychology, University of California San Diego, La Jolla, CA 92093, USA; [e]Department of Anthropology, Boston University, Boston, MA 02215, USA; [f]Department of Archaeology, University of Reading, Reading RG6 6AB, UK and [g]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; bronwyntarr01@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/

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.

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 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

CAMBRIDGE
UNIVERSITY PRESS

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.[1] 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,"[2] 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

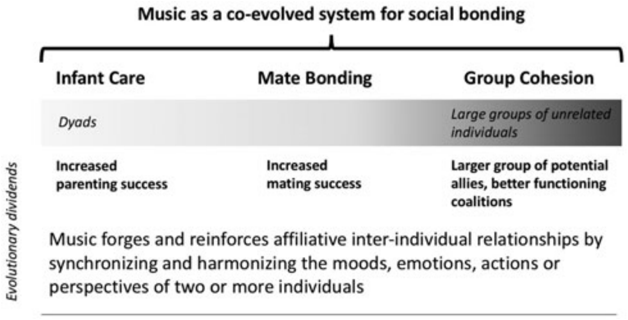


Figure 1. We propose that supposedly competing hypotheses for the evolution of human music, including mate bonding, infant care, and group cohesion (within both small coalitions and larger groups), are complementary sub-components of a broader social bonding function.

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 (Kroodsma, 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 interrelationships 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

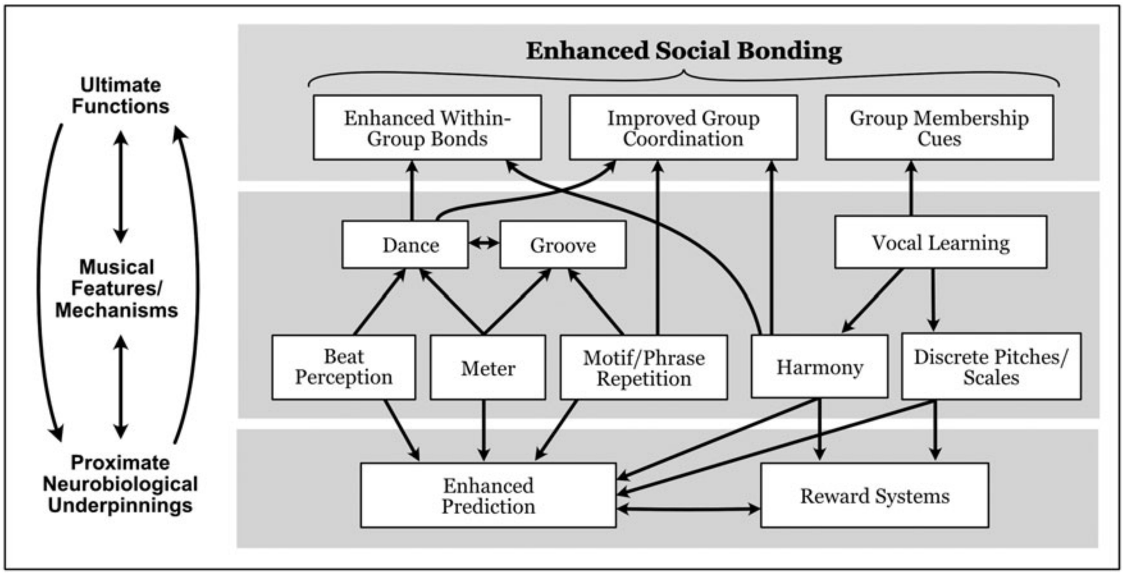


Figure 2. Proposed coevolutionary relationships among multiple musical features and mechanisms, indicating their contributions to ultimate functions by facilitating social bonding in multiple ways, their proximate neurobiological underpinnings in prediction and reward systems, and feedback loops among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

Functional context (from Mehr et al., 2019)	
(1) Dance	(15) War[a]
(2) Infancy	(16) Praise[a]
(3) Healing	(17) Love[a]
(4) Religious activity	(18) Group bonding[a]
(5) Play	(19) Marriage/weddings[a]
(6) Procession	(20) Art/creation[a]
(7) Mourning	
(8) Ritual	
(9) Entertainment	
(10) Children	
(11) Mood/emotions	
(12) Work	
(13) Storytelling	
(14) Greeting visitors	

Musical structure (from Savage et al., 2015)	
(1) Group performance	(15) Voice use
(2) Isochronous beat	(16) Modal register (chest voice)
(3) Metric hierarchy	*(17) Word use*
(4) 2- or 3-beat subdivisions	(18) Male performers
(5) 2-beat subdivisions	(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) ≤7 scale degrees	
(10) Unequal scales	
(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.

[a] 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, & Therriault, 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 (≤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 Feeuw, 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. playsong; 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

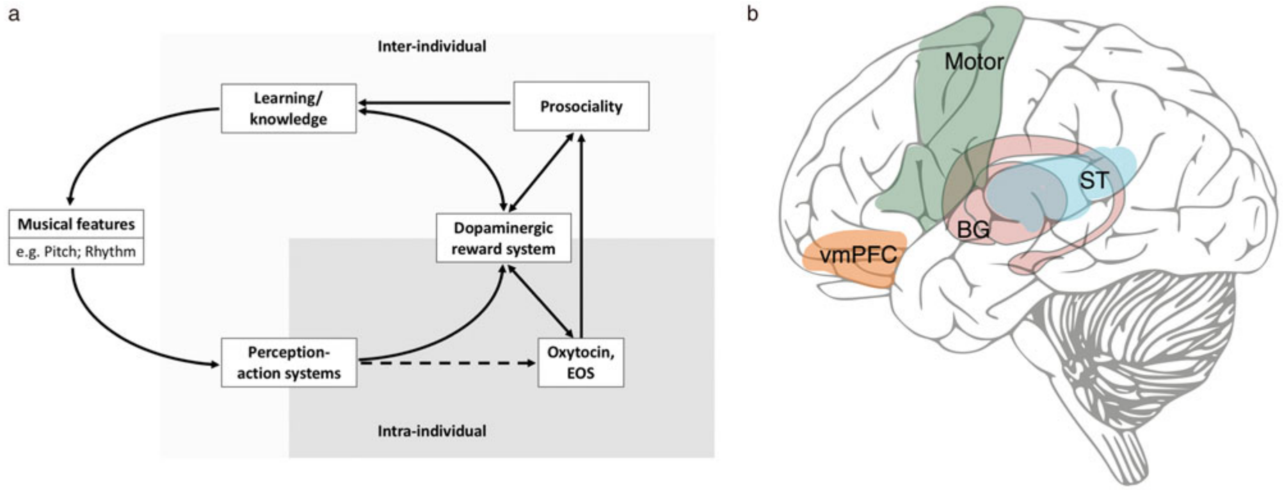


Figure 3. (a) Proposed neurobiological mechanisms underlying music's social bonding functions, showing intra- and inter-individual levels. We propose that the dopaminergic reward system interacts with the endogenous opioid system (EOS) and the release of oxytocin, ultimately providing opportunities for 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, Joanisse, & Grahn, 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, Rodriguez-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[3] (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; Hurlemann 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; Syal & 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).[4] 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, Grahn, 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).[5]

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).[6] 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, Rzeszutek, & 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.
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.mcg.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. http://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., & Therriault, 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/mto.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.

Feekes, F. (1982). Song mimesis within colonies of *Cacicus C. cela* (Icteridae: Aves): A colonial password? *Zeitschrift Tierpsychology*, *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. http://doi.org/10.3389/fpsyg.2011.00156.

Hanslick, E. (1858). *Vom Musikalisch-Schönen*. Ruoldf 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. http://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. Wigmore Hall Podcasts. https://wigmore-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., Mayseless, 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., & Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment – if the rhythm comes easily. *PLoS ONE*, *6*(11), e27272. http://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.

Kroodsma, 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., Pfordresher, 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. http://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., & Przysinda, 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.

Maestripieri, 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., Rodriguez-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., & Tlusty, 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., Grahn, 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., Schirrmacher, 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, southerneastern 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.

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/nxtbg.

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., Rzeszutek, 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/g6qxh.

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.

Tchalova, 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. 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., Joanisse, M. F., & Grahn, 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 spectacled parrotlets, *Forpus conspicillatus*. *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., Háden, 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/.

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.

Origins of music in credible signaling

Samuel A. Mehr[a,b,c]*, Max M. Krasnow[a]*, Gregory A. Bryant[d,e]* and Edward H. Hagen[f,]*

[a]Department of Psychology, Harvard University, Cambridge, MA 02138, USA; [b]Data Science Initiative, Harvard University, Cambridge, MA 02138; [c]School of Psychology, Victoria University of Wellington, Wellington 6012, New Zealand; [d]Department of Communication, University of California Los Angeles, Los Angeles, CA 90095, USA; [e]Center for Behavior, Evolution, & Culture, University of California Los Angeles, Los Angeles, CA 90095 and [f]Department of Anthropology, Washington State University, Vancouver, WA 98686, USA.
sam@wjh.harvard.edu; https://mehr.cz
krasnow@fas.harvard.edu; https://projects.iq.harvard.edu/epl
gabryant@ucla.edu; https://gabryant.bol.ucla.edu
edhagen@wsu.edu; https://anthro.vancouver.wsu.edu/people/hagen

doi:10.1017/S0140525X20000345, e60

Abstract

Music comprises a diverse category of cognitive phenomena that likely represent both the effects of psychological adaptations that are specific to music (e.g., rhythmic entrainment) and the effects of adaptations for non-musical functions (e.g., auditory scene analysis). How did music evolve? Here, we show that prevailing views on the evolution of music – that music is a byproduct of other evolved faculties, evolved for social bonding, or evolved to signal mate quality – are incomplete or wrong. We argue instead that music evolved as a credible signal in at least two contexts: coalitional interactions and infant care. Specifically, we propose that (1) the production and reception of coordinated, entrained rhythmic displays is a co-evolved system for credibly signaling coalition strength, size, and coordination ability; and (2) the production and reception of infant-directed song is a co-evolved system for credibly signaling parental attention to secondarily altricial infants. These proposals, supported by interdisciplinary evidence, suggest that basic features of music, such as melody and rhythm, result from adaptations in the proper domain of human music. The adaptations provide a foundation for the cultural evolution of music in its actual domain, yielding the diversity of musical forms and musical behaviors found worldwide.

1. Introduction

Thirty years ago, Steven Pinker and Paul Bloom made the "incredibly boring" (1990, p. 708) argument that language is the product of natural selection, resulting from adaptations for communication. This was, in fact, controversial: despite the facts that language is universally used to communicate information essential to survival and reproduction; that all people typically acquire language easily in infancy; that languages have deep computational structure unrelated to technological or societal progress; that neural injuries cause specific language impairments; and that specialized neuroanatomy enables speech production – many believed that language arose from byproducts of adaptations for cognition, not communication (e.g., Chomsky, 1968). The question of how language evolved is far from settled but it continues to generate testable hypotheses and productive results (e.g., Atkinson, Meade, Venditti, Greenhill, & Pagel, 2008; Christiansen & Chater, 2008; Fitch, 2017; Searcy, 2019).

Music shares many of the above facts with language but its contributions to survival and reproduction, if any, are less evident than those of language. As such, there is no consensus surrounding why humans make and listen to music; why music has its particular features and not others; or how music evolved. Three views on the evolution of music are prominent: a byproduct view, where music developed as a result of non-musical adaptations; an adaptationist view, where music evolved to create and maintain social bonds; and a second adaptationist view, where music evolved to signal mate quality.

We will argue that these views are incomplete or incorrect, proposing instead that the human psychology of music is built on adaptations for at least two categories of vocal signals common across species: territorial advertisements and contact calls. In these contexts, music can communicate overt information about covert properties of the human mind, functioning as a *credible signal*. This account explains some basic musical phenomena and the limited scope of music's proper domain, laying a foundation for cultural-evolutionary processes that shape the diversity of music worldwide.

2. What constitutes evidence for adaptation by natural selection?

Since antiquity, it has been recognized that unlike abiotic natural phenomena, the existence and form of many biological traits must be explained in reference to their "purpose." Rain does not fall in order to make corn grow, Aristotle wrote, but of necessity: "What is drawn up must cool, and what has been cooled must become water and descend, the result of this being that the corn grows" (Physics II, part 8). Teeth, in contrast, are "admirably constructed for their general office, the front ones being sharp, so as to cut the food into bits, and the hinder ones broad and flat, so as to grind it to a pulp" (Parts of Animals III, part 1). Human teeth universally grow this way, so this relation of means to ends cannot be due to chance, Aristotle argued; instead, these parts of animals can only be explained by their purpose, which benefits the animal itself (Ariew, 2002).

Two millennia later, William Paley described the organism as an intricate machine, "a cluster of contrivances" whose physical structures are best comprehended in relation to the useful functions they provide the organism (Paley, 1803, p. 185). He recognized that these contrivances must be understood in relation to their environments: "Can it be doubted, whether the wings of birds bear a relation to air, and the fins of fish to water?" (Paley, 1803, p. 291). Whereas Paley, arguing by exclusion, took evidence of design to be evidence for God, Darwin instead proposed that design evolved via heritable variation and differential reproduction, that is, adaptation by natural selection (Darwin, 1859).

Hypotheses for adaptation can be evaluated using criteria not so different from those of Aristotle, Darwin, or even Paley. Adaptations are generally characteristics of an entire species recognizable from a tight relation of means to ends: a fit between the features of the proposed adaptation and the features of the adaptive problem that it putatively solved. This constitutes evidence of *design* (Williams, 1966).

The human heart must be an adaptation to pump blood, for example, because it develops universally with properties that efficiently and reliably cause blood to circulate (e.g., muscles that compress chambers; valves; inlet and outlet ports; connections

---

*All authors contributed to this paper and are listed in order of reverse seniority.

to the circulatory system), an outcome essential to survival and reproduction. Conversely, a pumping function best explains why the heart has the structure it has, instead of other tissues in other arrangements, reliably and efficiently solving an adaptive problem (Darwin, 1859; Williams, 1966).

There are important differences between pre- and post-Darwinian conceptions of design, however. Selection among heritable variants generally optimized traits to increase *inclusive fitness*, the reproduction of self or close relatives (Hamilton, 1964), contra, for example, "well-being" or "longevity"; and did so in ancestral environments but not necessarily modern ones. In Williams's words, "the degree to which an organism actually achieves reproductive survival" is "rather trivial… The central biological problem is not survival as such, but design for survival" (Williams, 1966, p. 159).

A key issue when investigating the evolution of a trait – one central to questions of the evolution of music, as we will discuss – is the distinction between *proximate* and *ultimate*-level explanations (Mayr, 1961; Tinbergen, 1963). Proximate-level questions ask how a trait develops over ontogeny and what causal relationships it has with other parts of the organism. Ultimate-level questions, on the other hand, ask why a trait came to be and require identifying the phylogenetic history of the trait across ancestral and extant species, and the causal role it played, if any, in the reproduction of genes coding for it (discussion: Dickins & Barton, 2013; Laland, Sterelny, Odling-Smee, Hoppitt, & Uller, 2011; Scott-Phillips, Dickins, & West, 2011).

A proximate-level explanation for bitter taste, for instance, is that certain chemicals bind to bitter taste receptor proteins on the tongue, increasing intracellular calcium in the taste receptor cell, thereby stimulating a sensory afferent neuron, and so on. An ultimate-level explanation accounts for the presence of bitter taste receptor genes across vertebrates, and their expression in the oral cavity and other tissues, as part of a neurophysiological system to detect and avoid dietary toxins, which, if ingested, could reduce inclusive fitness (Roper & Chaudhari, 2017). Proximate-level explanation can also be applied to dysfunctions (such as cancer) and non-functions (such as the beating sound of the heart); they do not imply that a trait is an adaptation.

Ultimate-level analyses also do not presuppose adaptation. Adaptationist claims are onerous; there are infinitely many ways a phenotype can be carved into traits, most of which are unrelated to a genetic lineage's reproductive fitness. Supporting a claim of adaptation therefore requires evidence for design: evidence that a trait is improbably well-organized to efficiently, effectively, and reliably solve an adaptive problem (Williams, 1966).

Here we evaluate claims about the evolution of music using the approach outlined above, with particular attention to the psychological design of music.

Samuel A. Mehr is the Principal Investigator at The Music Lab, Department of Psychology, Harvard University. He studies how and why the design of the human mind leads us to perceive, create, and engage with music across human societies and across the lifespan. Originally a musician, Sam earned a B.M. in Music Education from the Eastman School of Music, followed by a doctorate in Human Development from Harvard. You can participate in Sam's research at https://themusiclab.org and follow him on Twitter @samuelmehr.

Max M. Krasnow is an Associate Professor of Psychology at Harvard University where he directs the Evolutionary Psychology Laboratory. He earned his Ph.D. in Evolutionary & Developmental Psychology at the University of California, Santa Barbara, in 2010. His research largely focuses on social evolution, using the tools of evolutionary and cognitive psychology to reveal the design of the mechanisms underlying human sociality such as our propensities for trust and generosity, yet also aggression and moral outrage. Other research lines include the evolution of music in humans, spatial adaptations for gathering, and the development of tool reasoning in children.

Gregory A. Bryant is a Professor of Communication at the University of California, Los Angeles. He received his Ph.D. in Cognitive Psychology from UC Santa Cruz in 2004 and completed a postdoctoral Fellowship in Psychology and Biological Anthropology at UCLA. He has published articles on vocal communication across cultures, psycholinguistics, bioacoustics, and music perception.

Edward H. Hagen, Professor of Anthropology at Washington State University, has over 70 publications on the evolution of music, drug use, mental health, leadership, and cognition.

3. Two claims regarding the origins of music are unconvincing

A successful account of music must provide evidence for the design of its principal features. Music is an auditory display built from melodies and rhythms. It can involve loud, elaborate, coordinated performances with voices and musical instruments, with many listeners; it also can involve quiet, simple lullabies between parents and infants. It appears in many behavioral contexts, across the sexes and across the lifespan, as a common element of daily life.

Before we proceed, readers should note a companion *BBS* target article, "Music as a coevolved system for social bonding," by Savage et al., which presents an alternate evolutionary scenario for the origins of music. Savage and colleagues propose that musicality arose fairly recently in human ancestry as a cultural invention to enhance social bonding, and was then elaborated via gene–culture evolution over tens of thousands of years (see also Podlipniak, 2017).

We appreciate the focus on gene–culture co-evolution, a phenomenon we do not explore in detail in this paper, but which dovetails nicely with our concluding ideas concerning cultural evolution. Nevertheless, the two approaches differ substantially. The theoretical justification for music as a social bonding mechanism relies primarily on the work of Dunbar and colleagues, who argued that grooming serves this function in smaller groups of non-human primates, but that larger human groups required more efficient mechanisms, namely laughter and music. On this idea, social bonds are created by the effects of joint musical performances on the neurobiology of the performers, rather than from information encoded in music. The costs of music production do not enter into this account, and Savage et al. (2020) mostly avoid theoretical or phylogenetic connections between human musicality and similar phenomena in other species (though they do offer some predictions concerning musicality in other species).

The theory we will describe differs substantially from this view. We propose that music has deep evolutionary roots in primate vocalizations, especially contact calls and territorial advertisements that were likely present in the last common ancestor of all primates, approximately 55–85 million years ago. We see music as a credible signal conveying information to listeners with whom signalers might have conflicts of interest, in a fashion similar to most work on non-human vocalizations. We draw theoretical and phylogenetic connections between human music and similar phenomena in other primate and non-primate species. We argue that unique aspects of human lifestyle, including multilevel social organization and high levels of parental investment (including from alloparents), selected for especially elaborate vocal signaling relative to most other species. Finally, we propose that the key features of musicality arising from adaptations in the proper domain of credible signaling serve as building blocks for cultural evolution, which shapes music into its actual domain.

To begin, we review two popular ideas about the origins of music, and ask whether they explain the core properties of music.

3.1 The byproduct hypothesis fails in light of six lines of evidence

The null hypothesis against which hypotheses for adaptation are tested claims that music has no evolved function, and instead is a byproduct of other adaptations that evolved for other functions unrelated to music. The *byproduct hypothesis* dates at least to William James, who wrote that music "is a pure incident of having a hearing organ" (James, 1890, p. 627); this view echoed other scholars of his time and before (Darwin, 1871; Monboddo, 1774; Rousseau, 1781; Spencer, 1902), and is common in the literature. Music has been proposed to be a byproduct of linguistic or emotive communication (Bryant, 2013; Cattell, 1891; Cross & Woodruff, 2009; Jackendoff, 2009; Panksepp, 2009; Patel, 2008; Pinker, 1997; Schulkin, 2013; Sievers, Polansky, Casey, & Wheatley, 2013); auditory scene analysis and habitat selection (Pinker, 1997; Trainor, 2015); signaling vocalizations (Bryant, 2013; Livingstone, 1973; Mithen, 2005; Pinker, 1997; Richman, 1993); mimicry of other animals' vocalizations (Benzon, 2001; Krause, 2012); physical or motor abilities (Geist, 1978; Larsson, 2014; Panksepp, 2009; Tierney, Russo, & Patel, 2011); theory of mind (Livingstone & Thompson, 2009); or general cognitive capacities (Cross, 2012; Honing & Ploeger, 2012; Jackendoff & Lerdahl, 2006; Justus & Hutsler, 2005; Marcus, 2012).

Pinker's (1997) framing is the best-known: "I suspect that music is auditory cheesecake, an exquisite confection crafted to tickle the sensitive spots of … our mental faculties" (p. 534). Six lines of evidence, taken together, call the byproduct hypothesis into question, however, and motivate theories of specific adaptations for music.

First, complex, song-like vocalizations have evolved convergently across distantly-related animals, including multiple clades of birds, marine mammals, primates, and insects; and provide important benefits related to mating and territorial defense (Coen, Xie, Clemens, & Murthy, 2016). In many cases, these are socially learned, like music (Schachner, Brady, Pepperberg, & Hauser, 2009). Moreover, at least some explicitly musical behaviors, such as entrainment to a beat, appear in many species (Phillips-Silver, Aktipis, & Bryant, 2010; Wilson & Cook, 2016). Music-like adaptations can therefore evolve, in principle.[1]

Second, music is a human universal: it appears throughout a representative sample of human societies (Mehr et al., 2019); plays an essential role in important activities, such as rituals and ceremonies (Nettl, 2015); and demonstrates cross-cultural links between form and function (Bainbridge & Bertolo et al., 2021; Hilton & Crowley-de Thierry et al., 2021; Mehr et al., 2019, 2018; Trehub, Unyk, & Trainor, 1993a). Music is not a byproduct of traits present in only some cultures.

Third, music shows evidence for complex design, including grammar-like structures analogous to those of language (Lerdahl & Jackendoff, 1983), some of which may be universal (Jacoby et al., 2019; Mehr et al., 2019). Moreover, music perception is computationally complex, such that artificial intelligence is currently at pains to emulate it (Benetos, Dixon, Giannoulis, Kirchhoff, & Klapuri, 2013). Music is unlikely to occur as a result of random chance.

Fourth, the motivation and ability to perceive music appear early in ontogeny: neonates are sensitive to rhythms (Winkler, Háden, Ladinig, Sziller, & Honing, 2009) and melodies (Granier-Deferre, Bassereau, Ribeiro, Jacquet, & DeCasper, 2011) and infant music cognition is precocial (e.g., infants have detailed long-term memory for music; Mehr, Song, & Spelke, 2016; Mehr & Spelke, 2017; Trainor, Wu, & Tsang, 2004; reviews: Hannon & Trainor, 2007; Trehub, 2001). Music perception develops naturally, does not require extensive training, and is not a byproduct of traits specific to adults.

Fifth, music perception displays evidence for neural specialization (Norman-Haignere, Kanwisher, & McDermott, 2015, 2019) and is impaired in specific deficits, such as tone-deafness (Peretz, Ayotte, Zatorre, & Jutras, 2002; Peretz & Vuvan, 2017). Music is unlikely to be a byproduct of other neural systems.

Last, music is ancient: flutes are at least 40,000 years old (Conard, Malina, & Münzel, 2009) and the human auditory and vocal production systems are far older (Fitch, 2006; Martínez, Rosa, Arsuaga, & Carbonell, 2004; Quam et al., 2013). Music is not a recent cultural invention.

While no one of these pieces of evidence is a sufficient condition for rejecting the byproduct hypothesis, taken together, they motivate a search for an alternative.

3.2 The social bonding hypothesis fails in light of three theoretical issues

The best-known evolutionary hypothesis for music is that it evolved to create and maintain "social bonds." Roederer (1984), for example, argued that music established "behavioral coherency in masses of people" to meet the demands of "coherent, collective actions on the part of groups of human society" (p. 356). Brown (2000b) asserted that "music-making has all the hallmarks of a group adaptation and functions as a device for promoting group identity, coordination, action, cognition, and emotional expression" (p. 296). These and similar claims (Barrow, 2005; Benzon, 2001; Brown, 2000a; Conard et al., 2009; Cross & Morley, 2009; Dissanayake, 2000, 2008, 2009; Dunbar, 1998, 2012a; Freeman, 2000; Fritz, Hardikar, Demoucron, & Leman, 2013; Geissmann, 2000; Huron, 2001; Jourdain, 1997; Kirschner & Tomasello, 2009, 2010; Koelsch & Siebel, 2005; Kogan, 1994; Launay, Tarr, & Dunbar, 2016; Loersch & Arbuckle, 2013; McNeill, 1995; Merker, Madison, & Eckerdal, 2009; Morley, 2012; Pearce, Launay, & Dunbar, 2015; Reddish, Fischer, & Bulbulia, 2013; Richman, 1993; Schulkin, 2013; Schulkin & Raglan, 2014; Weinstein, Launay, Pearce, Dunbar, & Stewart, 2016; Wiltermuth & Heath, 2009) together form the *social bonding hypothesis*.

This view was popularized in part by Dunbar's proposal of a role for social bonding in the evolution of many human social traits (Dunbar, 1991): he argued that in primates, manual grooming serves a social bonding function; as group size increased in the hominin lineage, manual grooming became prohibitively time-consuming, creating a selection pressure for a less costly bonding mechanism; and, as a consequence, new bonding mechanisms evolved. These mechanisms were first proposed to be language and gossiping, which could be broadcast to multiple individuals while doing other tasks, replacing grooming as the primary means of social bonding in humans (Dunbar, 1998). Later, Dunbar and colleagues revised this position (Dunbar & Lehmann, 2013), arguing that musical chorusing and laughter evolved instead (Dunbar, 2004, 2012a; Dunbar, Kaskatis, MacDonald, & Barra, 2012; Pearce et al., 2015, 2016, 2017; Tarr, Launay, Cohen, & Dunbar, 2015; Tarr, Launay, & Dunbar, 2014; Tarr, Launay, & Dunbar, 2016).

Most empirical tests of the hypothesis examine music's impact on prosociality and its hormonal mediators in laboratory experiments: participants are randomized into groups that engage in synchronized musical behavior (treatment) or another activity (control). The general finding is greater levels of prosociality and cooperation in the music/dancing conditions relative to controls[2] (Anshel & Kipper, 1988; Cirelli, Einarson, & Trainor, 2014; Kirschner & Tomasello, 2009, 2010; Pearce et al., 2015, 2016, 2017; Reddish et al., 2013; Schellenberg et al., 2015; Tarr et al., 2015).

The social bonding hypothesis has at least three key issues, however.

3.2.1 A "stress-reducing" social bonding mechanism is superfluous

The ultimate-level problem of sociality is that it imposes difficult-to-overcome inclusive fitness costs: increased competition with conspecifics for essential, limited resources; inbreeding depression; and increased exposure to pathogens (Alexander, 1974). Living with and cooperating with conspecifics requires that the inclusive fitness benefits of sociality outweigh its fitness costs. Often they do not: dispersal and solitary living are ubiquitous across species (Benton, Baguette, Clobert, & Bullock, 2017; Bowler & Benton, 2005; Duputié & Massol, 2013).

In primates, diurnal social living evolved about 52 million years ago (Shultz, Opie, & Atkinson, 2011). Because diurnal foraging increases predation risk, the joint evolution of diurnality and sociality supports the long-standing idea that primate sociality evolved as a defense against predators (Silk & Kappeler, 2017; Van Schaik, 1983). Advocates of the social bonding hypothesis claim that social living creates psychological stresses that threaten the cohesion of the group, necessitating a "bonding mechanism" (in non-human primates, grooming; in humans, music) that reduces stress:

> Since living in groups of any kind creates stresses that would normally result in the group disbanding, species that live in stable social groups have to circumvent this problem if they are to prevent group size collapsing. (Dunbar, 2012a, p. 1838)

This idea is superficially appealing because it draws attention to the fitness costs of social living, presenting them as proximate-level stresses, and implying a need for a behavioral response to relieve the stress.

But an ultimate-level analysis must consider alternative strategies. The alternative to sociality is solitary living, seen in ~70% of mammal species (Wilson & Reeder, 2005). On the hypothesis that sociality solves the adaptive problem of defense against predators, the *net* fitness benefits of sociality exceed those of solitary life (with its attendant high risk of predation). The stress-related benefits of a "social bonding mechanism" are superfluous.

For an analogy, consider a group of friends walking close together in a dangerous neighborhood at night. There are costs to this sociality: they bump into each other; they don't fit on the sidewalk, forcing some to risk injury from oncoming cars; it's harder for them to converse, and so on. An ultimate-level analysis recognizes that the benefits of their sociality – defense against getting mugged – outweigh the costs, and no bonding mechanism, such as grooming or singing, is required to keep them together.

Grooming does provide hygienic benefits to primates, such as removal of ectoparasites (Barton, 1985); perhaps with social functions beyond hygiene (McKenna, 1978; Seyfarth, 1977; Seyfarth & Cheney, 1984), because across species the proportion of time spent grooming is positively correlated with group size (Dunbar, 1991). But this association, the core empirical finding underlying the social bonding hypothesis, is poorly evidenced: its strength is modest, and, when adjusting for terrestriality and other ecological factors, is not distinguishable from zero (Jaeggi, Kramer, Hames, & Gurven, 2017). This may be because primate group size is confounded with terrestriality; if so, increased grooming time could instead be explained by some property of a terrestrial niche, such as increased parasite load (Grueter, Bissonnette, Isler, & van Schaik, 2013; Jaeggi et al., 2017; cf. Dunbar & Lehmann, 2013).

Whatever evolved social functions grooming might have, it is unlikely that they include stress reduction. Predation risk and perhaps communal resource defense, not grooming, are the ultimate-level "bonding forces" that likely explain primate sociality (Port et al., 2020), and the additional benefits of cooperative endeavors such as hunting, parenting, and territorial defense likely explain human sociality.

3.2.2 The social bonding hypothesis conflates proximate- and ultimate-level reasoning

Might grooming solve other problems of sociality? Defense against predators, territory defense, hunting, and parenting are compelling examples of cooperation whose benefits could offset sociality's costs. They raise profound theoretical challenges, however, involving free-riders: agents that receive benefits from others but do not provide any. Without countermeasures, free-riding is favored by natural selection (Nowak, 2006), so forming cooperative relationships with arbitrary individuals is untenable. Instead, these relationships must be targeted at specific categories of individuals, such as kin, neighbors, or those likely to reciprocate, and adjusted to local socio-ecological conditions (Markham, Gesquiere, Alberts, & Altmann, 2015) such that long-run benefits are provided only when they exceed long-run costs to the donor (Nowak, 2006).

Proponents of the social bonding hypothesis offer a proximate-level explanation, wherein the neurohormonal effects of music are a solution to the impediments to sociality and cooperation described above:

> My proposal is that music arose originally because it allows individuals to become more group-oriented. Music seems to achieve this through a capacity to produce endorphins which have a positive effect on our attitudes towards others. (Dunbar, 2012b, p. 208)

> We propose that synchrony might act as direct means to encourage group cohesion by causing the release of neurohormones that influence social bonding. (Launay et al., 2016, p. 779)

There are two problems with these claims. First, evidence that *X* causes *Y* is weak evidence that *X* evolved to cause *Y*. Recall Aristotle: rain causes corn to grow without implying any "purpose" for rain. Rain shows little evidence of special design for solving corn's hydration problem, it has many other, unrelated effects, and so on. By analogy, a proximate-level analysis shows that petting animals reduces human anxiety via hormonal and physiological effects (Beetz, Uvnäs-Moberg, Julius, & Kotrschal, 2012), but animal-petting did not evolve to reduce anxiety or the threats that trigger it, of course.

Second, proximate mechanisms, such as release of neurohormones, are themselves subject to selection, and therefore cannot serve as ultimate-level explanations for the genetic evolution of a social bonding strategy. In order for a social strategy to evolve, it must outperform conceivable mutant strategies (a well-recognized criterion for claims of adaptation, the evolutionarily stable strategy; Smith & Price, 1973). A mutation that prevented music from increasing endorphins and/or reduced endorphins' effects on prosociality would have allowed humans with that mutation to free-ride: they could gain from the prosocial behavior of others (becoming more bonded with the group) without being prosocial themselves. Such a free-rider mutation would be selected for (Nowak, 2006).

How can an unconditional social bonding mechanism like music be stabilized against free-riders? Confusion between proximate- and ultimate-level analyses in the social bonding hypothesis leaves this question unanswered.

3.2.3 Music is poorly designed to coordinate groups

Another version of the social bonding hypothesis proposes that music evolved by genetic group selection to enable humans to act as coordinated superorganisms: music increased group fitness by promoting group identity, cognition, coordination, and catharsis. These within-group functions are proposed to increase the ability of groups to compete with other groups (Brown, 2000a).

While music does play a universal role in rituals (e.g., shamanistic trance; Mehr et al., 2019; Singh, 2018), the problem with this view is that it equates proximate social "functions" or "effects" with adaptations shaped by natural selection.[3] Because any behavior has effects, and some of those effects may be incidentally "useful" (e.g., animal-petting reduces anxiety), the proper criterion is that music be well-designed for the proposed within-group function.

The superorganism model is based on an explicit analogy with multicellular organisms, where energy and time are sharply constrained resources. Within-organism signaling, cognition, and coordination evolved to be as efficient as possible, to maximize between-organism competitiveness. In neural signaling, for example, time and energy trade off: higher information rates use more energy, so at all levels of neural organization, strategies evolved to reduce energy consumption by filtering out predictable inputs, reducing the amount of redundant encoding (Laughlin, 2001; Niven, 2016; Niven & Laughlin, 2008).

But music takes considerable time and energy to produce. People who produce music incur opportunity costs (Mehr & Krasnow, 2017) and expend energy that could be used for other activities that directly increase reproductive success, such as food production (Hagen & Bryant, 2003). Music is also often loud, and could attract predators or allow competing groups to eavesdrop. These costs also accrue to the variants of the social bonding hypothesis discussed earlier. Indeed, while music and other synchronous, ritualistic behaviors are often argued to be unambiguously beneficial for groups, the "neglected dark side of synchrony" (p. 3) shows that synchrony increases conformity and groupthink while reducing creativity and productive dissent (Gelfand, Caluori, Jackson, & Taylor, 2020).

Because natural selection shapes traits to perform specific functions by selecting among alternatives, a criterion for claiming adaptation is that a trait is uniquely suited to causing certain effects, relative to feasible alternatives. In the case of the social bonding hypothesis, an obvious alternative to music that serves the same proposed within-group functions is language, a low-cost signaling system that efficiently facilitates the coordination of collective action and other social behaviors (Pinker & Bloom, 1990). Consider that the coxswain, whose job is to maintain the coordination of rowers, does not sing, nor does the crew; the efficient vocalization "row!" minimizes the energy required for within-group coordination, while maximizing the rowers' ability to win a race.[4] Moreover, in a sample of six small-scale human societies, *conversation* time was close to the expected grooming time for a terrestrial primate with recent ape ancestry (Jaeggi et al., 2017), suggesting that language adequately provides whatever social functions grooming may have. As a social coordination or bonding mechanism, music thus appears to have no advantages over language and many disadvantages.

The weak case for music as an adaptation for social bonding does not mean that music has no evolved social functions. In the rest of this paper, we outline an alternative social hypothesis for the origins of music.

4. Origins of music in credible signaling

The social bonding hypothesis proposes that the fitness benefit of music arises from the neurophysiological effects of music production on music-makers themselves. Signaling hypotheses, in contrast, propose that fitness benefits arise from the information communicated by music-makers, via their music, to various categories of listeners.

Acoustic communication has evolved repeatedly and independently in many clades of tetrapods. It appeared 200 million years ago in therian mammals and is found in ~95% of mammal species (Chen & Wiens, 2020). If music is an adaptation, it likely evolved from ancestral vocalizations, an idea foreshadowed by Lucretius two millennia ago:

> To imitate the liquid notes of birds
> Was earlier far amongst men than power to make
> By measured song, melodious verse and give
> Delight to ears.
> (De Rerum Natura, Book V)[5]

In non-human animals, most vocal adaptations evolve to send *signals*, which are defined as "any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved" (Maynard Smith & Harper, 2003, p. 3).[6] On average, receivers benefit from responding to the signal, and signalers benefit from the receivers' response.[7] Cues, in contrast, convey information about one organism to another but did not evolve to do so (e.g., bleeding is a cue of injury but did not evolve to signal injury). Common functions of signals include species identification in mate choice; individual recognition in interactions among conspecifics, such as territoriality, dominance, and cooperation; and conveying information on formidability, health, or behavioral type (Tibbetts, Mullen, & Dale, 2017).

Why do animals believe the vocal signals they hear? What maintains their credibility? If the interests of signaler and receiver are aligned, as in cells in an organism or agents in a superorganism, then selection for dishonesty is absent and signals evolve to be as efficient as possible (Maynard Smith & Harper, 2003). If not, then selection can drive signalers to deceive and receivers to be vigilant against manipulation.

Some signals are necessarily credible because they are causally related to the quality being signaled. A wolf howl credibly indicates that a wolf is present, for example, and the number of distinct, simultaneous wolf howls credibly indicates a lower bound on the size of the pack (an "index"; Maynard Smith & Harper, 2003). *Costly* signals, in contrast, are credible because to send them imposes a fitness cost that is lower for individuals with the quality than those without it: faking the signal is more costly than it is worth[8] (Spence, 1973; Zahavi, 1975).

In addition to credibility, multiple selection pressures can shape signals, including biases in the sensory systems of receivers; receiver abilities to discriminate signals; the structure of the environment; social challenges; and arms races between signalers and receivers (Krebs & Dawkins, 1984), where signalers are selected to produce the signal at lower cost and receivers are selected to better discriminate the quality of signalers (Bradbury & Vehrencamp, 1998; Cummings & Endler, 2018; Hill, 1994; Lindsay, Andersson, Bererhi, & Edwards, 2019; McCoy & Haig, 2020; van Doorn & Weissing, 2006).

Here, we emphasize the importance of conflicts of interest between music producers and the audience, private information, and the features of music that underlie its ability to overtly signal covert information about the minds of those producing it.

4.1 The mate quality hypothesis is poorly supported

An early theory of music, first proposed by Darwin (1871) and endorsed by many others (Barrow, 2005; Charlton, 2014; Dutton, 2009; Merker, 2000a; Miller, 2000a, 2000b; Miranda, Kirby, & Todd, 2003; Orians, 2014; Sluming & Manning, 2000; Todd, 2000; Todd & Werner, 1999; van den Broek & Todd, 2009) is that male musical abilities and female musical preferences coevolved, with music functioning as a credible signal of male mate quality.

If musical production requires a brain and body relatively unperturbed by genetic mutation, infection, or developmental instability, plus time to cultivate one's talent (properties that are difficult to perceive directly), the *mate quality hypothesis* argues that mates who prefer music-producers will benefit. This increases selection for music-producers to generate more impressive, complex, or interesting music (so as to improve the chance of being chosen as a mate). Given the sex difference in the amount of investment required of human parents for an offspring to be reproductively viable (Trivers, 1972), signal production should be accentuated in the sex with lower obligate parental investment (males) and choosiness should be accentuated in the sex with the higher obligate parental investment (females).

Sexually dimorphic signals of mate quality are common across species (e.g., coloration, ornaments) and they play key roles in mate attraction (Andersson, 1994; Dale, Dey, Delhey, Kempenaers, & Valcu, 2015; O'Brien, Allen, Van Kleeck, & Emlen, 2018); for example, male birdsong functions in part to attract mates (Catchpole & Slater, 2018). Some mammals show this pattern too. In sac-winged bats, males produce complex songs that females may use in mating decisions (Behr et al., 2006). Adult house mice produce sexually dimorphic ultrasonic vocalizations with song-like features in response to the presence of novel female urine, but not the scents of immature females or other males (Musolf, Hoffmann, & Penn, 2010). If human music evolved in the context of signaling mate quality, it should have retained similar, signature features of a sexually selected adaptation.[9]

Ironically, the mate quality hypothesis is easy to refute precisely because it is so well-specified. Music is tenuously linked to mate quality. While love/courtship songs are common across cultures, they are only weakly identifiable as such relative to other forms of song (Mehr et al., 2019, 2018). A large twin study found inconsistent relations between measures of reproductive success or sociosexuality and measures of music production or music perception abilities (Mosing et al., 2015); the few positive relations reported were weak, and no stronger in men than in women. In another study, music performance quality was positively associated with indices of mate quality and attractiveness, but the effects and sample size were small and did not differ by sex (Madison, Holmquist, & Vestin, 2018). Musical preferences can vary across the menstrual cycle (Charlton, 2014), perhaps indicating a role in mating, but this effect has failed to replicate[10] (Charlton, Filippi, & Fitch, 2012).

Sexually-selected traits that function for display and choice in mating contexts are often developmentally and contextually calibrated to mating (Kokko, 1997). In humans, for instance, puberty coordinates the developmental timing of physical and psychological traits that support mating: menarche and spermarche coincide with the development of secondary sexual characteristics and the relative onset of mating psychological systems (Kaplan & Gangestad, 2005). But humans of all ages produce and listen to music; no part of the music faculty emerges at or around puberty.[11] Young children enjoy the music of sexually mature conspecifics, and vice versa, a pattern contrasting with that of sexual attraction, which begins in late childhood (Herdt & McClintock, 2000). While mating-related behaviors tend to be produced only in mating-relevant contexts, music is produced and consumed in multifarious contexts, universally, that are completely unrelated to mating (e.g., work, healing, greeting visitors, mourning; Mehr et al., 2019).

Last, many mating-related traits in humans are sexually dimorphic, such as male biases for traits useful in physical competition (Puts, 2010; Sell et al., 2008) and female biases for traits useful in discerning investment potential (Buss, 1989; Conroy-Beam, Buss, Pham, & Shackelford, 2015; Kenrick & Keefe, 1992). If music evolved to signal mate quality, then adaptations for music production should be more developed in men and adaptations for music perception should be more developed in women.

Little evidence supports this pattern. Dimorphisms in human vocalizations and vocal anatomy – lower voices in males, signaling threat potential (Puts, Apicella, & Cárdenas, 2011), and higher voices in females, signaling fecundity (Apicella & Feinberg, 2009) – appear beginning at puberty (McDermott, 2012), but are neither more exaggerated nor more honestly signaled via song rather than via speech (cf. Keller, König, & Novembre, 2017). Auditory perception skills are comparable in males and females, with only small and inconsistent sex differences (Müllensiefen, Gingras, Musil, & Stewart, 2014; Shuter-Dyson & Gabriel, 1981). Musical disorders, such as specific musical anhedonia and congenital amusia, are found just as frequently in males as in females (Mas-Herrero, Zatorre, Rodriguez-Fornells, & Marco-Pallarés, 2014; Peretz & Vuvan, 2017). A lone report of sex differences in the frequency of music performance across human societies (Savage, Brown, Sakai, & Currie, 2015) is likely the result of sampling bias (for discussion, see Mehr et al. 2018, 2019). If anything, female musicians produce more novel songs than their male peers (Askin, Mauskapf, Koppman, & Uzzi, 2020).

The pervasiveness of music across the sexes is evident in daily life: both males and females seek out and enjoy the performances of both male and female musicians (Hagen & Bryant, 2003), and some evidence suggests that musical preferences are biased toward performers of the same sex as the listener (Greenberg, Matz, Schwartz, & Fricke, 2020). Male and female performers are both well-represented, historically, on the Billboard Top 100, albeit with an advantage toward males (Lafrance, Worcester, & Burns, 2011). While many of the highest-grossing musical artists of all time are male, sex differences in success as a musician likely have little to do with biology – a half-century ago, virtually all professional orchestral musicians were male, for example, whereas now the world's top orchestras are approaching gender parity (Sergeant & Himonides, 2019).

This pattern of evidence has contributed to a growing consensus that links between music and mate quality are weak (Mosing et al., 2015; Ravignani, 2018).

4.2 Music as a credible signal of cooperative intent

We agree with proponents of the mate quality hypothesis that music is a credible signal. But song-like vocalizations in nonhuman animals often signal much more than mate quality. Even in songbirds, the poster-species for the sexual selection of male song, singing can serve other functions, such as territorial advertisements (Tobias et al., 2016).

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. Cooperation often fails, making it useful to have a credible signal indicating that, by various (non-musical) means, one or more agents have decided to cooperate. Credible signals of cooperative intent, in turn, can produce decisions by signal receivers that benefit the signalers.

We will discuss two behavioral contexts where complex vocal signals have evolved in numerous other species; where unique characteristics of the human species created selection pressures for an elaborate credible signal; and where music universally appears.

First, in the context of territorial advertisements, we consider pressures of coordinated territorial defense across coalitions and in the context of cooperative alliances with other groups. We propose that music could function as a credible signal of coalition strength, size, and coordination ability.

Second, in the context of contact calls, we consider pressures of helpless infants requiring substantial parental investment, relative to other primates; and multiple dependent siblings competing for parental investment. We propose that music could function as a credible signal of parental attention.

4.2.1 Synchronous coordinated music as a credible signal of coalition strength, size, and cooperation ability

In mammals, loud auditory signals are frequently agonistic, and territorial advertisements are a prime example (Gustison & Townsend, 2015). Territoriality is common in taxa ranging from bacteria to vertebrates (Maher & Lott, 2000; Smith & Dworkin, 1994), including primates (Willems & van Schaik, 2015). Territory owners have a consistent advantage over intruders, often retaining their territory without a fight (Kokko, Jennions, & Brooks, 2006). It is thus in the interest of owners to advertise their residence in a territory to deter intruders and avoid a fight.

Territorial calls, which credibly signal that a territory is occupied, are found in many species, including birds, primates, and other mammals (Bates, 1970; Gustison & Townsend, 2015; Ladich & Winkler, 2017; Wich & Nunn, 2002). Loud primate calls are a plausible evolutionary precursor to human music (Geissmann, 2000) because they appear to have existed in the last common ancestor of all primates and are often produced by both sexes and directed at both sexes (Wich & Nunn, 2002). Some African apes display drumming-like behaviors as part of territorial signals (Goodall, 1986; Hagen & Hammerstein, 2009). In humans, vocal and instrumental music are reliably associated with war, procession, and ritual across a representative sample of societies (Mehr et al., 2019, Table 1); appears in political and military contexts with analogs to territorial signaling (Hagen & Bryant, 2003; Hagen & Hammerstein, 2009); is generally not sexually differentiated (see above); and, of course, is often loud.

Social species that collectively defend territory, such as chimpanzees and several species of social carnivores (e.g., lions, wolves), produce coordinated vocal territorial advertisements (e.g., roars, howls), which credibly signal group size to potential intruders (Harrington, 1989; Harrington & Mech, 1979; Krebs, 1977; McComb, Packer, & Pusey, 1994; Wilson, Hauser, & Wrangham, 2001). In a study of nearly 10,000 bird species, the presence of communal signaling was associated with territoriality, typically in conjunction with stable social bonds (Tobias et al., 2016). Moreover, the effect of territoriality was more than twice the size of that of social bonds, and territoriality was a crucial precursor to communal signaling, suggesting that long-term social bonds might evolve after communal signaling.

Some coordinated vocal signals, like bird duets, involve complex, temporally synchronized displays. A high level of synchronous coordination among signalers requires considerable effort to achieve, and thus credibly signals a willingness and ability to cooperate over time, thereby serving as an index of the quality of the coalition defending the territory, above and beyond coalition size (critical information otherwise not apparent to intruders; Hagen & Bryant, 2003; Hall & Magrath, 2007; Wiley & Wiley, 1977). If synchronous coordination is a signal of coalition quality, selection should push receivers to better discriminate differences in degrees of coordination, and signalers to produce more complex coordinated signals, leading to signal elaboration.

Several primate species also produce highly synchronized song-like duets. As in birds, song-like calls are characteristic of species living in small, monogamous groups (Schruth, Templeton, & Holman, 2019). Although duetting and coordinated vocalizations might have some role in pair-bond formation and strengthening in a few monogamous species, such as gibbons and titi monkeys, most evidence suggests these calls primarily function to exclude intruders and maintain spacing: they are territorial advertisements (Snowdon, 2017).

Experimental evidence suggests that higher levels of coordination in such signals indicate higher coalition quality. Duetting magpie-larks that had been paired for a longer time were more likely to produce highly coordinated displays, and in an experimental loudspeaker study on natural territories, playbacks of highly coordinated duets, which simulated territorial intrusions, evoked significantly higher song rates by resident males than poorly coordinated duets (indicating that the highly coordinated duets were perceived as more threatening; Hall & Magrath, 2007).

Humans are both primates and social hunters, so we expect human ancestors to have advertised territory ownership in a similar fashion: using loud, coordinated vocalizations, perhaps with drumming. We propose that such territorial vocalizations are an evolutionary precursor to music, especially rhythmic music (Hagen & Bryant, 2003; Hagen & Hammerstein, 2009; cf. Merker, 2000b). Signatures of this function might persist in modern humans in coordinated group dances that are universal across cultures (Mehr et al., 2019; Nettl, 2015; e.g., the Māori *haka*; Best, 1924). The group music of Aka Congo Basin hunter-gatherers, for instance, is audible to groups living some distance away.[12]

Complex forms of social organization likely set the stage for the evolution of complex credible signals, including synchronized and coordinated vocalizations. Sometime after diverging from other apes, the human lineage underwent a major transition to a multilevel society. In multilevel societies, small family units regularly aggregate with other family units, forming a higher-level unit, which in some species aggregate to form an even higher-level unit. This societal structure occurs in some other primate species (e.g., hamadryas baboons; Swedell & Plummer, 2019), and some evidence suggests that higher degrees of social complexity are correlated with increased vocal diversity and flexibility (e.g., in macaques; Rebout, De Marco, Lone, & Thierry, 2020).

A notable attribute of some multilevel-society species is that, in addition to the agonism or tolerance exhibited between units, units also cooperate. *Homo sapiens* exhibits particularly rich cooperative behavior between units: cooperative families are nested within cooperative residential groups that often form cooperative alliances with other residential groups to obtain food, buffer resource variation, raise children, defend territory, and so on (Chapais, 2013; Hamilton, Milne, Walker, Burger, & Brown, 2007; Pisor & Surbeck, 2019; Rodseth, Wrangham, Harrigan, & Wolpoff, 1991; Swedell & Plummer, 2019).

Between-group cooperation likely created many new selection pressures. In particular, if human groups varied in the benefits they could provide other groups as allies, and the number of alliances a group could maintain was limited, a biological market would have arisen (Hammerstein & Noë, 2016), wherein groups evaluated the coalition quality of potential allies by assessing their size, cooperation ability, and willingness to cooperate, and potential allies had incentives to exaggerate these qualities (Hagen & Bryant, 2003; Hagen & Hammerstein, 2009). Common properties of music, especially those found in rhythmic, coordinated performances, provide a close fit to the necessary criteria for a credible signal of such otherwise difficult-to-observe group-level features. The time needed to create and practice group complex musical performances and achieve complex synchrony necessarily corresponds to a dimension of the underlying quality of the coalition: the amount of time coalition members have cooperated with one another.

In summary, we propose that music evolved, in part, as a means for groups to credibly show off their qualities to other groups.

There is substantial ethnographic, historical, and archaeological evidence of credible signaling of coalition quality among human groups, typically in the context of *feasting*. In feasting, two or more individuals share special types or quantities of foods, for a special purpose or event (Hayden, 2014). In addition to food, feasting often includes special clothing, ornaments, and other artifacts – and music and dance. Feasting has been documented in societies of all levels of social complexity, ranging from band-level hunter-gatherers to nation-states, including at archaeological sites throughout the Holocene (reviewed in Hayden, 2014; Hayden & Villeneuve, 2011). While many functions of feasting have been proposed (Hayden, 2014; Wiessner & Schiefenhövel, 1998), there is widespread agreement that feasts play a critical role in the formation of alliances between groups (reviewed in Hayden, 2014; Hayden & Villeneuve, 2011). As Sosis (2000) observed, the goal of ritualized foraging and feasting and other forms of food distribution is often to enhance the reputation of an entire group by displaying its productivity.

It is notable, then, that music and dance co-occur with feasting frequently in the ethnographic record.[13] For example, Congo Basin hunter-gatherers are renowned for their music, which they perform in many social contexts, including at spirit plays and large inter-community dances following big game kills (Fürniss, 2017; Lewis, 2013, 2017), as in the Mbendjele BaYaka:

> Sharing [food] between camps is less frequent, but will occur when big game is killed and during massana forest spirit performances. When an elephant is killed, Mbendjele in the area go rapidly to where the carcass is lying. Large camps grow, and feasting and dancing go on until the elephant has been consumed. (Lewis, 2017, p. 227)

It is plausible that similar events regularly occurred during human evolution because there is archaeological evidence for domestic spaces, large game hunting, mass kills, cooking, large aggregations, burials, ornaments, use of pigments, and musical instruments throughout the Upper Pleistocene, with some evidence appearing earlier (Barham, 2002; Conard et al., 2009; Kuhn, 2014; Kuhn & Stiner, 2019; Maher & Conkey, 2019; Stiner, 2019, 2013, 2017).

A function of music in the context of alliance formation may also help to explain why music is often produced for and enjoyed by strangers, as in modern recorded music and live concerts. The selective dynamics of "social foraging" in the hominin niche, where strangers have an uncertain but non-zero possibility of becoming lucrative social partners, frame strangers as the appropriate targets of social foraging tactics (Delton, Krasnow, Cosmides, & Tooby, 2011; Delton & Robertson, 2012; Rand, Peysakhovich, Kraft-Todd, & Greene, 2014). Moreover, some data show that observers can infer coalition quality and fighting ability from observations of musical performances and other coordinated behaviors. People who listened to a musical performance with instruments mixed either in-sync, consistently out-of-sync, or scrambled rated coalition quality higher in the in-sync versus out-of-sync (but not scrambled) conditions (Hagen & Bryant, 2003). When listening to people marching asynchronously or synchronously, judges rated the synchronous groups as more formidable, better able to coordinate a physical attack, and higher in social closeness; judgments of formidability were mediated by judgments of coordination, not bonding (Fessler & Holbrook, 2016).

On this view, music is clearly rooted in sociality. In contrast to the social bonding hypothesis, however, we predict that music does not *directly* cause social cohesion: rather, it signals existing social cohesion that was obtained by other means (Hagen & Bryant, 2003, p. 30).

We do not think this is the only social context in which music can act as a credible signal. Within groups, musical performances might also create common knowledge of decisions to cooperate, which could serve group coordination and cooperation (Chwe, 2001; Freitas, Thomas, DeScioli, & Pinker, 2019; see Hagen & Bryant, 2003 for other possibilities); credibly signal qualities guiding same-sex partner choice in a biological market (Hammerstein & Noë, 2016), and perhaps informing mate choice by *both* sexes; and as a group analog of emotional expression (Hagen & Bryant, 2003; Hagen & Hammerstein, 2009). Producing music that is specific to a group might also credibly signal membership in that group (Mehr et al., 2016; Mehr & Spelke, 2017) in a fashion similar to food preferences and dialects (see Kinzler, Dupoux, & Spelke, 2007; Liberman, Woodward, Sullivan, & Kinzler, 2016).

Next, we examine a case where we believe within-group rather than between-group credible signaling has shaped music.

4.2.2 Infant-directed song as a co-evolved system for negotiating parental investment of attention

Contact calls are a common vocalization across many species, distinct from territorial signals. In primates, these include loud calls between separated group members, and frequent quiet calls during heightened risk of separation (e.g., in dense vegetation). Contact calls rank among the most diverse and complex call types across species (Bouchet, Blois-Heulin, & Lemasson, 2013; Leighton, 2017), enabling individuals to recognize, estimate distance to, and maintain contact with their social partners (Kondo & Watanabe, 2009; Rendall, Cheney, & Seyfarth, 2000).

One important class of contact calls are those between parents and offspring. These serve functions of mutual interest to parents and offspring, for example, enabling parents to be available to solve problems their offspring are ill-suited to solve on their own. Chacma baboon barks, for example, range from tonal, harmonically rich variants that are used for contact calls, to barks with a noisier, harsher structure that are used for alarm calls. By the age of 6 months, infants learn to discriminate call types and to discriminate their mothers' contact barks from those of unrelated females (Fischer, Cheney, & Seyfarth, 2000); and mothers recognize their infants' contact calls (Rendall et al., 2000).

We propose that in the human lineage, maternal contact calls evolved to encode credible information beyond identity and distance, namely attention to the infant. There are few relationships where inclusive fitness interests overlap as much as they do between parents and offspring – but even these are not perfectly aligned. Because of the mechanics of diploid sexual reproduction, a parent is equally related to all her offspring, whereas each offspring is twice and four times as related to itself as it is to each of its full and half siblings, respectively. A strategy that optimizes the parent's inclusive fitness (e.g., equal food distribution across offspring) does not necessarily optimize an offspring's inclusive fitness, and vice versa. This possibility, parent–offspring conflict (Trivers, 1974), implies differences in the interests of supply genes in the parent and demand genes in the offspring (Bossan, Hammerstein, & Koehncke, 2013).

Some aspects of human reproduction suggest that selection pressures for complex contact calls have increased relative to those in apes. First, human brain size is about triple that of other apes (Schoenemann, 2006) and most brain growth occurs postnatally, implying that human infants are born helpless and have a very long juvenile period. Second, human forager interbirth intervals are about half those of chimpanzees (Marlowe, 2005; Thompson, 2013), requiring ancestral human mothers to simultaneously care for multiple dependent offspring (in contrast to chimpanzee mothers, who typically care for a single dependent offspring). Third, unlike other great apes, humans rely heavily on alloparenting in a multilevel society, requiring ancestral human infants to establish relationships with multiple caregivers and vice versa (Hrdy, 2009). Unlike chimpanzee infants, ancestral human infants typically competed with multiple juveniles for the attention of multiple caregivers.

Human parents increase their offspring's fitness by attending to them and protecting them from harm. Attention is a limited resource, however; many other challenges require attention, and solving those may benefit the parent more than the infant (relative to the provisioning of attention to maintain infant safety). The interests of infants and their parents conflict, in terms of the optimal provisioning of attention: infants often "prefer" more attention than a parent would "prefer" to provide.

How does this conflict of interests play out? Infants have bargaining power to extract parental investment (in the form of material investment, like nursing, or parental attention); they demand attention by crying (for discussion of evolutionary scenarios, see Soltis, 2004 and commentaries). Parents lack perfect access to their infant's internal state, so crying provides information about when investment can be provided. Care-eliciting infant vocalizations (e.g., distress calls, separation calls) are common across mammals (Newman, 2007), including humans, and mothers reliably respond to these vocalizations by providing care (Bornstein, Putnick, Rigo, & Venuti, 2017).

Whereas infants can easily detect when material investment has been provisioned, attention is a covert property of the parent's mind, with unreliable cues. Infants can infer that parents are attending to them from estimating the parent's gaze direction, but this only provides partial information (the parent could be concentrating on something else). Touch is also a good cue that a parent is nearby; but the parent could be asleep, or attending to something else.

Better than these cues would be a credible signal from the parent, reliably indicating that the infant has their attention (Mehr & Krasnow, 2017). A vocal signal is a good candidate because its acoustic properties allow the proximity of the producer of the signal to be reliably inferred by the target. To the extent that the signal monopolizes the vocal apparatus, producing it is incompatible with other activities (such as speaking to another adult) that could co-opt the parent's attention. And aspects of the vocal signal can be modulated in real time, in response to the infant's state and behavior, which cannot be done without attending to the infant.

Here again we expect an evolutionary arms race, driven by partially conflicting fitness interests between senders and receivers, producing an elaborated signal. We propose that this process could lead to key features of music: in particular, contrasting with the rhythmic features developing from territorial signals, we expect the rather more subdued context of soothing parent–infant contact calls to give rise to melodic features, tokens of which are the lullabies we sing to infants today (Mehr & Krasnow, 2017).

Three sets of results support this idea. First, if adaptations support the production of song in parents and alloparents, and the appetite for and ability to perceive song in infants, then music should appear universally in the context of infant care and infant-directed songs should share features worldwide. These predictions, long discussed in the music cognition literature (Hannon & Trainor, 2007; Peretz, 2006; Trehub, 2001), are well-evidenced. In an analysis of high-quality ethnography from a representative sample of human societies, text concerning vocal music was significantly associated with infant care and children, over and above base rates of reporting (this finding replicates both with expert annotations of the ethnography and automated text analysis; Mehr et al., 2019). Moreover, infant-directed songs are found in 100% of a pseudorandom sample of field recordings in mostly-small-scale societies; and naïve listeners, who are unfamiliar with the languages or cultures involved, reliably recognize them as infant-directed, with remarkable consistency (Mehr & Singh et al., 2018, 2019). This finding replicates prior cross-cultural work (Trehub, Unyk, & Trainor, 1993a, 1993b).

Second, the genetic architecture of musical perception and motivation should be regulated, in part, by parent-of-origin epigenetic mechanisms, such as genomic imprinting. Humans are sexually reproducing but not obligately monogamous, which differentiates the conflict of interest between parents and offspring by parental sex: because maternity certainty is greater than paternity certainty, genes of maternal origin are more likely to be found in an offspring's siblings than genes of paternal origin (Haig & Wilkins, 2000). Genes of maternal origin are thus under selection to bias the tradeoff in demand for parental investment in the direction of the offspring's siblings and away from the offspring; on average, maternally inherited genes should reduce investment demands on mothers, and vice versa. This prediction is confirmed by the fact that genes with parent-of-origin effects tend to affect demands for parental investment, such as intrauterine growth (Haig, 1993).

Genomic imprinting disorders, where genetic dysregulation is differentiated by parent-of-origin, provide a unique test of the relation between a trait and its putative link to parental investment (Haig & Wharton, 2003). Angelman and Prader-Willi syndromes result from opposing dysregulation at the same genetic region (15q11-13), with a loss of genes expressing maternal interest resulting in Angelman syndrome, and the reverse, a relative loss of genes expressing paternal interest resulting in Prader-Willi syndrome. The behavioral phenotypes reflect the different effects of maternally versus paternally inherited genes: infants with Angelman syndrome have a voracious appetite while nursing, are awake for more hours of the day than typically developing infants, and attract more attention via smiling than do typically developing children (Ubeda, 2008; Williams, Beaudet, Clayton-Smith, & Wagstaff, 2006), increasing investment demands on the mother. Infants with Prader-Willi syndrome, in contrast, are born with low birth weight, sleep more than typically developing infants, and often lack a suckle reflex (Cassidy & Driscoll, 2008; Holm, Cassidy, Butler, & Greenberg, 1993; Peters, 2014), with the opposite effect (decreasing investment demands on the mother).

Recent findings show that these effects extend to the domain of music, demonstrating a genetic link between music perception and parental investment. People with Angelman syndrome have a suppressed relaxation response to music (Kotler, Mehr, Egner, Haig, & Krasnow, 2019); while people with Prader-Willi syndrome have a potentiated relaxation response to music, along with pitch perception deficits (Mehr, Kotler, Howard, Haig, & Krasnow, 2017). These results support the idea that music signals attention: suppressed relaxation in Angelman syndrome implies increased maternal demands, while potentiated relaxation in Prader-Willi syndrome implies reduced maternal demands, in line with other findings concerning parental investment demands in genomic imprinting disorders.

Last, we also expect relationships between the acoustic features of non-human primate contact calls and human infant-directed song. While few data exist with which to test these relationships, preliminary findings suggest that similarities do exist. For example, baboon contact calls are harmonically rich, whereas alarm calls are harsh and noisy (Fischer et al., 2000); in a vocalization corpus from 21 human societies, infant-directed song was acoustically distinct from infant-directed speech across many pitch, rhythmic, phonetic, and timbral attributes (Moser et al., 2020), with a similar pattern of results to the acoustic differences between baboon contact calls and alarm calls. Moreover, several acoustic features driving these effects were related to vocal exertion (e.g., temporal modulation, pitch rate, vowel rate), perhaps honestly signaling additional costs incurred by the signaler.

5. Discussion

A comprehensive understanding of music requires that proximate-level explanations are distinguished from ultimate-level explanations uniquely linked to music; that proposed adaptations explain the core features of music that are putatively shaped by natural selection, and distinguish them from features that are byproducts of other adaptations; and finally, that the results of evolutionary analyses provide a foundation on which cultural-evolutionary processes can plausibly act.

The credible signaling account meets these criteria, whereas other accounts of the origins of music do not.

5.1 Credible signaling may explain some basic features of music

Early in this paper we noted some properties of human music that need explanation. While we find it implausible that any one theory can explain all of them, two core features of music are directly related to the ideas presented here.

An evolved system for quickly and reliably signaling coalition quality, which might otherwise be difficult to perceive, especially during territorial advertisements, agonistic intergroup encounters (e.g., war songs, dances), and alliance-forging feasts, provides a functional explanation for rhythm: selection pressures toward synchronized isochronous sounds, with complex internal design. An evolved system for credibly signaling parental attention to infants provides a functional explanation for melody: selection pressures toward manipulating affective prosody in vocalizations, constrained by the physics of the vocal production system and inherent features of the auditory world.

These "building blocks" appear universally in music (Mehr et al., 2019; Nettl, 2015; Savage et al., 2015), like "building blocks" of language (e.g., Baker, 2001). They provide a grammar-like, combinatorially generative interface through which musical content can be created, improvised, and elaborated upon, through hierarchical organization of meter and tonality[14] (Krumhansl, 2001; Lerdahl & Jackendoff, 1983), in fashions that themselves have universal signatures (Jacoby & McDermott, 2017; Jacoby et al., 2019; Mehr et al., 2019).

The importance of rhythm and pitch in human music perception – and the degree to which these features of music are unique to human vocalizations – may be directly tied to their evolutionary history.

5.2 Music is culturally evolved but cultural evolution has to start somewhere

We understand culture as information that affects individuals' behavior and that is acquired from conspecifics through teaching, imitation, and other types of social transmission (Boyd & Richerson, 2004; Tooby & Cosmides, 1992). Because information is transmitted with some degree of fidelity through non-genetic means (e.g., memory, learning), information is cumulative. Some cultural information is passed on with greater frequency and higher regularity than other information. For example, social learners tend to pay attention to information sources that have established prestige more than sources that do not (Henrich & Boyd, 2002). Similarly, some information is easier to learn than other information; children exhibit interest about information associated with danger and retain it with greater fidelity and over longer periods than related information unassociated with danger (Barrett, Peterson, & Frankenhuis, 2016; Wertz, 2019).

One characteristic of cumulative culture is ritualization (Lorenz, 1966), analogous to co-evolutionary processes underlying animal communication systems (Krebs & Dawkins, 1984). Cultural signals can develop extravagant physical features resulting from arms race dynamics, particularly in cases when there is a conflict of interest between senders and receivers. Examples from modern environments include conspicuously branded luxury goods, which can signal wealth (Han, Nunes, & Drèze, 2010); or businesses that engage in one-upmanship by incorporating exaggerated sensory features in competitive advertisements (Dunham, 2011).

Such cultural ritualization is likely at play in the musical domain, especially given the increasingly important role of elaborate feasting in the cultural evolution of social complexity across the globe throughout the Holocene (Hayden, 2014), and given the highly variable musical features that continually unfold over time across compositional styles, instrumentation and orchestration, improvisatory motifs, setting lyrics to music, and so on. Music must be shaped by culture in all contexts, however, not only those of coalition signaling and parental care.

We propose that the adaptations proposed here provide a foundation for cultural-evolutionary processes. These traits – particular grammar-like structures, for instance, such as tonalities and meters – gravitate toward certain forms, or "attractors," and away from others (Sperber, 1996; Sperber & Hirschfeld, 2004). These attractors will interact with evolved capacities for non-musical traits resulting in sensitivity and attraction to features in communicative acts that trigger them (e.g., musical phenomena that evoke the sound of an emotional voice), increasing variability in music.

As this process repeats within and across cultures, the diversity of music increases, while underlaid by universals that can be traced back to music's adaptive functions in credible signaling. This pattern of universality and diversity is exactly what is observed in systematic analyses of music across cultures (Mehr et al., 2019) and, we believe, is what continues to shape music, worldwide, today.

Understanding this variability has been a longstanding interest of ethnomusicologists, who document musical traditions as they are shaped by social environments, politics, and ethnolinguistic history (Blacking, 1973; Feld, 1984; Nettl, 2015), but it has strong parallels in the study of cultural evolution and social transmission. In particular, the prevalence of specific musical features (a particular scale, musical instrument, ornament, vocal practice, and so on) in a given society's music is likely to be shaped by that society's relation to other societies, just as the presence or absence of linguistic features is predictable by lineage (Dunn, Greenhill, Levinson, & Gray, 2011).

We expect that studying the cultural evolution of musical features will be a productive endeavor (with promising first steps already underway; e.g., Savage et al., 2015). We predict, however, that those features least likely to be shaped by culture are those core features predicted by the evolutionary account described here. For example, whereas we expect few musical systems worldwide to *lack* melody and rhythm as core features, we expect many to have rather different *instantiations* of those features. This is uncontroversial: while scales commonly used in music differ across cultures, they nevertheless are mutually intelligible, implying shared psychological mechanisms for music perception surrounding the interpretation of melodies (Castellano, Bharucha, & Krumhansl, 1984; Krumhansl et al., 2000; Mehr et al., 2019).

5.3 Auditory cheesecake: not wrong, but not right either

A key difficulty of studying the evolution of music, to which we alluded throughout this paper, is that the present environment has diverged from the environment in which humans evolved. In this context, Pinker's (1997) "auditory cheesecake" analogy for a byproduct account of music is neither surprising nor controversial. We should expect *many* human behaviors to have cheesecake-like features. Just as the world's great writers have stretched the bounds of human language far beyond language's original adaptive functions, the boundless creativity of composers and performers have created an actual domain of music that, we believe, is quite far from its proper domain.

In this sense, we agree with Pinker that many musical inventions are byproducts, plain and simple: auditory cheesecake is not wrong. But in light of the adaptations proposed here, auditory cheesecake isn't right either: as we have argued, in at least two contexts, music exhibits design features consistent with adaptations for credible signaling, which give rise to a universal human psychology of music.

6. Conclusions

Why study the origins of music, language, or any other human behavior? It's unlikely that anyone will ever explain the full extent to which a particular behavior is accounted for by one or more adaptations because, given its complexity, human behavior cannot be exhaustively measured.

Nevertheless, we think that inching toward a functional understanding of complex behavior helps determine what the phenomena in question are, exactly, by isolating the core psychological representations and cultural processes underlying the phenomena from those that are merely associated with them. In the case of music, the analyses presented here lay out a roadmap for understanding the phenomenon of human musicality.

Music-like behaviors occur in a broad swath of species, including our ape relatives, and increasing evidence indicates that these serve important credible signaling functions among agents with conflicts of interest, such as territorial advertisements and mate attraction. In humans, across cultures, music is associated with social behaviors that directly involve credible signaling private information among agents with conflicts of interest, especially coalitional interactions and infant care, but perhaps others too. Accordingly, the psychological mechanisms for processing and producing features of music that are implied by those contexts, such as melody and rhythm, should also be universal; all of this is proposed to constitute music's proper domain. In music's actual domain, in contrast, we should expect the engine of cultural evolution to develop and expand these features, producing a diverse set of musical manifestations worldwide that retain some key features of their evolved functions.

Additional mechanisms likely interact with these core features. These may include psychological mechanisms that enable the perception of higher-level features of music, such as implied harmony or musical emotions; linguistic mechanisms that shape the ways in which language and music are intertwined; cultural mechanisms that drive musical traditions and are shaped historically as cultures mix and combine to form new cultures; technological mechanisms that directly alter the feature space of musicality, including musical inventions, such as instruments and music production software, or new musical forms, such as microtonal music; and, not least, aesthetic mechanisms that drive the preferences and interests of those who make and listen to music worldwide.

Understanding these mechanisms in isolation and as they interact with each other to produce the phenomenon of human musicality is a key challenge for the field – a challenge that will be served well by a clear explanation for the origins of music, which can then be built upon using the interdisciplinary toolkit of modern science.

Acknowledgments. We thank Paul Seabright, the Institute for Advanced Study in Toulouse, and the Fondation Royaumont, whose meeting "Origins of music in non-state societies" sparked the idea for this paper; and Mila Bertolo, Courtney Hilton, Cody Moser, and Manvir Singh for feedback on the manuscript.

Financial support. The "Origins of music in non-state societies" meeting was supported by the Agence Nationale de la Recherche (Investissements d'Avenir ANR-17-EURE-0010). S.A.M. is supported by the Harvard Data Science Initiative and the National Institutes of Health Director's Early Independence Award DP5OD024566.

Conflict of interest. None.

Notes

1. The degree to which music-like behaviors in non-human species are homologous to music is up for debate (see Bertolo, Singh, & Mehr, 2021; Honing, Bouwer, Prado, & Merchant, 2018; McDermott & Hauser, 2005), especially given surprising differences in auditory cognition and auditory preferences across species (Bregman, Patel, & Gentner, 2016; McDermott & Hauser, 2004, 2007). For discussion, see Kotz, Ravignani, and Fitch (2018) and Patel (2017).
2. We leave aside a serious issue: most studies of prosocial effects of music-making are vulnerable to participant expectancy effects, which may account for the literature's poor reproducibility (Atwood, Mehr, & Schachner, 2020).
3. We leave aside intense debates over whether or not genetic group selection is tenable; see Pinker (2012) and commentaries.
4. We thank anonymous Reviewer 5 for this example.
5. We thank Cody Moser for suggesting this quotation.
6. A fascinating exception is the phenomenon of echolocation, wherein the sender and receiver of a vocal signal are the same organism.
7. We leave aside deceptive signals, which benefit the signaler at the expense of the receiver.
8. For discussion of cues, indices, costly signals, and their relationships, see Biernaskie, Perry, and Grafen (2018) and references therein.
9. This is true even of adaptations that subsequently change; bird feathers served as insulation before supporting flight but they retain features revealing their original function (see Persons & Currie, 2019 and your duvet).
10. Cycle effects on mate preferences, in general, have been questioned by recent studies (Gangestad, Haselton, Welling, & Puts, 2016; Jones, Hahn, & DeBruine, 2018).
11. Musical preferences change modestly during middle childhood (e.g., Hargreaves, Comber, & Colley, 1995) but whether the *frequency* of musical behaviors also changes is unknown.
12. This statement is supported by personal interviews in E.H.'s fieldwork.
13. Whereas evidence of feasting is abundant in agricultural and complex hunter-gatherer societies throughout the Holocene, it is less well-documented in simple hunter-gatherer societies, with some exceptions (Hayden, 2014; Wallis & Blessing, 2015).
14. Here, Temperley's (2004) discussion of communicative structure in the evolution of musical style may have surprising parallels in the biological evolution of music.

References

Alexander, R. D. (1974). The evolution of social behavior. *Annual Review of Ecology and Systematics*, *5*(1), 325–383.

Andersson, M. B. (1994). *Sexual selection*. Princeton University Press.

Anshel, A., & Kipper, D. A. (1988). The influence of group singing on trust and cooperation. *Journal of Music Therapy*, *25*(3), 145–155.

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, *276* (1659), 1077–1082.

Ariew, A. (2002). Platonic and Aristotelian roots of teleological arguments. In A. Ariew, R. Cummins, & M. Perlman (Eds.) *Functions: New essays in the philosophy of psychology and biology* (pp. 7–32). Oxford University Press.

Askin, N., Mauskapf, M., Koppman, S., & Uzzi, B. (2020). Do women produce more novel work than men? Gender differences in musical creativity. *Working Paper*.

Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. J., & Pagel, M. (2008). Languages evolve in punctuational bursts. *Science (New York, N.Y.)*, *319*(5863), 588.

Atwood, S., Mehr, S., & Schachner, A. (2020). Expectancy effects threaten the inferential validity of synchrony-prosociality research. *PsyArXiv*. https://doi.org/10.31234/osf.io/zjy8u.

Bainbridge, C. M., Bertolo, M., Youngers, J., Atwood, S., Yurdum, L., Simson, J., … Mehr, S. A. (2021). Infants relax in response to unfamiliar foreign lullabies. *Nature Human Behaviour*, *5*(2), 256–64. doi: 10.1038/s41562-020-00963-z.

Baker, M. C. (2001). *The atoms of language*. Basic Books.

Barham, L. S. (2002). Systematic pigment use in the middle pleistocene of South-Central Africa. *Current Anthropology*, *43*(1), 181–190.

Barrett, H. C., Peterson, C. D., & Frankenhuis, W. E. (2016). Mapping the cultural learnability landscape of danger. *Child Development*, *87*(3), 770–781.

Barrow, J. D. (2005). *The artful universe expanded*. Oxford University Press.

Barton, R. (1985). Grooming site preferences in primates and their functional implications. *International Journal of Primatology*, *6*(5), 519–532.

Bates, B. C. (1970). Territorial behavior in primates: A review of recent field studies. *Primates*, *11*(3), 271–284.

Beetz, A., Uvnäs-Moberg, K., Julius, H., & Kotrschal, K. (2012). Psychosocial and psychophysiological effects of human-animal interactions: The possible role of oxytocin. *Frontiers in Psychology*, *3*, 234.

Behr, O., von Helversen, O., Heckel, G., Nagy, M., Voigt, C. C., & Mayer, F. (2006). Territorial songs indicate male quality in the sac-winged bat *Saccopteryx bilineata* (Chiroptera, Emballonuridae). *Behavioral Ecology*, *17*(5), 810–817.

Benetos, E., Dixon, S., Giannoulis, D., Kirchhoff, H., & Klapuri, A. (2013). Automatic music transcription: Challenges and future directions. *Journal of Intelligent Information Systems*, *41*(3), 407–434.

Benton, T. G., Baguette, M., Clobert, J., Bullock, J. M., & Oxford University Press. (2017). *Dispersal ecology and evolution*. Oxford University Press.

Benzon, W. (2001). *Beethoven's anvil: Music in mind and culture*. Basic Books.

Bertolo, M., Singh, M., & Mehr, S. A. (2021). Sound-induced motion in chimpanzees does not imply shared ancestry for music or dance. *Proceedings of the National Academy of Sciences*, *118*(2), e2015664118.

Best, E. (1924). *The Māori* (Vol. 2). H.H. Tombs.

Biernaskie, J. M., Perry, J. C., & Grafen, A. (2018). A general model of biological signals, from cues to handicaps. *Evolution Letters*, *2*(3), 201–209.

Blacking, J. (1973). *How musical is man?* University of Washington Press.

Bornstein, M. H., Putnick, D. L., Rigo, P., Esposito, G., Swain, J. E., Suwalsky, J. T. D., … Venuti, P. (2017). Neurobiology of culturally common maternal responses to infant cry. *Proceedings of the National Academy of Sciences*, *114*(45), E9465–E9473.

Bossan, B., Hammerstein, P., & Koehncke, A. (2013). We were all young once: An intragenomic perspective on parent–offspring conflict. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1754), 20122637.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, *4*, 390.

Bowler, D. E., & Benton, T. G. (2005). Causes and consequences of animal dispersal strategies: Relating individual behaviour to spatial dynamics. *Biological Reviews*, *80*(2), 205–225.

Boyd, R., & Richerson, P. J. (2004). *The origin and evolution of cultures*. Oxford University Press.

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication*. Sinauer Associates.

Bregman, M. R., Patel, A. D., & Gentner, T. Q. (2016). Songbirds use spectral shape, not pitch, for sound pattern recognition. *Proceedings of the National Academy of Sciences*, *113*(6), 1666–1671.

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*, pp. 231–281). Springer.

Brown, S. (2000). The "Musilanguage" model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). MIT Press.

Bryant, G. A. (2013). Animal signals and emotion in music: Coordinating affect across groups. *Frontiers in Psychology*, *4*, 990.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*(1), 1–14.

Cassidy, S. B., & Driscoll, D. J. (2008). Prader-Willi syndrome. *European Journal of Human Genetics*, *17*(1), 3–13.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of north India. *Journal of Experimental Psychology: General*, *113*(3), 394–412.

Catchpole, C., & Slater, P. J. B. (2018). *Bird song: Biological themes and variations*. Cambridge University Press.

Cattell, J. M. (1891). On the origin of music. *Mind; A Quarterly Review of Psychology and Philosophy*, *16*(63), 375–388.

Chapais, B. (2013). Monogamy, strongly bonded groups, and the evolution of human social structure. *Evolutionary Anthropology: Issues, News, and Reviews*, *22*(2), 52–65.

Charlton, B. D. (2014). Menstrual cycle phase alters women's sexual preferences for composers of more complex music. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1784), 20140403.

Charlton, B. D., Filippi, P., & Fitch, W. T. (2012). Do women prefer more complex music around ovulation? *PLoS ONE*, *7*(4), e35626.

Chen, Z., & Wiens, J. J. (2020). The origins of acoustic communication in vertebrates. *Nature Communications*, *11*(1), 369.

Chomsky, N. A. (1968). *Language and mind*. Harcourt, Brace and World.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489–509.

Chwe, M. S.-Y. (2001). *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011.

Coen, P., Xie, M., Clemens, J., & Murthy, M. (2016). Sensorimotor transformations underlying variability in song intensity during Drosophila courtship. *Neuron*, *89*(3), 629–644.

Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, *460*(7256), 737–740.

Conroy-Beam, D., Buss, D. M., Pham, M. N., & Shackelford, T. K. (2015). How sexually dimorphic are human mate preferences? *Personality and Social Psychology Bulletin*, *41* (8), 1082–1093.

Cross, I. (2012). Cognitive science and the cultural nature of music. *Topics in Cognitive Science*, *4*(4), 668–677.

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. 61–81). Oxford University Press.

Cross, I., & Woodruff, G. E. (2009). Music as a communicative medium. In R. Botha & C. Knight (Eds.), *The prehistory of language* (Vol. *1*, pp. 113–144). Oxford University Press.

Cummings, M. E., & Endler, J. A. (2018). 25 Years of sensory drive: The evidence and its watery bias. *Current Zoology*, *64*(4), 471–484.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., & Valcu, M. (2015). The effects of life history and sexual selection on male and female plumage colouration. *Nature*, *527* (7578), 367–370.

Darwin, C. (1859). *On the origin of species by means of natural selection*. J. Murray.

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

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, *108*(32), 13335–13340.

Delton, A. W., & Robertson, T. E. (2012). The social cognition of social foraging: Partner selection by underlying valuation. *Evolution and Human Behavior*, *33*(6), 715–725.

Dickins, T. E., & Barton, R. A. (2013). Reciprocal causation and the proximate–ultimate distinction. *Biology & Philosophy*, *28*(5), 747–756.

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. (2008). If music is the food of love, what about survival and reproductive success? *Musicae Scientiae*, *12*(1 Suppl.), 169–195.

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.

Dunbar, R. I. M. (1991). Functional significance of social grooming in primates. *Folia Primatologica*, *57*(3), 121–131.

Dunbar, R. I. M. (1998). *Grooming, gossip, and the evolution of language*. Harvard University Press.

Dunbar, R. I. M. (2004). Language, music, and laughter in evolutionary perspective. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 257–274). MIT Press.

Dunbar, R. I. M. (2012a). Bridging the bonding gap: The transition from primates to humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1837–1846.

Dunbar, R. I. M. (2012b). 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., Kaskatis, K., MacDonald, I., & Barra, V. (2012). Performance of music elevates pain threshold and positive affect: Implications for the evolutionary function of music. *Evolutionary Psychology*, *10*(4), 688–702.

Dunbar, R. I. M., Lehmann, J. (2013). Grooming and social cohesion in primates: A comment on Grueter et al. *Evolution and Human Behavior*, *34*(6), 453–455.

Dunham, B. (2011). The role for signaling theory and receiver psychology in marketing. In G. Saad (Ed.), *Evolutionary psychology in the business sciences* (pp. 225–256). Springer Berlin Heidelberg.

Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, *473*(7345), 79–82.

Duputié, A., & Massol, F. (2013). An empiricist's guide to theoretical predictions on the evolution of dispersal. *Interface Focus*, *3*(6), 20130028.

Dutton, D. (2009). *The art instinct: Beauty, pleasure, & human evolution*. Bloomsbury Press.

Feld, S. (1984). Sound structure as social structure. *Ethnomusicology*, *28*(3), 383–409.

Fessler, D. M. T., & Holbrook, C. (2016). Synchronized behavior increases assessments of the formidability and cohesion of coalitions. *Evolution and Human Behavior*, *37*(6), 502–509.

Fischer, J., Cheney, D. L., & Seyfarth, R. M. (2000). Development of infant baboons' responses to graded bark variants. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1459), 2317–2321.

Fitch, W. T. (2006). Production of vocalizations in mammals. *Visual Communication*, *3*, 145.

Fitch, W. T. (2017). Preface to the special issue on the biology and evolution of language. *Psychonomic Bulletin & Review*, *24*(1), 1–2.

Freeman, W. J. (2000). A neurobiological role of music in social bonding. In N. Wallin, B. Merkur & S. Brown (Eds.), *The origins of music* (pp. 411–424). MIT Press.

Freitas, J. D., Thomas, K., DeScioli, P., & Pinker, S. (2019). Common knowledge, coordination, and strategic mentalizing in human social life. *Proceedings of the National Academy of Sciences*, *116*(28), 13751–13758.

Fritz, T. H., Hardikar, S., Demoucron, M., Niessen, M., Demey, M., Giot, O., … Leman, M. (2013). Musical agency reduces perceived exertion during strenuous physical performance. *Proceedings of the National Academy of Sciences*, *110*(44), 17784–17789.

Fürniss, S. (2017). Diversity in pygmy music: A family portrait. In B. S. Hewlett (Ed.), *Hunter-gatherers of the Congo basin: Cultures, histories, and biology of African pygmies* (pp. 217–248). Routledge.

Gangestad, S. W., Haselton, M. G., Welling, L. L. M., Gildersleeve, K., Pillsworth, E. G., Burriss, R., … Puts, D. A. (2016). How valid are assessments of conception probability in ovulatory cycle research? Evaluations, recommendations, and theoretical implications. *Evolution and Human Behavior*, *37*(2), 85–96.

Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 103–123). MIT Press.

Geist, V. (1978). *Life strategies, human evolution, environmental design: Toward a biological theory of health*. Springer-Verlag.

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*(1805), 20190432.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Belknap Press of Harvard University Press.

Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A.-Y., & DeCasper, A. J. (2011). A melodic contour repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. *PLoS ONE*, *6*(2), e17304.

Greenberg, D. M., Matz, S. C., Schwartz, H. A., & Fricke, K. R. (2020). The self-congruity effect of music. *Journal of Personality and Social Psychology*. doi: 10.1037/pspp0000293.

Grueter, C. C., Bissonnette, A., Isler, K., & van Schaik, C. P. (2013). Grooming and group cohesion in primates: Implications for the evolution of language. *Evolution and Human Behavior*, *34*(1), 61–68.

Gustison, M. L., & Townsend, S. W. (2015). A survey of the context and structure of high- and low-amplitude calls in mammals. *Animal Behaviour*, *105*, 281–288.

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*(1), 21–51.

Hagen, E. H., & Hammerstein, P. (2009). Did Neanderthals and other early humans sing? Seeking the biological roots of music in the territorial advertisements of primates, lions, hyenas, and wolves. *Musicae Scientiae*, *13*(2 Suppl.), 291–320.

Haig, D. (1993). Genetic conflicts in human pregnancy. *The Quarterly Review of Biology*, *68*(4), 495–532.

Haig, D., & Wharton, R. (2003). Prader-Willi syndrome and the evolution of human childhood. *American Journal of Human Biology*, *15*(3), 320–329.

Haig, D., & Wilkins, J. F. (2000). Genomic imprinting, sibling solidarity and the logic of collective action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *355*(1403), 1593–1597.

Hall, M. L., & Magrath, R. D. (2007). Temporal coordination signals coalition quality. *Current Biology*, *17*(11), R406–R407.

Hamilton, M. J., Milne, B. T., Walker, R. S., Burger, O., & Brown, J. H. (2007). The complex structure of hunter–gatherer social networks. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1622), 2195–2203.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I & II. *Journal of Theoretical Biology*, *7*(1), 1–16.

Hammerstein, P., & Noë, R. (2016). Biological trade and markets. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1687), 20150101.

Han, Y. J., Nunes, J. C., & Drèze, X. (2010). Signaling status with luxury goods: The role of brand prominence. *Journal of Marketing*, *74*(4), 15–30.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*(11), 466–472.

Hargreaves, D. J., Comber, C., & Colley, A. (1995). Effects of age, gender, and training on musical preferences of British secondary school students. *Journal of Research in Music Education*, *43*(3), 242–250.

Harrington, F. H. (1989). Chorus howling by wolves: Acoustic structure, pack size and the Beau Geste effect. *Bioacoustics*, *2*(2), 117–136.

Harrington, F. H., & Mech, L. D. (1979). Wolf howling and its role in territory maintenance. *Behaviour*, *68*(3–4), 207–249.

Hayden, B. (2014). *The power of feasts*. Cambridge University Press.

Hayden, B., & Villeneuve, S. (2011). A century of feasting studies. *Annual Review of Anthropology*, *40*(1), 433–449.

Henrich, J., & Boyd, R. (2002). On modeling cognition and culture: Why cultural evolution does not require replication of representations. *Journal of Cognition and Culture*, *2*(2), 87–112.

Herdt, G., & McClintock, M. (2000). The magical age of 10. *Archives of Sexual Behavior*, *29*(6), 587–606.

Hill, G. E. (1994). Trait elaboration via adaptive mate choice: Sexual conflict in the evolution of signals of male quality. *Ethology Ecology & Evolution*, *6*(3), 351–370.

Hilton, C. B., Crowley, L., Yan, R., Martin, A., & Mehr, S. A. (2021). Children infer the behavioral contexts of unfamiliar foreign songs. *PsyArXiv*. https://doi.org/10.31234/osf.io/rz6qn.

Holm, V. A., Cassidy, S. B., Butler, M. G., Hanchett, J. M., Greenswag, L. R., Whitman, B. Y., & Greenberg, F. (1993). Prader-Willi syndrome: Consensus diagnostic criteria. *Pediatrics*, *91*(2), 398–402.

Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus monkeys (*Macaca mulatta*) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *12*, 475. doi: 10.3389/fnins.2018.00475.

Honing, H., & Ploeger, A. (2012). Cognition and the evolution of music: Pitfalls and prospects. *Topics in Cognitive Science*, *4*(4), 513–524.

Hrdy, S. B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. Harvard University Press.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, *930*(1), 43–61.

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception: An Interdisciplinary Journal*, *26*(3), 195–204.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*(1), 33–72.

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, *27*(3), 359–370.

Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & McDermott, J. H. (2019). Universal and non-universal features of musical pitch perception revealed by singing. *Current Biology*, *29*(19), 3229–3243.

Jaeggi, A. V., Kramer, K. L., Hames, R., Kiely, E. J., Gomes, C., Kaplan, H., & Gurven, M. (2017). Human grooming in comparative perspective: People in six small-scale societies groom less but socialize just as much as expected for a typical primate. *American Journal of Physical Anthropology*, *162*(4), 810–816.

James, W. (1890). *The principles of psychology*. H. Holt.

Jones, B. C., Hahn, A. C., & DeBruine, L. M. (2018). Ovulation, sex hormones, and women's mating psychology. *Trends in Cognitive Sciences*, *23*(1), 51–62.

Jourdain, R. (1997). *Music, the brain, and ecstasy: How music captures our imagination*. W. Morrow.

Justus, T., & Hutsler, J. J. (2005). Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain specificity. *Music Perception: An Interdisciplinary Journal*, *23*(1), 1–27.

Kaplan, H. S., & Gangestad, S. W. (2005). Life history theory and evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 68–95). John Wiley & Sons Inc.

Keller, P. E., König, R., & Novembre, G. (2017). Simultaneous cooperation and competition in the evolution of musical behavior: Sex-related modulations of the singer's formant in human chorusing. *Frontiers in Psychology*, *8*, 1559.

Kenrick, D. T., & Keefe, R. C. (1992). Age preferences in mates reflect sex differences in human reproductive strategies. *Behavioral and Brain Sciences*, *15*(1), 75–91.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577–12580.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299–314.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*(12), 578–584.

Kogan, N. (1994). On aesthetics and its origins: Some psychobiological and evolutionary considerations. *Social Research*, *61*(1), 139–165.

Kokko, H. (1997). Evolutionarily stable strategies of age-dependent sexual advertisement. *Behavioral Ecology and Sociobiology*, *41*(2), 99–107.

Kokko, H., Jennions, M. D., & Brooks, R. (2006). Unifying and testing models of sexual selection. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 43–66.

Kondo, N., & Watanabe, S. (2009). Contact calls: Information and social function. *Japanese Psychological Research*, *51*(3), 197–208.

Kotler, J., Mehr, S. A., Egner, A., Haig, D., & Krasnow, M. M. (2019). Response to vocal music in Angelman syndrome contrasts with Prader-Willi syndrome. *Evolution and Human Behavior*, *40*(5), 420–426.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, *22*(10), 896–910.

Krause, B. L. (2012). *The great animal orchestra: Finding the origins of music in the world's wild places* (1st ed.). Little, Brown.

Krebs, J., & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. Krebs & N. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 380–402). Blackwell.

Krebs, J. R. (1977). The significance of song repertoires: The Beau Geste hypothesis. *Animal Behaviour*, *25*, 475–478.

Krumhansl, C. L. (2001). *Cognitive foundations of musical pitch*, Oxford University Press.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, *76*(1), 13–58.

Kuhn, S. L. (2014). Signaling theory and technologies of communication in the Paleolithic. *Biological Theory*, *9*(1), 42–50.

Kuhn, S. L., & Stiner, M. C. (2019). Hearth and home in the Middle Pleistocene. *Journal of Anthropological Research*, *75*(3), 305–327.

Ladich, F., & Winkler, H. (2017). Acoustic communication in terrestrial and aquatic vertebrates. *Journal of Experimental Biology*, *220*(13), 2306–2317.

Lafrance, M., Worcester, L., & Burns, L. (2011). Gender and the Billboard Top 40 charts between 1997 and 2007. *Popular Music and Society*, *34*(5), 557–570.

Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and effect in biology revisited: Is Mayr's proximate-ultimate dichotomy still useful? *Science (New York, N.Y.)*, *334*(6062), 1512–1516.

Larsson, M. (2014). Self-generated sounds of locomotion and ventilation and the evolution of human rhythmic abilities. *Animal Cognition*, *17*(1), 1–14.

Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology*, *11*(4), 475–480.

Launay, J., Tarr, B., & Dunbar, R. I. M. (2016). Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology*, *122*(10), 779–789.

Leighton, G. M. (2017). Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proceedings of the Royal Society B: Biological Sciences*, *284* (1868), 20171508.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.

Lewis, J. (2013). A cross-cultural perspective on the significance of music and dance to culture and society. In M. A. Arbib (Ed.), *Language, music, and the brain* (pp. 45–66). MIT Press.

Lewis, J. (2017). Egalitarian social organization: The case of the Mbendjele BaYaka. In B. S. Hewlett (Ed.), *Hunter-Gatherers of the Congo basin* (pp. 249–274). Routledge.

Liberman, Z., Woodward, A. L., Sullivan, K. R., & Kinzler, K. D. (2016). Early emerging system for reasoning about the social nature of food. *Proceedings of the National Academy of Sciences*, *113*(34), 9480–9485.

Lindsay, W. R., Andersson, S., Bererhi, B., Höglund, J., Johnsen, A., Kvarnemo, C., … Edwards, S. V. (2019). Endless forms of sexual selection. *PeerJ*, *7*, e7988.

Livingstone, F. B. (1973). Did the australopithecines sing? *Current Anthropology*, *14*(1/2), 25–29.

Livingstone, R. S., & Thompson, W. F. (2009). The emergence of music from the theory of mind. *Musicae Scientiae*, *13*(2 Suppl.), 83–115.

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.

Lorenz, K. Z. (1966). Evolution of ritualization in the biological and cultural spheres. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *251*(772), 273–284.

Madison, G., Holmquist, J., & Vestin, M. (2018). Musical improvisation skill in a prospective partner is associated with mate value and preferences, consistent with sexual selection and parental investment theory: Implications for the origin of music. *Evolution and Human Behavior*, *39*(1), 120–129.

Maher, C. R., & Lott, D. F. (2000). A review of ecological determinants of territoriality within vertebrate species. *The American Midland Naturalist*, *143*(1), 1–29.

Maher, L. A., & Conkey, M. (2019). Homes for hunters? Exploring the concept of home at hunter-gatherer sites in Upper Paleolithic Europe and Epipaleolithic Southwest Asia. *Current Anthropology*, *60*(1), 91–137.

Marcus, G. F. (2012). Musicality: Instinct or acquired skill? *Topics in Cognitive Science*, *4*(4), 498–512.

Markham, A. C., Gesquiere, L. R., Alberts, S. C., & Altmann, J. (2015). Optimal group size in a highly social mammal. *Proceedings of the National Academy of Sciences*, *112*(48), 14882–14887.

Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *14*(2), 54–67.

Martínez, I., Rosa, M., Arsuaga, J.-L., Jarabo, P., Quam, R., Lorenzo, C., … Carbonell, E. (2004). Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences of the USA*, *101*(27), 9976–9981.

Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A., & Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia. *Current Biology*, *24*(6), 699–704.

Maynard Smith, J., & Harper, D. (2003). *Animal signals*. Oxford University Press.

Mayr, E. (1961). Cause and effect in biology: Kinds of causes, predictability, and teleology are viewed by a practicing biologist. *Science (New York, N.Y.)*, *134*(3489), 1501–1506.

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, *47*(2), 379–387.

McCoy, D. E., & Haig, D. (2020). Embryo selection and mate choice: Can "honest signals" be trusted? *Trends in Ecology & Evolution*, *35*(4), 308–318.

McDermott, J., & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition*, *94*(2), B11–B21.

McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception*, *23*(1), 29–59.

McDermott, J., & Hauser, M. (2007). Nonhuman primates prefer slow tempos but dislike music overall. *Cognition*, *104*(3), 654–668.

McDermott, J. H. (2012). Auditory preferences and aesthetics: Music, voices and everyday sounds. In *Neuroscience of preference and choice* (pp. 227–256). Elsevier.

McKenna, J. J. (1978). Biosocial functions of grooming behavior among the common Indian langur monkey (*Presbytis entellus*). *American Journal of Physical Anthropology*, *48*(4), 503–509.

McNeill, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Harvard University Press.

Mehr, S. A., Kotler, J., Howard, R. M., Haig, D., & Krasnow, M. M. (2017). Genomic imprinting is implicated in the psychology of music. *Psychological Science*, *28*(10), 1455–1467.

Mehr, S. A., & Krasnow, M. M. (2017). Parent-offspring conflict and the evolution of infant-directed song. *Evolution and Human Behavior*, *38*(5), 674–684.

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*, *366*(6468), 957–970.

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

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

Mehr, S. A., & Spelke, E. S. (2017). Shared musical knowledge in 11-month-old infants. *Developmental Science*, *21*(2), e12542.

Merker, B. (2000a). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 315–327). MIT Press.

Merker, B. (2000b). Synchronous chorusing and the origins of music. *Musicae Scientiae*, *3* (1 Suppl.), 59–73.

Merker, B., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*(1), 4–17.

Miller, G. (2000a). 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.

Miller, G. F. (2000b). *The mating mind: How sexual choice shaped the evolution of human nature*. Doubleday.

Miranda, E. R., Kirby, S., & Todd, P. (2003). On computational models of the evolution of music: From the origins of musical taste to the emergence of grammars. *Contemporary Music Review*, *22*(3), 91–111.

Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body*. Weidenfeld Nicolson.

Monboddo, J. B. L. (1774). *Of the origin and progress of language* (Vol. *1*, 2nd ed.). J. Balfour.

Morley, I. (2012). Hominin physiological evolution and the emergence of musical capacities. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 109–141). Oxford University Press.

Moser, C. J., Lee-Rubin, H., Bainbridge, C. M., Atwood, S., Simson, J., Knox, D., … Mehr, S. A. (2020). Acoustic regularities in infant-directed vocalizations across cultures. *BioRxiv*. https://doi.org/10.1101/2020.04.09.032995.

Mosing, M. A., Verweij, K. J. H., Madison, G., Pedersen, N. L., Zietsch, B. P., & Ullén, F. (2015). Did sexual selection shape human music? Testing predictions from the sexual selection hypothesis of music evolution using a large genetically informative sample of over 10,000 twins. *Evolution and Human Behavior*, *36*(5), 359–366.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS ONE*, *9*(2), e89642.

Musolf, K., Hoffmann, F., & Penn, D. J. (2010). Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*. *Animal Behaviour*, *79*(3), 757–764.

Nettl, B. (2015). *The study of ethnomusicology: Thirty-three discussions*. University of Illinois Press.

Newman, J. D. (2007). Neural circuits underlying crying and cry responding in mammals. *Behavioural Brain Research*, *182*(2), 155–165.

Niven, J. E. (2016). Neuronal energy consumption: Biophysics, efficiency and evolution. *Current Opinion in Neurobiology*, *41*, 129–135.

Niven, J. E., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *Journal of Experimental Biology*, *211*(11), 1792–1804.

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*(6), 1281–1296.

Norman-Haignere, S. V., Feather, J., Brunner, P., Ritaccio, A., McDermott, J. H., Schalk, G., & Kanwisher, N. (2019). Intracranial recordings from human auditory cortex reveal a neural population selective for musical song. *bioRxiv*. https://www.biorxiv.org/content/10.1101/696161v1.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science (New York, N.Y.)*, *314*(5805), 1560–1563.

O'Brien, D. M., Allen, C. E., Van Kleeck, M. J., Hone, D., Knell, R., Knapp, A., … Emlen, D. J. (2018). On the evolution of extreme structures: Static scaling and the function of sexually selected signals. *Animal Behaviour*, *144*, 95–108.

Orians, G. H. (2014). *Snakes, sunrises, and Shakespeare: How evolution shapes our loves and fears*. The University of Chicago Press.

Paley, W. (1803). *Natural theology: Or, evidences of the existence and attributes of the deity, collected from the appearances of nature*. R. Faulder.

Panksepp, J. (2009). The emotional antecedents to the evolution of music and language. *Musicae Scientiae*, *13*(2 Suppl.), 229–259.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.

Patel, A. D. (2017). Why doesn't a songbird (the European starling) use pitch to recognize tone sequences? The informational independence hypothesis. *Comparative Cognition & Behavior Reviews*, *12*, 19–32. Retrieved from http://comparative-cognition-and-behavior-reviews.org/2017/vol12_patel/.

Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science*, *2*(10), 150221.

Pearce, E., Launay, J., MacCarron, P., & Dunbar, R. I. M. (2017). Tuning in to others: Exploring relational and collective bonding in singing and non-singing groups over time. *Psychology of Music*, *45*(4), 496–512.

Pearce, E., Launay, J., van Duijn, M., Rotkirch, A., David-Barrett, T., & Dunbar, R. I. M. (2016). Singing together or apart: The effect of competitive and cooperative singing on social bonding within and between sub-groups of a university fraternity. *Psychology of Music*, *44*(6), 1255–1273.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*, *100*(1), 1–32.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, *33*(2), 185–191.

Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal of Human Genetics*, *25*(5), 625–630.

Persons, W. S., & Currie, P. J. (2019). Feather evolution exemplifies sexually selected bridges across the adaptive landscape. *Evolution*, *73*(9), 1686–1694.

Peters, J. (2014). The role of genomic imprinting in biology and disease: An expanding view. *Nature Reviews Genetics*, *15*(8), 517–530.

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception*, *28*(1), 3–14.

Pinker, S. (1997). *How the mind works*. Norton.

Pinker, S. (2012). The false allure of group selection. Retrieved from http://edge.org/conversation/the-false-allure-of-group-selection.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*(4), 707–727.

Pisor, A. C., & Surbeck, M. (2019). The evolution of intergroup tolerance in nonhuman primates and humans. *Evolutionary Anthropology*, *28*, 210–223.

Podlipniak, P. (2017). The role of the Baldwin Effect in the evolution of human musicality. *Frontiers in Neuroscience*, *11*, 452. doi: 10.3389/fnins.2017.00542.

Port, M., Hildenbrandt, H., Pen, I., Schülke, O., Ostner, J., & Weissing, F. J. (2020). The evolution of social philopatry in female primates. *American Journal of Physical Anthropology*, *173*, 397–410. doi: 10.1002/ajpa.24123.

Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, *31*(3), 157–175.

Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2011). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 601–609.

Quam, R. M., de Ruiter, D. J., Masali, M., Arsuaga, J.-L., Martínez, I., … Moggi-Cecchi, J. (2013). Early hominin auditory ossicles from South Africa. *Proceedings of the National Academy of Sciences*, *110*(22), 8847–8851.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, *5*(1), 3677.

Ravignani, A. (2018). Darwin, sexual selection, and the origins of music. *Trends in Ecology & Evolution*, *33*(10), 716–719.

Rebout, N., De Marco, A., Lone, J.-C., Sanna, A., Cozzolino, R., Micheletta, J., … Thierry, B. (2020). Tolerant and intolerant macaques show different levels of structural complexity in their vocal communication. *Proceedings of the Royal Society B: Biological Sciences*, *287* (1928), 20200439.

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: Synchrony, shared intentionality and cooperation. *PLoS ONE*, *8*(8), e71182.

Rendall, D., Cheney, D. L., & Seyfarth, R. M. (2000). Proximate factors mediating 'contact' calls in adult female baboons (*Papio cynocephalus ursinus*) and their infants. *Journal of Comparative Psychology*, *114*(1), 36–46.

Richman, B. (1993). On the evolution of speech: Singing as the middle term. *Current Anthropology*, *34*(5), 721–722.

Rodseth, L., Wrangham, R. W., Harrigan, A. M., Smuts, B. B., Dare, R., Fox, R., … Wolpoff, M. H. (1991). The human community as a primate society. *Current Anthropology*, *32*(3), 221–254.

Roederer, J. G. (1984). The search for a survival value of music. *Music Perception*, *1*(3), 350–356.

Roper, S. D., & Chaudhari, N. (2017). Taste buds: Cells, signals and synapses. *Nature Reviews. Neuroscience*, *18*(8), 485–497.

Rousseau, J.-J. (1781). *Essai sur l'origine des langues*. (J. H. Moran, Trans.), F. Ungar.

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*, *112*(29), 8987–8992.

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*(10), 831–836.

Schellenberg, E. G., Corrigall, K. A., Dys, S. P., & Malti, T. (2015). Group music training and children's prosocial skills. *PLOS ONE*, *10*(10), e0141449.

Schoenemann, P. T. (2006). Evolution of the size and functional areas of the human brain. *Annual Review of Anthropology*, *35*(1), 379–406.

Schruth, D., Templeton, C. N., & Holman, D. J. (2019). A definition of song from human music universals observed in primate calls. *bioRxiv*. Retrieved from https://www.biorxiv.org/content/10.1101/649459v1.

Schulkin, J. (2013). *Reflections on the musical mind: An evolutionary perspective*. Princeton University Press.

Schulkin, J., & Raglan, G. B. (2014). The evolution of music and human social capability. *Frontiers in Neuroscience*, *8*, 292.

Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary theory and the ultimate–proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science*, *6*(1), 38–47.

Searcy, W. A. (Ed.). (2019). Animal communication, cognition, and the evolution of language. *Animal Behaviour*, *151*, 203–205.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2008). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1656), 575–584.

Sergeant, D. C., & Himonides, E. (2019). Orchestrated sex: The representation of male and female musicians in world-class symphony orchestras. *Frontiers in Psychology*, *10*, 1760.

Seyfarth, R. M. (1977). A model of social grooming among adult female monkeys. *Journal of Theoretical Biology*, *65*(4), 671–698.

Seyfarth, R. M., & Cheney, D. L. (1984). Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature*, *308*(5959), 541–543.

Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature*, *479*(7372), 219–222.

Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability* (2nd ed.). Methuen.

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

Silk, J. B., & Kappeler, J. M. (2017). Sociality in primates. In D. R. Rubenstein & P. Abbot (Eds.), *Comparative social evolution* (pp. 253–83). Cambridge University Press.

Singh, M. (2018). The cultural evolution of shamanism. *Behavioral and Brain Sciences*, *41*, 1–62.

Sluming, V. A., & Manning, J. T. (2000). Second to fourth digit ratio in elite musicians: Evidence for musical ability as an honest signal of male fitness. *Evolution and Human Behavior*, *21*(1), 1–9.

Smith, D. R., & Dworkin, M. (1994). Territorial interactions between two *Myxococcus* species. *Journal of Bacteriology*, *176*(4), 1201–1205.

Smith, J. M., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*(5427), 15–18.

Snowdon, C. T. (2017). Vocal communication in family-living and pair-bonded primates. In R. Quam, M. Ramsier, R. R. Fay & A. N. Popper (Eds.), *Primate hearing and communication* (Vol. *63*, 1st ed., pp. 141–174). Springer International Publishing.

Soltis, J. (2004). The signal functions of early infant crying. *Behavioral and Brain Sciences*, *27*(04), 443–458.

Sosis, R. (2000). Costly signaling and torch fishing on Ifaluk atoll. *Evolution and Human Behavior*, *21*(4), 223–244.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Spencer, H. (1902). *Facts and comments* (*two essays*). D. Appleton and Co.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Blackwell.

Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, *8*(1), 40–46.

Stiner, M. (2019). *Human predators and prey mortality*. Routledge.

Stiner, M. C. (2013). An unshakable middle paleolithic? Trends versus conservatism in the predatory niche and their social ramifications. *Current Anthropology*, *54*(S8), S288–S304.

Stiner, M. C. (2017). Love and death in the stone age: What constitutes first evidence of mortuary treatment of the human body? *Biological Theory*, *12*(4), 248–261.

Swedell, L., & Plummer, T. (2019). Social evolution in Plio-Pleistocene hominins: Insights from hamadryas baboons and paleoecology. *Journal of Human Evolution*, *137*, 102667.
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*(10), 20150767.
Tarr, B., Launay, J., & Dunbar, R. I. M. (2014). Music and social bonding: "self-other" merging and neurohormonal mechanisms. *Frontiers in Psychology*, *5*(1096).
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.
Temperley, D. (2004). Communicative pressure and the evolution of musical styles. *Music Perception*, *21*(3), 313–337.
Thompson, M. E. (2013). Reproductive ecology of female chimpanzees. *American Journal of Primatology*, *75*(3), 222–237.
Tibbetts, E. A., Mullen, S. P., & Dale, J. (2017). Signal function drives phenotypic and genetic diversity: The effects of signalling individual identity, quality or behavioural strategy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1724), 20160347.
Tierney, A. T., Russo, F. A., & Patel, A. D. (2011). The motor origins of human and avian song structure. *Proceedings of the National Academy of Sciences*, *108*(37), 15510–15515.
Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Für Tierpsychologie*, *20*(4), 410–433.
Tobias, J. A., Sheard, C., Seddon, N., Meade, A., Cotton, A. J., & Nakagawa, S. (2016). Territoriality, social bonds, and the evolution of communal signaling in birds. *Frontiers in Ecology and Evolution*, *4*, 74. doi: 10.3389/fevo.2016.00074.
Todd, P. (2000). Simulating the evolution of musical behavior. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 361–388). MIT Press.
Todd, P. M., & Werner, G. M. (1999). Frankensteinian methods for evolutionary music composition. In N. Griffith & P. M. Todd (Eds.), *Musical networks: Parallel distributed perception and performance* (pp. 313–339). MIT Press.
Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). Oxford University Press.
Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140089.
Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, *7*(3), 289–296.
Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences*, *930*(1), 1–16.
Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993a). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, *16*(2), 193–211.
Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, *16*(3), 285–295.
Trivers, R. L. (1972). Parental investment and sexual selection. In B. G. Campbell (Ed.), *Sexual selection and the descent of man* (pp. 136–179). Aldine.
Trivers, R. L. (1974). Parent-offspring conflict. *American Zoologist*, *14*(1), 249–264.
Ubeda, F. (2008). Evolution of genomic imprinting with biparental care: Implications for Prader-Willi and Angelman syndromes. *PLoS Biology*, *6*(8), 1678–1692.
van den Broek, E. M. F., & Todd, P. M. (2009). Evolution of rhythm as an indicator of mate quality. *Musicae Scientiae*, *13*(2 Suppl.), 369–386.
van Doorn, G. S., & Weissing, F. J. (2006). Sexual conflict and the evolution of female preferences for indicators of male quality. *The American Naturalist*, *168*(6), 742–757.
Van Schaik, C. P. (1983). Why are diurnal primates living in groups? *Behaviour*, *87*(1/2), 120–144.
Wallis, N. J., & Blessing, M. E. (2015). Big feasts and small scale foragers: Pit features as feast events in the American Southeast. *Journal of Anthropological Archaeology*, *39*, 1–18.
Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Singing and social bonding: Changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior*, *37*(2), 152–158.
Wertz, A. E. (2019). How plants shape the mind. *Trends in Cognitive Sciences*, *23*(7), 528–531.
Wich, S. A., & Nunn, C. L. (2002). Do male "long-distance calls" function in mate defense? A comparative study of long-distance calls in primates. *Behavioral Ecology and Sociobiology*, *52*(6), 474–484.
Wiessner, P., & Schiefenhövel, W. (1998). *Food and the status quest: An interdisciplinary perspective*. Berghahn Books.
Wiley, R. H., & Wiley, M. S. (1977). Recognition of neighbors' duets by stripe-backed wrens *Campylorhynchus nuchalis*. *Behaviour*, *62*(1/2), 10–34.
Willems, E. P., & van Schaik, C. P. (2015). Collective action and the intensity of between-group competition in nonhuman primates. *Behavioral Ecology*, *26*(2), 625–631.
Williams, C. A., Beaudet, A. L., Clayton-Smith, J., Knoll, J. H., Kyllerman, M., Laan, L. A., … Wagstaff, J. (2006). Angelman syndrome 2005: Updated consensus for diagnostic criteria. *American Journal of Medical Genetics*, *140*(5), 413–418.
Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton University Press.
Wilson, D. E., & Reeder, D. M. (Eds.). (2005). *Mammal species of the world: A taxonomic and geographic reference* (Vols. *1–2*, 3rd ed.). Johns Hopkins University Press.
Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin & Review*, *23*(6), 1647–1659.
Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, *61*(6), 1203–1216.
Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5.
Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, *106*(7), 2468–2471.
Zahavi, A. (1975). Mate selection – a selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214.

Open Peer Commentary

Music, groove, and play

Richard D. Ashley

Department of Music Studies, Northwestern University, Evanston, IL 60208, USA.
r-ashley@northwestern.edu

doi:10.1017/S0140525X20001727, e61

Abstract

Savage et al. include groove and dance among musical features which enhance social bonds and group coherence. I discuss groove as grounded in structure and performance, and relate musical performance to play in nonhuman animals and humans. The interplay of individuals' contributions with group action is proposed as the common link between music and play as contributors to social bonding.

Social bonding is the primary focus of Savage et al.'s approach to "why music?" During their discussion of ancestral bonding mechanisms, they mention play (sects. 2.1 and 6.2) but only in passing. Play is, like music, a human universal, and is also found in other species. The literature on the nature of play across the animal kingdom (cf. Burghardt, 2005; Graham & Burghardt, 2010) parallels many of the questions and issues surrounding music's evolutionary purpose as discussed in these target articles. To connect play with music and social bonding, let us begin with the place of groove in Savage et al.'s framework.

In Figure 2 of Savage et al., note how musical features and mechanisms connect with one another through their inputs and outputs. At the level of "Musical Features," the box marked "Groove" is almost unique in that its influences on the system – through its connections to "Dance" – are bidirectional and at its own level; only the units at the level of "Proximate

Neurobiological Underpinnings" parallel this. Musical features suggested to subserve groove include entrainment to a recurring pattern of beats and larger rhythmic units, hierarchically organized (sect. 4.1, para. 2). These give rise to expectations which guide and govern responses to the music. Groove, then, is described as an emergent property of interactions of the predictable and the unpredictable in music; groove feeds into dance, which in turn serves group bonding and coherence. Dance is action, and Savage et al. focus on synchrony and shared, coordinated actions. But, let us consider how variable and asynchronous action also play a part in the construction of groove and provide a connection with social play.

The focus on synchronous group action in these articles deemphasizes the ways in which the variability of human performance – variability stemming from intention, habit, or random factors such as motor noise – allows the contribution of individuals to be audible even in relatively large group settings. Music, an auditory display of human action, leverages the rapidity of processing in the auditory system to enable the perception of the presence, action, and interaction of multiple performers, concurrent rather than turn-by-turn as in speech. Swift and omnidirectional, audition registers smaller and larger variations in sound coming from throughout the acoustic environment. Such variations may be at the compositional level (the main perspective in these target articles) or at the level of performance.

From the perspective of auditory scene analysis (Mehr et al., sect. 3.1 via Trainor, 2018; see also Bregman, 1994), listeners partition musical soundscapes – continuous, quickly-changing amplitudes – into features or events such as notes, lines, and motives. Spectral similarity/dissimilarity, frequency proximity/distance, and temporal proximity/distance are crucial low-level features the auditory system processes, along with learned and culturally-situated musical schemas. From a schematic-processing standpoint, predictability and variability of rhythmic patterning are primary sources of groove (e.g., Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014), exemplary of "levels of rhythmic complexity and expectation violation" (sect. 4.1, para. 2). Attention can, however, be captured and guided and expectation manipulated by other kinds of musical patterning. These include sequences of timbres, whether produced by one performer's drum kit or by an ensemble of different instrumental or vocal timbres (Ashley, 2014) as well as timespans that begin with relatively predictable elements before becoming more variable and then more predictable once more (Ashley, 2014; Danielsen, 2006).

The flux of fixedness and variability serve to differentiate individual performers' actions from one another and the sounding whole. Individuals' contributions may be detected from: chorusing, or not-quite unison in some musical line; distinctive timbres or musical figures identifying individual, discrete "sound sources"; and asynchrony in attack or spectral evolution, where joint action does not completely efface minute, audible differences between performers' actions. The ways in which performers are asynchronous with one another has been theorized as "participatory discrepancies" (Keil, 1987) and proposed as fundamental to emergent groove. Such divergences from synchrony are additional instances of "levels of rhythmic complexity and expectation violation," albeit at a "microtiming" temporal level which is faster than that of musical notation. Performance, in addition to composition, creatively and continuously modulates music as auditory display, with the actions and interactions of performers both predictable and unexpected – essential to groove and musical engagement beyond groove. There's a reason why we speak of "playing" music: it captures the nature of creative action between individuals well.

And so, I turn again to some suggestions of parallels between music and play. The social/group outcomes in Savage et al.'s framework parallel those proposed by researchers studying social play in nonhuman species. Theories of the origins and functions of animal play are varied (Bekoff & Byers, 1998; Burghardt, 2005) and our consideration here is, perforce, limited. We note that nonhuman species' actions in social play have been proposed as social bonding mechanisms (Bekoff, 1984). One prominent theory of Spinka, Newberry, and Bekoff (2001) draws intriguingly close to Savage et al.'s framework, proposing that social play in animals trains organisms for *responding to the unexpected*. In playful interaction, animals find themselves not in complete control and so Spinka and colleagues propose that play "enhances the ability of animals to cope emotionally with unexpected situations," is "emotionally exciting … and rewarding, maybe even pleasurable, while at the same time being relaxed … producing the complex emotional state that is referred to as 'having fun' in human folk psychology." Finally, they consider play a "cognitively demanding activity" which "requires frequent and rapid assessment and reassessment of qualitatively different situations" (pp. 144–145). These descriptions fit human musicking as well as they do animal play.

Although one may seek insights into evolutionary processes by comparative studies, it seems well to conclude by connecting human play to musicking. A recent review proposes that "What might be distinctive about human play is that people not only exploit their existing knowledge about how to explore … but also explore new ways to explore" (Chu & Schulz, 2020, p. 332). The myriad ways humans have made, and continue to make, so many kinds of music from a handful of common starting points richly testifies to such explorations.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Ashley, R. (2014). Expressiveness in funk. In D. Fabian, R. Timmers & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 154–169). Oxford University Press.

Bekoff, M. (1984). Social play behavior. *Bioscience*, *34*(4), 228–233.

Bekoff, M., & Byers, J. A. (Eds.). (1998). *Animal play: Evolutionary, comparative and ecological perspectives*. Cambridge University Press.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.

Burghardt, G. M. (2005). *The genesis of animal play: Testing the limits*. MIT Press.

Chu, J., & Schulz, L. E. (2020). Play, curiosity, and cognition. *Annual Review of Developmental Psychology*, *2*, 317–43.

Danielsen, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament*. Wesleyan University Press.

Graham, K. L., & Burghardt, G. M. (2010). Current perspectives on the biological study of play: Signs of progress. *The Quarterly Review of Biology*, *85*(4), 393–418.

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, *2*(3), 275–283.

Spinka, M., Newberry, R. C., & Bekoff, M. (2001). Mammalian play: Training for the unexpected. *The Quarterly Review of Biology*, *76*(2), 141–168.

Trainor, L. J. (2018). The origins of music: Auditory scene analysis, evolution, and culture in musical creation. In H. Honing (Ed.), *The origins of musicality* (pp. 81–112). MIT Press.

Witek, M. A., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*(4), e94446.

Musicality was not selected for, rather humans have a good reason to learn music

Shir Atzil and Lior Abramson

Department of Psychology, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel.
shir.atzil@mail.huji.ac.il; https://scholars.huji.ac.il/atzillab
lior.abramson@mail.huji.ac.il

doi:10.1017/S0140525X2000134X, e62

Abstract

We propose that not social bonding, but rather a different mechanism underlies the development of musicality: being unable to survive alone. The evolutionary constraint of being dependent on other humans for survival provides the ultimate driving force for acquiring human faculties such as sociality and musicality, through mechanisms of learning and neural plasticity. This evolutionary mechanism maximizes adaptation to a dynamic environment.

Both Mehr et al. and Savage et al. agree that music supports social behaviors. Although Mehr and colleagues propose that musicality evolved to support specific behaviors, Savage et al. construct a general framework, suggesting that musicality evolved to promote sociality at large. These ideas rely on the assumption that sociality, in itself, was selected for in evolution as an inborn faculty. However, in contrary to a cumbersome evolutionary solution implementing inborn faculties, the brain could have evolved with one ultimate feature: to be able to wire ad hoc to the environment, optimizing survival in any cultural nice. Such an evolutionary plan is sufficient to ensure the acquisition of profound human characteristics through learning if they are relevant for survival, including (but not limited to) sociality (Atzil, Gao, Fradkin, & Barrett, 2018) and musicality.

Learning is guided by the need of an organism to optimize the internal milieu (Pezzulo, Rigoli, & Friston, 2015). *Allostasis* is the ongoing adjustment of the internal milieu necessary for survival, growth, and reproduction (Sterling, 2012). In social animals, allostasis is fundamentally social because of one evolutionary constraint: offspring cannot survive alone (Atzil et al., 2018). They depend on a conspecific, a caregiver, to regulate their allostatic processes, including energy expenditure, temperature (Winberg, 2005), immunity (Arrieta, Stiemsma, Amenyogbe, Brown, & Finlay, 2014), and arousal (Cirelli, Jurewicz, & Trehub, 2019). By providing these needs, the caregiver's allostatic support is rewarding and reinforces bonding. Moreover, infants learn to regulate allostasis via social interactions (Atzil & Gendron, 2017). At first, social regulation is physical, relying on touch (Feldman, Keren, Gross-Rozval, & Tyano, 2004) and vocalizations (Cirelli et al., 2019). With development, cultural constructs are prone to be learned and become salient and rewarding as well. Gradually, guided by allostatic motivation, young humans acquire social behaviors and concepts, and become social experts (Atzil et al., 2018). Consequently, any behavior or concept secondary to sociality, such as musicality, would be further learned rather than genetically selected for.

Not only social stimuli, but rather any concrete or abstract stimulus that impacts allostasis will be stabilized through learning. These include natural rewards such as food, but in humans also abstract cultural products such as language, art, religion, money, and most relevantly-music. These abstract constructs become meaningful to humans' survival through learning and stabilize across societies and generations, even though they are not heritable in a domain-specific manner. Yet, music is special because it is dual-valued, and holds both concrete physical faculties and abstract information, which was culturally imbued. For example, an anthem, which contains both concrete sensory input and abstract information about group identity, is extremely powerful in regulating humans (Konečni, Wanic, & Brown, 2007). This dual concrete–abstract effect of music on allostasis aligns with Savage et al.'s idea that musicality is more powerful than grooming (concrete stimulus) or language (abstract stimulus). It is also in line with Mehr et al.'s hypothesis on the role of musicality in parental care. As infants develop and acquire abstract knowledge, musicality provides a developmental bridge between concrete and abstract regulation.

Savage et al. propose that musicality and sociality share an underlying mechanism of prediction and reward. However, prediction and reward are domain-general processes of learning. In any context, the brain learns statistical regularities in the environment to actively generate predictions (Clark, 2013; Friston, 2005). Prediction is crucial for survival and partly what separates the terms *allostasis* from *homeostasis* (Sterling, 2012). Although in homeostasis a deviation from a set-point elicits reaction, allostasis is a process where the brain integrates prior knowledge to anticipate upcoming needs, preparing the organism to react in advance (McEwen & Wingfield, 2003). Thus, the evolutionary development of domain-general functions such as prediction and reward, can underlie the learning of any statistical pattern relevant for predicting allostasis. In social context, infants rely on statistical regularities of parental care for allostasis. For example, for a hungry infant, the parent is a reliable predictor for an upcoming increase in glucose levels, which is in turn rewarding and reinforces further interaction. Similarly, listening to a fast bit tune predicts changes in levels of arousal (Bernardi, Porta, & Sleight, 2006). Thus, beyond the idea that music is rewarding because it improves social predictions, suggested by Savage et al., music is rewarding because it directly improves allostatic predictions. In this sense, sociality and musicality can be considered cultural patterns, learned because they are useful to improve allostatic predictions.

Accordingly, the neural mechanism underlying reward and prediction, which process music and social information, is not selective but rather underlies every motivated behavior that is relevant for allostasis (Kleckner et al., 2017; Seth, 2013), such as feeding (Feldstein Ewing et al., 2017), avoiding pain (Scott, Heitzeg, Koeppe, Stohler, & Zubieta, 2006), and seeking drugs (Volkow & Morales, 2015). In this domain-general circuit, cortico-limbic structures rely on prior knowledge to propagate prediction signals to sensory-motor cortices to regulate perception-action (Den Ouden, Daunizeau, Roiser, Friston, & Stephan, 2010; Pezzulo et al., 2015). Predictions also projects downstream to the body via the hypothalamus-pituitary-adrenal (HPA) axis and autonomic nervous system to regulate allostasis (Barrett & Simmons, 2015; Pezzulo et al., 2015) (Fig. 1). Although prediction and reward are domain-general functions

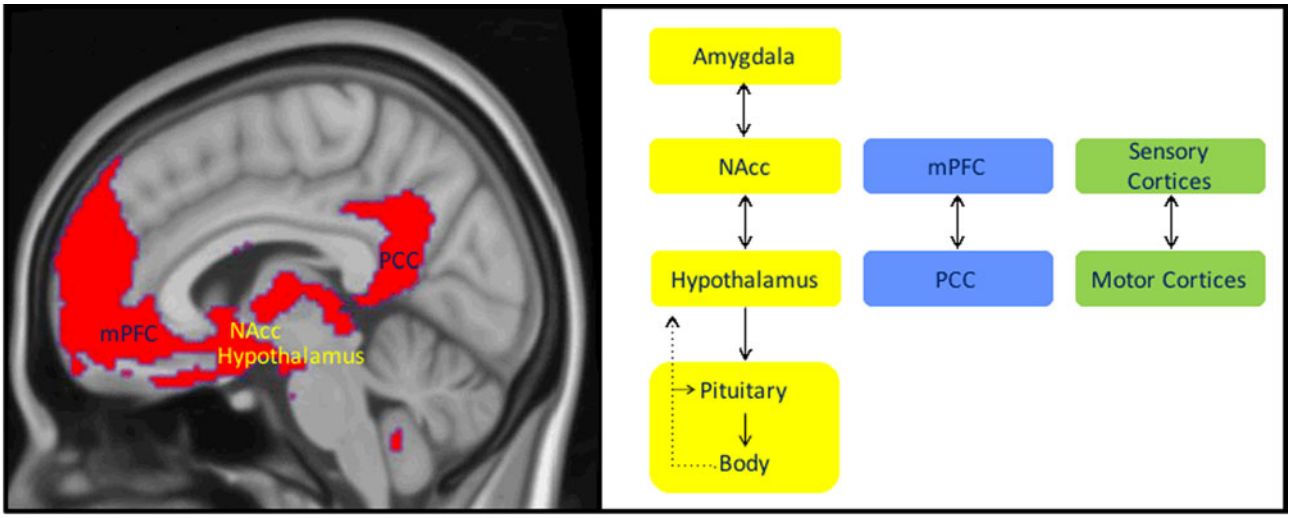


Figure 1. (Atzil & Abramson) A domain-general neural system for motivated behaviors in humans. This domain-general neural system contains midline cortices in the medial prefrontal cortex (mPFC) (including the anterior cingulate cortex [ACC]), and posterior cingulate cortex (PCC); subcortical structures of the amygdala, nucleus accumbens (NAcc) and hypothalamus; and sensory-motor cortices. The mPFC and PCC represent abstract concepts (Baldassano, Hasson, & Norman, 2018), like a human or a song (in blue). These associative cortices, along with other cortico-limbic structures, rely on past experience to issue predictions to sensory-motor cortices (in green), to regulate behavior and perception (Den Ouden et al., 2010; Pezzulo et al., 2015). They also project downstream to the body, predicting and regulating allostasis (Barrett & Simmons, 2015) (in yellow). This circuitry is plastic and can associate any learned concept or behavior, which is relevant for reward and prediction, to allostasis regulation in the body. In that sense, this circuitry is not exclusively related to sociality or musicality, but rather is a domain-general mechanism, which can be wired to support any motivated behavior with allostatic consequences.

that are relevant for survival in any niche, the particulars of a certain niche are dynamic. Accordingly, an adaptive brain is one that can learn them ad hoc. Given the plasticity of the brain, every statistical regularity that allows reliable prediction about allostasis will be represented in the cortex and neurally associated with allostasis through this circuit. Because social information and music are extremely useful for allostatic predictions, they efficiently attain this neural circuitry.

To conclude, music is a cultural product. Because of its rewarding and predictive values, humans are repeatedly reinforced to learn it and to pass it on to the next generations, not necessarily through specific genetic adaptations. All animals rely on prediction and reward to survive. Humans developed an especially wide and complex range of cues in multiple levels of abstraction that support prediction of allostasis. Although musicality and sociality rely on shared mechanisms, it does not mean that these mechanisms evolved specifically for them. Alternatively, a rather simple evolutionary constraint of social dependency for survival, along with neural plasticity, is sufficient to ensure the cross-generation transmission of both sociality and musicality. We propose an alternative mechanism, by which musicality and sociality are dynamic and were thus not genetically selected for in a domain-specific manner. Instead, a domain-general mechanism of reward and prediction was selected for to improve survival by being flexible and wired to the environment. This ensures that behavior is culturally sculptured in an ongoing process that maximizes adaptation. This theoretical approach provides a parsimonious explanation for the role of music in human culture.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Arrieta, M. C., Stiemsma, L. T., Amenyogbe, N., Brown, E., & Finlay, B. (2014). The intestinal microbiome in early life: Health and disease. *Frontiers in Immunology*, *5*, 427. https://doi.org/10.3389/fimmu.2014.00427.

Atzil, S., Gao, W., Fradkin, I., & Barrett, L. F. (2018). Growing a social brain. *Nature Human Behaviour*, *2*(9), 624–636. https://doi.org/10.1038/s41562-018-0384-6.

Atzil, S., & Gendron, M. (2017). Bio-behavioral synchrony promotes the development of conceptualized emotions. *Current Opinion in Psychology*, *17*, 162–169. https://doi.org/10.1016/j.copsyc.2017.07.009.

Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of real-world event schemas during narrative perception. *Journal of Neuroscience*, *38*(45), 9689–9699. https://doi.org/10.1523/JNEUROSCI.0251-18.2018.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*(7), 419–429. https://doi.org/10.1038/nrn3950.

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*, *92*(4), 445–452. https://doi.org/10.1136/hrt.2005.064600.

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*(7), 1213–1220. https://doi.org/10.1162/jocn_a_01402.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477.

Den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., & Stephan, K. E. (2010). Striatal prediction error modulates cortical coupling. *Journal of Neuroscience*, *30*(9), 3210–3219. https://doi.org/10.1523/JNEUROSCI.4458-09.2010.

Feldman, R., Keren, M., Gross-Rozval, O., & Tyano, S. (2004). Mother–child touch patterns in infant feeding disorders: Relation to maternal, child, and environmental factors. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(9), 1089–1097. https://doi.org/10.1097/01.chi.0000132810.98922.83.

Feldstein Ewing, S. W., Claus, E. D., Hudson, K. A., Filbey, F. M., Yakes Jimenez, E., Lisdahl, K. M., & Kong, A. S. (2017). Overweight adolescents' brain response to sweetened beverages mirrors addiction pathways. *Brain Imaging and Behavior*, *11*(4), 925–935. https://doi.org/10.1007/s11682-016-9564-z.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622.

Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., … Feldman Barrett, L. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Human Behaviour*, *1*(5), 1–14. https://doi.org/10.1038/s41562-017-0069.

Konečni, V. J., Wanic, R. A., & Brown, A. (2007). Emotional and aesthetic antecedents and consequences of music-induced thrills. *American Journal of Psychology*, *120*(4), 619–643. https://doi.org/10.2307/20445428.

McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, *43*(1), 2–15. https://doi.org/10.1016/S0018-506X(02)00024-7.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35. https://doi.org/10.1016/j.pneurobio.2015.09.001.

Scott, D. J., Heitzeg, M. M., Koeppe, R. A., Stohler, C. S., & Zubieta, J. K. (2006). Variations in the human pain stress experience mediated by ventral and dorsal basal ganglia dopamine activity. *Journal of Neuroscience*, *26*(42), 10789–10795. https://doi.org/10.1523/JNEUROSCI.2577-06.2006.

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565–573. https://doi.org/10.1016/j.tics.2013.09.007.

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology and Behavior*, *106*(1), 5–15. https://doi.org/10.1016/j.physbeh.2011.06.004.

Volkow, N. D., & Morales, M. (2015). The brain on drugs: From reward to addiction. *Cell*, *162*(4), 712–725. https://doi.org/10.1016/j.cell.2015.07.046.

Winberg, J. (2005). Mother and newborn baby: Mutual regulation of physiology and behavior – A selective review. *Developmental Psychobiology*, *47*(3), 217–229. https://doi.org/10.1002/dev.20094.

Social bonding and music: Evidence from lesions to the ventromedial prefrontal cortex

Amy M. Belfi

Department of Psychological Science, Missouri University of Science and Technology, Rolla, MO 65409, USA.
amybelfi@mst.edu; www.amybelfi.com

doi:10.1017/S0140525X2000103X, e63

Abstract

The music and social bonding (MSB) hypothesis suggests that damage to brain regions in the proposed neurobiological model, including the ventromedial prefrontal cortex (vmPFC), would disrupt the social and emotional effects of music. This commentary evaluates prior research in persons with vmPFC damage in light of the predictions put forth by the MSB hypothesis.

Although both target articles develop compelling theories on the evolutionary origins of music, the current commentary will focus on the neurobiological model underlying the music and social bonding (MSB) hypothesis. In their model, the authors propose key neural systems underlying the social bonding function of music. Accordingly, they predict that individuals with disruptions to these neural systems may show altered responses to music. One such prediction is that "special populations with high sociability may respond well to musical features especially when coupled with social stimuli" (sect. 5.4, para. 2). Although it is not exactly clear what "respond well" means in this case (let's assume this means that individuals with heightened sociability will also show heightened social and/or emotional responsiveness to music), one could also consider the converse prediction, that individuals with abnormally low social functioning may not respond emotionally to music.

When thinking of populations with disrupted social functioning, one can look toward the neuropsychological literature: A rich history of literature demonstrates that individuals with damage to the ventromedial prefrontal cortex (vmPFC) show substantial changes in their social and affective behaviors (for review, see Schneider & Koenigs, 2017). The vmPFC is highlighted by Savage and colleagues as a critical component of the neuroanatomical model underlying the MSB hypothesis (see Fig. 3). Because the authors claim that this system is "involved causally in the link between music and social bonding" (sect. 4.2, para. 2), a useful approach might be to consider whether the behavior of individuals with damage to the vmPFC aligns with its proposed role underlying the social function of music.

To address this point, the current commentary will focus on research investigating the behavior of individuals with lesions to the vmPFC and consider these findings with regard to the predictions put forth by the MSB hypothesis. To start, substantial prior research has indicated that individuals with damage to the vmPFC have deficits in various socioemotional domains, including emotion and reward, social behavior, and value-based decision-making (Abel et al., 2016; Jenkins et al., 2014). For example, individuals with vmPFC damage show both functional and structural disruptions to the reward network, including decreased nucleus accumbens volumes and reduced activity in the ventral striatum during the anticipation of reward (Pujara, Philippi, Motzkin, Baskaya, & Koenigs, 2016). In addition, individuals with vmPFC damage display deficits in social cognition and behavior, including theory of mind (Leopold et al., 2012), moral decision-making (Koenigs et al., 2007), and empathy (Beadle, Paradiso, & Tranel, 2018).

The authors of the MSB hypothesis imply that individuals with disrupted social functioning, such as persons with vmPFC damage, should also display considerable deficits in the socioemotional aspects of music. Some evidence does point toward impaired music processing in individuals with vmPFC damage. For example, persons with vmPFC damage show reduced skin conductance responses in response to music, indicating a lack of physiological arousal (Johnsen, Tranel, Lutgendorf, & Adolphs, 2009). Individuals with vmPFC damage also produce music-evoked autobiographical memories that are less episodically rich, as compared to memories evoked by visual cues (Belfi, Karlan, & Tranel, 2018). Other research in patients with behavioral variant frontotemporal dementia (bvFTD) suggests that the vmPFC plays a critical role in "mentalizing" in music. That is, patients with bvFTD have difficulties attributing complex mental states to music (e.g., identifying whether music sounds "dreamy," "dreading," or "adventurous"). Interestingly, this deficit in musical emotion attribution was correlated with deficits in social inference and empathy more broadly, and was associated with gray matter loss in the vmPFC (Downey et al., 2013). Taken together, this research suggests that individuals with damage to the vmPFC have both deficits in social cognition and affect more broadly, as well as in their emotional responses to music.

Although the aforementioned research suggests that the vmPFC plays a critical role underlying the social and emotional aspects of musicality, it does not seem that vmPFC damage leads to a complete disregard for music. That is, individuals with vmPFC damage do not exhibit higher levels of musical anhedonia (a selective lack in pleasure from music). No published cases of acquired musical anhedonia include damage to the vmPFC (Griffiths, Warren, Dean, & Howard, 2004; Hirel et al., 2014; Mazzoni et al., 1993; Satoh, Nakase, Nagata, & Tomimoto, 2011; Satoh et al., 2016) and a large-scale lesion study did not identify any persons with vmPFC damage as musically anhedonic (Belfi, Evans, Heskje, Bruss, & Tranel, 2017). This may be explained by the fact that acquired musical anhedonia is typically characterized by self-reported changes in emotional responses to music. That is, it may be the case that individuals with vmPFC damage do not *self-report* changes in musical reward, although they might *experience* reduced musical reward. This coincides with evidence indicating that individuals with vmPFC damage show reduced physiological responsiveness, but not reductions in subjective feelings when listening to music (Johnsen et al., 2009).

To conclude, the authors of the MSB hypothesis highlight the role of the reward system, including the vmPFC, as a critical network supporting the coevolved processes of social bonding and music. This model subsequently predicts that damage to structures in this network would disrupt the social bonding effects of music. Although not testing the social bonding effects of music directly, some research indicates that damage to the vmPFC is associated with deficits in socioemotional aspects of music, including physiological responsiveness, music-evoked autobiographical memories, and attributing mental states to music. When taken together, these results provide tentative support for the role of the vmPFC in the socioemotional components of musicality. Although this seems promising for the neurobiological model underlying the MSB hypothesis, there is one final but important point to note – the fact that two cognitive processes are disrupted from damage to the same region (i.e., vmPFC damage is associated with deficits in both social behavior *and* emotional responses to music) does not necessarily speak directly to the relationship between those processes. Therefore, although it should be interpreted cautiously, this research provides initial evidence for similar neural systems underlying both social behavior and emotional responses to music.

Financial support. This research was funded by the Missouri S&T Center for Biomedical Research.

Conflict of interest. None.

References

Abel, T. J., Manzel, K., Bruss, J., Belfi, A. M., Howard, M. A., & Tranel, D. (2016). The cognitive and behavioral effects of meningioma lesions involving the ventromedial prefrontal cortex. *Journal of Neurosurgery*, *124*(6), 1568–77. https://doi.org/10.3171/2015.5.JNS142788.

Beadle, J. N., Paradiso, S., & Tranel, D. (2018). Ventromedial prefrontal cortex is critical for helping others who are suffering. *Frontiers in Neurology*, *9*, 1–13. https://doi.org/10.3389/fneur.2018.00288.

Belfi, A. M., Evans, E., Heskje, J., Bruss, J., & Tranel, D. (2017). Musical anhedonia after focal brain damage. *Neuropsychologia*, *97*, 29–37.

Belfi, A. M., Karlan, B., & Tranel, D. (2018). Damage to the medial prefrontal cortex impairs music-evoked autobiographical memories. *Psychomusicology: Music, Mind, and Brain*, *28*, 201–208.

Downey, L. E., Blezat, A., Nicholas, J., Omar, R., Golden, H. L., Mahoney, C. J., … Warren, J. D. (2013). Mentalising music in frontotemporal dementia. *Cortex*, *49*, 1844–55.

Griffiths, T. D., Warren, J. D., Dean, J. L., & Howard, D. (2004). "When the feeling's gone": A selective loss of musical emotion. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*, 341–345.

Hirel, C., Leveque, Y., Deiana, G., Richard, N., Cho, T.-H., Mechtouff, L., & Nighoghossian, N. (2014). Acquired amusia and musical anhedonia. *Revue Neurologique*, *170*, 536–540. https://doi.org/doi:10.1016/j.neurol.2014.03.015.

Jenkins, L. M., Andrewes, D. G., Nicholas, C. L., Drummond, K. J., Moffat, B. A., Phal, P., … Kessels, R. P. C. (2014). Social cognition in patients following surgery to the prefrontal cortex. *Psychiatry Research – Neuroimaging*, *224*(3), 192–203. https://doi.org/10.1016/j.pscychresns.2014.08.007.

Johnsen, E. L., Tranel, D., Lutgendorf, S., & Adolphs, R. (2009). A neuroanatomical dissociation for emotion induced by music. *International Journal of Psychophysiology*, *72*, 24–33. https://doi.org/10.1016/j.ijpsycho.2008.03.011.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*, 908–911. https://doi.org/10.1038/nature05631.

Leopold, A., Krueger, F., Dal monte, O., Pardini, M., Pulaski, S. J., Solomon, J., & Grafman, J. (2012). Damage to the left ventromedial prefrontal cortex impacts affective theory of mind. *Social Cognitive and Affective Neuroscience*, *7*(8), 871–880. https://doi.org/10.1093/scan/nsr071.

Mazzoni, M., Moretti, P., Pardossi, L., Vista, M., Muratorio, A., & Puglioli, M. (1993). A case of music imperception. *Journal of Neurology, Neurosurgery & Psychiatry*, *56*, 322–322.

Pujara, M. S., Philippi, C. L., Motzkin, J. C., Baskaya, M. K., & Koenigs, M. (2016). Ventromedial prefrontal cortex damage is associated with decreased ventral striatum volume and response to reward. *Journal of Neuroscience*, *36*(18), 5047–5054. https://doi.org/10.1523/JNEUROSCI.4236-15.2016.

Satoh, M., Kato, N., Tabei, K., Nakano, C., Abe, M., Fujita, R., … Kondo, K. (2016). A case of musical anhedonia due to right putaminal hemorrhage: A disconnection syndrome between the auditory cortex and insula. *Neurocase*, *22*, 518–25. https://doi.org/10.1080/13554794.2016.1264609.

Satoh, M., Nakase, T., Nagata, K., & Tomimoto, H. (2011). Musical anhedonia: Selective loss of emotional experience in listening to music. *Neurocase*, *17*, 410–417.

Schneider, B., & Koenigs, M. (2017). Human lesion studies of ventromedial prefrontal cortex. *Neuropsychologia*, *107*, 84–93. https://doi.org/10.1016/j.neuropsychologia.2017.09.035.

Evolutionary linguistics can help refine (and test) hypotheses about how music might have evolved

Antonio Benítez-Burraco

Department of Spanish, Linguistics, and Theory of Literature (Linguistics), Faculty of Philology, University of Seville, 41004-Seville, Spain.
abenitez8@us.es
http://antoniobenitez.wix.com/benitez-burraco

doi:10.1017/S0140525X20000680, e64

Abstract

Both the music and social bonding (MSB) hypothesis and the music as a credible signal hypothesis emerge as solid views of how human music and human musicality might have evolved. Nonetheless, both views could be improved (and tested in better ways) with the consideration of the way in which human language(s) might have evolved under the effects of our self-domestication.

Savage and colleagues provide a compelling argument in favor of the coevolution of music and musicality via iterative niche construction driven by their positive effect on human social bonding (the music and social bonding [MSB] hypothesis). By contrast, Mehr and colleagues argue that music evolved to provide increasingly sophisticated credible signals that are needed to cope with progressively complex social conflicts of interest, like those resulting from multi-level social organization or higher levels of (allo) parental investment. These two hypotheses are presented as somehow irreconcilable. This commentary paper brings language evolution to the forefront with the aim of discussing the plausibility of both views, suggesting potential ways of improving and

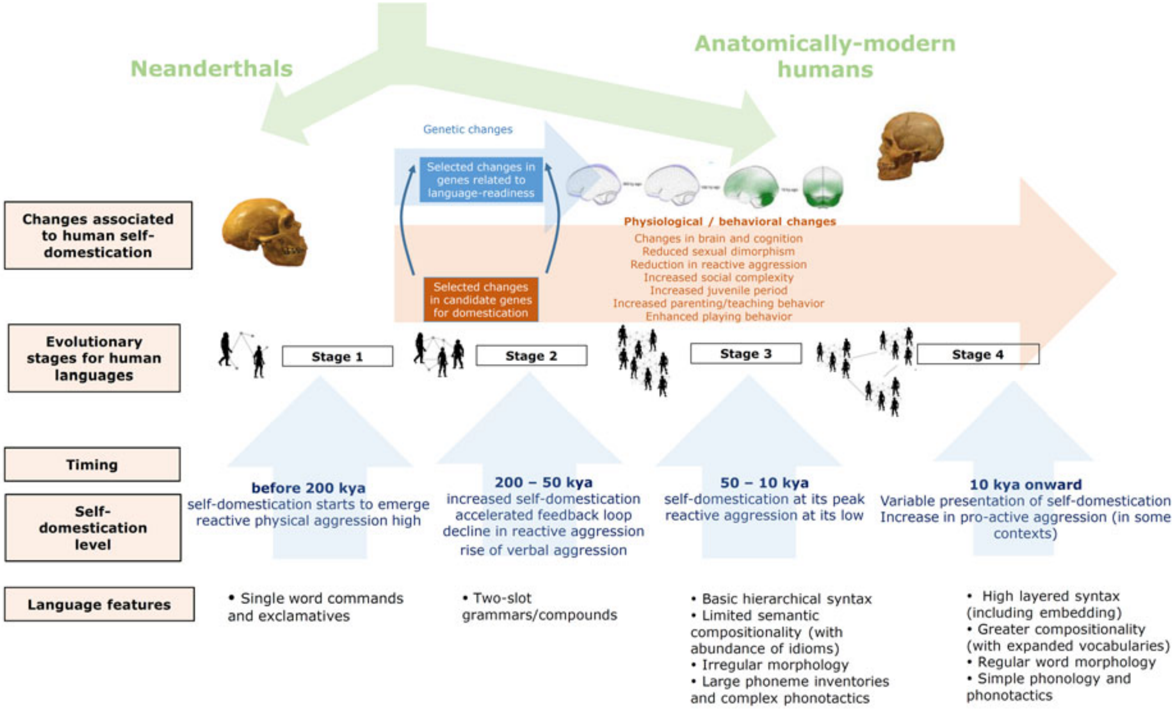


Figure 1. (Benítez-Burraco) A graphical overview of a model of language evolution under the effects of human self-domestication (see Benítez-Burraco, 2017; Benítez-Burraco & Kempe, 2018; Benítez-Burraco & Progovac, 2020; Langley, Benítez-Burraco, & Kempe, 2020; Progovac & Benítez-Burraco, 2019 for details).

testing them, and eventually, reconcile them under the light of new views of human evolution.

Recently, the self-domestication hypothesis of human evolution has (re)emerged as a promising account of many of our species-specific traits. According to this view, we evolved similarly to domesticated varieties of mammals. But, although animal domestication usually results from selection for tameness, our domestication might have been triggered by external factors such as changes in our foraging ecology, the rise of community living, or the advent of co-parenting (hence, *self*-domestication) (Hare, Wobber, & Wrangham, 2012; Pisor & Surbeck, 2019). These factors are hypothesized to have impacted the neurobiological mechanisms controlling aggression, resulting in individuals that were less emotionally reactive and more tolerant for strangers, with these changes ultimately favoring the emergence of many of our distinctive features, including our enhanced social cognition, increased cooperation, and extended social networks, and eventually, our sophisticated culture and advanced technology (see Hare, 2017 for details). Our own research supports the view that self-domestication could account as well for most aspects of the cultural niche that enables the sophistication of language structure and use via a cultural mechanism, mostly through the potentiation of the cognitive and behavioral abilities involved in language learning and use. We have equally argued for an intense feedback loop (the "virtuous spiral" mentioned by Savage and colleagues) between self-domestication processes, grammar complexity, and pragmatics sophistication, and ultimately, for a gradual co-evolution of human language (and human languages) under the effects of self-domestication (Fig. 1).

Most of the evidence discussed by Savage and colleagues supports a potential link between the evolution of musicality and self-domestication mechanisms, particularly, regarding the biological underpinnings. To mention just one instance, they suggest that the positive effect of musicality on social boding might result in part from the activation of the dopaminergic reward system (their Fig. 3a). Interestingly, changes in the dopamine systems have been regularly documented in domesticated animals (Komiyama et al., 2014; Sato et al., 2020). Evidence of selection of pathways related to dopaminergic synapse have been found in European samples during the past 6,000 years (Chekalin et al., 2019), a time period when genes involved in animal domestication have been selected too (Benítez-Burraco, et al., in press). Overall, including self-domestication in the equation could help the authors address what they anticipate the key criticism to the MSB hypothesis, namely, the degree to which the evolution of musicality and social bonding are causally linked. One could say that it was our self-domestication, via its inhibitory impact on reactive aggression, that set in motion the "virtuous spiral" involving social bonding and music. Similarly, the cognitive and behavioral outcomes of self-domestication might have stimulated the evolution of music as a credible signal for cooperating between groups (coalitions) and inside groups (child–adult relationships), as suggested by Mehr and colleagues. In fact, these authors acknowledge that "music does not *directly* cause social cohesion: rather, it signals existing social cohesion that was obtained by other means" (sect. 4.2.1, para. 14). Self-domestication could be one (or the most important) of such means, as the reduced levels of reactive aggression brought about by self-domestication might have favored longer, more frequent, and more diverse contacts with others, including strangers and caregivers, setting the scene for the selection of music as a credible signal for cooperative exchanges. Overall, considering self-domestication forces could help reconcile these two divergent views of music evolution. One could hypothesize that music was initially selected because of its contribution to the enhanced social bonding brought about by self-domestication (stage 2 in Fig. 1), and later, because of its role as a credible signal, once human groups became larger and more complex as self-domestication reached its peak (stages 3 and 4 in Fig. 1).

The evolutionary scenario sketched above could help as well improve the less developed aspect of both papers, specifically, how the different types of music emerged through a cultural mechanism. Both papers expect some sort of link between music complexity and social complexity. Savage and colleagues expect music to be more effective as a social bond mechanism in bigger, more complex human groups, with larger and more hierarchical societies preferring "presentational" music over "participatory" music. Similarly, Mehr and colleagues (seem to) expect more elaborated forms of music in multi-level societies. Considering again the domain of evolutionary linguistics, and briefly summarizing a vast body of research (e.g., Bolender, 2007; Lupyan & Dale, 2010; Nettle, 2012; Trudgill, 2011; Wray & Grace, 2007), one finds that the languages spoken by small, isolated human groups forming close-knit social networks exhibit quite the opposite structural features (from grammar, to vocabulary, to sound patterns) to the languages spoken by large human groups forming extensive and complex social networks with higher rates of cultural exchange (Fig. 1; stages 3 and 4, respectively). These opposite features seemingly result from the dissimilar amount of knowledge shared by speakers (i.e., the common ground), in turn a consequence of the different nature of the social bonds they maintain, in turn a consequence of the levels of reactive and proactive aggression, an aspect that is at the core of the self-domestication hypothesis (see also Fig. 1). Consequently, our understanding of the cultural evolution of music could benefit from the comparative analysis of the structural and functional features of the types of music produced by these two main types of human societies, looking for the sort of correlations (and causality) found in languages. If musical compositions parallel what we observe in languages, we would expect that the musical compositions created by (let's say) English-speaking or Chinese-speaking societies, which are outstanding examples of languages of the second type, exhibit a more elaborated structure, put more information in the sheet, and are thus more "understandable" by people with different cultural backgrounds, that the music created by (let's say) the Tsimane people (McDermott, Schultz, Undurraga, & Godoy, 2016).

In summary, we regard these two hypotheses as two solid, complementary views of how music/musicality evolved, but they could be improved with the consideration of fresh models of language(s) evolution, particularly, those based on the self-domestication account of human evolution.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Benítez-Burraco, A. (2017). Grammaticalization and language evolution: Focusing the debate. *Language Sciences*, *63*, 60–68. https://doi.org/10.1016/j.langsci.2017.03.003.

Benítez-Burraco, A., Chekalin, E., Bruskin, S., Tatarinova, T., & Morozova, I. (in press). Recent selection of candidate genes for domestication in Europeans and language change in Europe: A hypothesis. *Annals of Human Biology*.

Benítez-Burraco, A., & Kempe, V. (2018). The emergence of modern languages: Has human self-domestication optimized language transmission?. *Frontiers in Psychology*, *9*, 551. https://doi.org/10.3389/fpsyg.2018.00551.

Benítez-Burraco, A., & Progovac, L. (2020). A four-stage model for language evolution under the effects of human self-domestication. *Language & Communication*, *73*, 1–17. https://doi.org/10.1016/j.langcom.2020.03.002.

Bolender, J. (2007). Prehistoric cognition by description: A Russellian approach to the upper Paleolithic. *Biology & Philosophy*, *22*, 383–399.

Chekalin, E., Rubanovich, A., Tatarinova, T. V., Kasianov, A., Bender, N., Chekalina, M., … Morozova, I. (2019). Changes in biological pathways during 6,000 years of civilization in Europe. *Molecular Biology and Evolution*, *36*(1), 127–140. https://doi.org/10.1093/molbev/msy201.

Hare, B. (2017). Survival of the friendliest: *Homo sapiens* evolved via selection for prosociality. *Annual Review of Psychology*, *68*, 155–186. https://doi.org/10.1146/annurev-psych-010416-044201.

Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: Evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573–585. https://doi.org/10.1016/j.anbehav.2011.12.007.

Komiyama, T., Iwama, H., Osada, N., Nakamura, Y., Kobayashi, H., Tateno, Y., & Gojobori, T. (2014). Dopamine receptor genes and evolutionary differentiation in the domestication of fighting cocks and long-crowing chickens. *PLoS ONE*, *9*(7), e101778. https://doi.org/10.1371/journal.pone.0101778.

Langley, M. C., Benítez-Burraco, A., & Kempe, V. (2020). Playing with language, creating complexity: Has play contributed to the evolution of complex language? *Evolutionary Anthropology*, *29*(1), 29–40. https://doi.org/10.1002/evan.21810.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559. https://doi.org/10.1371/journal.pone.0008559.

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*(7613), 547–550. https://doi.org/10.1038/nature18635.

Nettle D. (2012). Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1597), 1829–1836. https://doi.org/10.1098/rstb.2011.0216.

Pisor, A. C., & Surbeck, M. (2019). The evolution of intergroup tolerance in nonhuman primates and humans. *Evolutionary Anthropology*, *28*(4), 210–223. https://doi.org/10.1002/evan.21793.

Progovac, L., & Benítez-Burraco, A. (2019). From physical aggression to verbal behavior: Language evolution and self-domestication feedback loop. *Frontiers in Psychology*, *10*, 2807. https://doi.org/10.3389/fpsyg.2019.02807.

Sato, D. X., Rafati, N., Ring, H., Younis, S., Feng, C., Blanco-Aguiar, J. A., … Andersson, L. (2020). Brain transcriptomics of wild and domestic rabbits suggests that changes in dopamine signalling and ciliary function contributed to evolution of tameness. *Genome Biology and Evolution*, *12*(10), 1918–1928. https://doi.org/10.1093/gbe/evaa158.

Trudgill, P. (2011). Sociolinguistic *typology: Social determinants of linguistic complexity*. Oxford university Press.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua. International Review of General Linguistics. Revue Internationale De Linguistique Generale*, *117*, 543–578.

Music and dance are two parallel routes for creating social cohesion

Steven Brown

Department of Psychology, Neuroscience & Behaviour, McMaster University,
Hamilton, Ontario L8S 4K1, Canada.
stebro@mcmaster.ca; neuroarts.org

doi:10.1017/S0140525X20000977, e65

Abstract

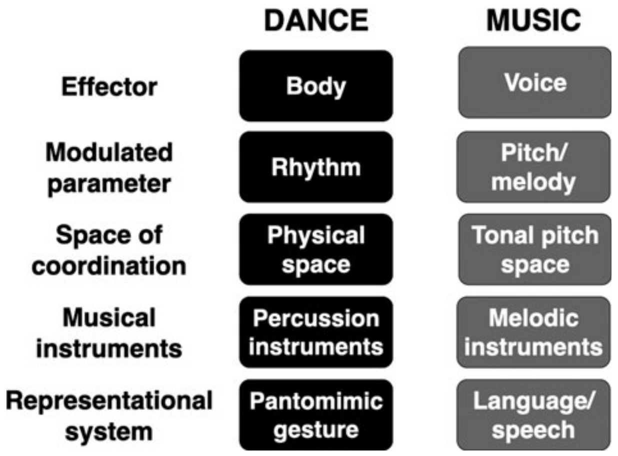
Savage et al. do an excellent job of making the case for social bonding in general, but do a less good job of distinguishing the manners by which dance and music achieve this. It is important to see dance and music as two parallel and interactive mechanisms that employ the "group body" and "group voice," respectively, in engendering social cohesion.

Darwin, in *The Descent of Man* (1871), was the first person to point out that music was utterly mysterious from the standpoint of natural selection as a mechanism to support individual survival. He suggested instead that music could potentially be an example of his alternative mechanism of sexual selection, whereby music would function as an esthetic display to attract mates (Miller, 2000; Prum, 2012), rather than promoting survival per se. However, an examination of the musical anthropology literature – totally separate from concerns about evolutionary mechanisms – casts strong doubt on this hypothesis. It shows that music in traditional cultures is not used for courtship in the manner that Darwin imagined, but that it is instead produced overwhelmingly by and for groups (Lomax, 1968; Merriam, 1964), principally in contexts of group-wide concern (Arom & Khalfa, 1998).

By the time that our book *The Origins of Music* was published in 2000 (Wallin, Merker, & Brown, 2000), the scorecard was pretty much at this point. Similar to Darwin, nobody was seriously advocating a pure natural-selection argument for music, and so the debate was polarized between sexual selection for mate attraction and theoretical notions about group functionality. I described this in detail in Brown (2000). Two decades later, the debate has not really budged from where it was in 2000. The only difference is that the group-functionalist account has been given new clothing in the name of social bonding and coalition building. But, it's the same old story. If music's evolutionary advantage is not about individuals competing with one another for survival, and not about males competing with one another for mates, then it's about enhancing group solidarity by increasing people's altruistic tendency to cooperate with group members – both non-kin and kin – by means of an engagement in collective, participative rituals involving music.

Another thing that hasn't budged over the years is the visceral resistance to seeing music's cooperative advantages as implicating some form of group or multi-level selection. However, no matter how you slice the issue, it ultimately comes down to the evolutionary problem of altruism and cooperation (Bell, Richerson, & McElreath, 2009; Boyd & Richerson, 1985; Dugatkin & Reeve, 1994; Smaldino, 2014; Sober & Wilson, 1998; Wilson, 2012; Wilson & Wilson, 2007). Group music-making is no less problematic in this regard than the phenomenon of an individual running into a burning building to save a stranger. Social bonding is merely a means to an end. That end is something along the lines of being willing to participate in a potentially deadly hunt of a large mammal. The music-making ritual comes *before* the hunt (not during it) in order to influence people's willingness to make self-sacrifices on behalf of the group and engage in a life-threatening endeavor.

The article by Savage et al. is an excellent exposition of the group-functionalist account of music, while even including the terrain of sexual selection for mate attraction and kin-selection accounts based on mother–infant interaction. There is something for everyone in this article, except for a hard-core natural selectionist arguing for individual-level survival advantages of music. The breadth of the cross-disciplinary coverage is impressive and highly commendable. The article is an outstanding addition to other group-functionalist accounts of music, whether from an evolutionary perspective or from the anthropological perspective that has dominated ethnomusicology since classic texts such as Merriam's (1964) *The Anthropology of Music* and Blacking's (1974) *How Musical is Man*. An important idea that is implicit in the social bonding hypothesis but that is not really stated is that music evolved as something choral from its inception and that its most significant design features are its textural properties (Jordania, 2006).



	DANCE	MUSIC
Effector	Body	Voice
Modulated parameter	Rhythm	Pitch/ melody
Space of coordination	Physical space	Tonal pitch space
Musical instruments	Percussion instruments	Melodic instruments
Representational system	Pantomimic gesture	Language/ speech

Figure 1. (Brown) Parallel properties of dance and music.

My main critique of the article is that music and dance are not adequately distinguished, and that a better version of the social bonding argument would view group dancing and group chorusing as *two parallel routes* for achieving bonding via different manners of interpersonal coordination. I can't think of a more insightful description of this duality than Radcliffe-Brown's (1922) classic analysis of the group rituals of the Andaman Islanders: "the primary function of [the] rhythmical nature of the dance is to enable a number of persons to join in the same actions and perform them *as one body*," and "song first came into general use in human society because it provides a means by which a number of persons can utter the same series of sounds together and *as with one voice*" (pp. 247–248, emphases added). Savage et al. do an excellent job of making the case for social bonding in general, but do a less good job of distinguishing the distinct manners by which dance and music achieve this, instead giving most of the attention to music. I think it is more reasonable to see dance and music as two parallel routes to social bonding that employ the group body and group voice, respectively (Fig. 1). Dance and music differ not only in the principal motor effectors used, but also in the types of parameters that are modulated, the space in which coordination occurs, associations with musical instruments, and associations with semantic representational systems. It is, thus, important to see social bonding as occurring by means of two parallel and interactive mechanisms, rather than calling this collective suite of corporeal and vocal behaviors "music," as the authors tend to. As I've argued previously, body percussion provides the ideal interface between dance and music as rhythmic phenomena (Brown & Parsons, 2008).

Two decades after evolutionary musicology took off, I feel that we are more or less at the same point as where we started. However, Savage et al. have made a very important contribution to the group-functionalist account of music. My suggestions for future study are to better define the contributions of dance and music to social bonding, as well as to overcome the resistance to invoking group selection mechanisms in explaining the most synchronous and prosocial group behaviors in the entire human repertoire, namely group dancing and chorusing.

Financial support. This study was funded by grants from the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

Conflict of interest. None.

References

Arom, S., & Khalfa, J. (1998). Une raison en acte: Pensée formelle et systématique musicale dans les sociétés de tradition orale. *Revue de Musicologie*, *84*, 5–17.

Bell, A. V., Richerson, P. J., & McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences*, *106*, 17671–17674. https://doi.org/10.1073/pnas.0903232106.

Blacking, J. (1974). *How musical is man?* University of Washington Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.

Brown, S. (2000). Evolutionary models of music: From sexual selection to group selection. In F. Tonneau & N. S. Thompson (Eds.), *Perspectives in ethology. 13: Behavior, evolution and culture* (pp. 231–281). Plenum Publishers.

Brown, S., & Parsons, L. M. (2008). The neuroscience of dance. *Scientific American*, *299*, 32–37.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. Princeton University Press.

Dugatkin, L. A., & Reeve, H. K. (1994). Behavioral ecology and the levels of selection: Dissolving the group selection controversy. *Advances in the Study of Behavior*, *23*, 101–133.

Jordania, J. (2006). *Who asked the first question? The origins of human choral singing, intelligence, language and speech*. Logos.

Lomax, A. (1968). *Folk song style and culture*. American Association for the Advancement of Science.

Merriam, A. P. (1964). *The anthropology of music*. Northwestern University Press.

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.

Prum, R. O. (2012). Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 2253–2265. https://doi.org/10.1098/rstb.2011.0285.

Radcliffe-Brown, A. (1922). *The Andaman islanders: A study in social anthropology*. Cambridge University Press.

Smaldino, P. E. (2014). The cultural evolution of emergent group-level traits. *Behavioral and Brain Sciences*, *37*, 243–295. https://doi.org/10.1017/S0140525X13001544.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Harvard University Press.

Wallin, N., Merker, B., & Brown, S. (Eds.). (2000). *The origins of music*. MIT Press.

Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology*, *82*, 327–348. https://doi.org/10.1086/703580.

Wilson, E. O. (2012). *The social conquest of the earth*. Liveright Publishing Corporation.

Music, attachment, and uncertainty: Music as communicative interaction

Ian Cross

Faculty of Music, Centre for Music & Science, University of Cambridge,
Cambridge CB3 9DP, UK.
ic108@cam.ac.uk; https://www.mus.cam.ac.uk/directory/ian-cross

doi:10.1017/S0140525X20001028, e66

Abstract

Both papers – to different degrees – underplay the interactive dimensions of music, and both would have benefited from integrating the concept of attachment into their treatments of social bonding. I further suggest that their treatment of music as a discrete domain of human experience and behaviour weakens their arguments concerning its functions in human evolution.

These two papers present contrasting views of music's relationship to evolutionary theory, drawing on a wide range of diverse sources. Savage et al. suggest that music is a mechanism for enhancing social bonds, its role in evolution being bound to its efficacy in establishing group cohesion. Mehr et al. reject this idea, claiming that music had dual adaptive functions in human evolution concerned with (i) coalition signalling, and (ii) signalling parental attention as a substitute for parental proximity. Both papers represent significant contributions to the literature on music and evolutionary theory, although both suffer from similar problems in the ways in which they operationalize conceptions of music that in one case weaken the force of the argument and in the other critically undermine the conclusions reached.

Both papers make frequent reference to the ethnomusicological literature and to findings with respect to cross-cultural musical universals. However, in outlining and defending their theses both tend to rely on ideas of music that reflect Western conceptions, substantially in Mehr et al., much less so in Savage et al. For Savage et al., this is most evident in their treatment of music in development, where they suggest that singing to infants allows "parents to communicate specific emotional messages to infants"; what's missing here is any sense of the interactive and reciprocal nature of musical engagement in infancy and, indeed, adulthood (see Malloch & Trevarthen, 2009). This is unfortunate, as in building a case for a close relationship between music and social bonding their paper otherwise engages fruitfully with the distinction between music as an interactive and as a presentational medium (after Turino, 2008) although their characterization of that distinction is too categorical and could have benefited from reference to studies by Slobin (1992) and Finnegan (1989).

In the case of Mehr et al., the conception of music that is at the heart of the paper – "Music is an auditory display built from melodies and rhythms" – is critically flawed by its curious neglect of the literature on music as a participatory and interactive medium. Although one of their two hypotheses concerns a signalling function of music in group behaviour, the members of the "group" that engages in musical signalling seem more like fireflies in their undifferentiatedly synchronous signalling than like the "Aka Congo Basin hunter-gatherers" to whom the authors refer, whose group musical behaviour is far away from synchrony in its spontaneously supple polyphony. The processes through which the members of a group interact with each other in creating and sustaining a "musical signal" are here treated almost as occurring in a black box, whereas the primary argument with respect to their musical signalling hypothesis really requires that the dynamics of those processes be understood and explained (see, e.g., Gratier, 2008; Keller, Novembre, & Hove, 2014; Turino, 2008).

Another curious omission from both papers is any consideration of the concept of attachment, mentioned in passing by Savage et al. and absent from Mehr et al. A substantial literature supports the idea that attachment constitutes a process grounded in mammalian neurophysiology that is powerfully and flexibly present in humans across the lifespan, extending to a wide variety of social relationships (see, e.g., Coan, 2008; Feldman, 2017). Reference to that literature could have strengthened the argument in Savage et al. although problematizing that presented in Mehr et al. A reorientation of the arguments towards attachment rather than bonding, a focus on infant–parent attachment, an acknowledgement that music plays a significant role in reinforcing this as an interactive medium, and an extension of the neural and behavioural correlates of parent–infant attachment behaviours to interactions with adults, kin, and strangers can yield what, according to Mehr et al., appears unviable; the instauration in the human behavioural repertoire of a mechanism – which we can call "something-like-music" – that can facilitate dyadic and group bonding.

Finally, both papers appear to conceive of music as a discrete domain of human experience, having parallels with but being distinct from language. Savage et al. suggest that because "music and language are both found universally in all known societies," they independently fulfil adaptive functions, whereas Mehr et al. assert that "language adequately provides whatever social functions grooming may have. As a social coordination or bonding mechanism, music thus appears to have no advantages over language and many disadvantages" (sect. 3.2.3, para 5). Ethnomusicological and ethnolinguistic evidence, together with a growing number of studies of real-time communicative interaction (e.g., Gorisch, Wells, & Brown, 2012; Robledo Del Canto, Hawkins, Cross, & Ogden, 2016) would suggest that a boundary between music and language is not clear-cut and that it may, in fact, not exist.

A fascinating recent instance is study by Senft (2018), who finds that the preferred mode of communicative interaction among the Trobriand islanders is *biga sopa*, "joking or lying speech … speech which is not vouched for," which is "characterised by four genres; *sopa* (joke, lie, trick), *kukwanebu sopa* (story), *kasilam* (gossip) and *wosi* (songs)." Here, speech in a register that is recognizable as phatic (after Malinowski, 1923) shades into music in the form of song; its function is strategic and social. As Senft notes (2018, pp. 211–212), "the *biga sopa* variety channels emotions, it keeps aggression under control, and it keeps possibilities of contact open … this concept with its tension-releasing functions secures harmony in Trobriand society."

The *biga sopa* speech register fulfils a function that I and others (Cross, 2006, 2014; McLeod, 1974) have proposed as integral to music across cultures and times: the management of situations of social uncertainty. Thinking of music as overlapping significantly – as an interactive medium – with speech registers oriented towards establishing and sustaining social relationships, the function of "something-like-music" in social bonding appears to fall out of its potential for managing situations of social uncertainty from dyads to groups, and from situations ranging across caregiver–infant interaction (Malloch & Trevarthen, 2009), the Aka's performance of *mokondi massana* to delight the forest (Lewis, 2009), and the sophisticatedly crude chanting of football crowds (Kytö, 2011).

And, of course, in whatever form something-like-music manifests itself, we do it together because it's fun!

Conflict of interest. None.

References

Coan, J. A. (2008). Toward a neuroscience of attachment. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (2nd ed., Vol. *2*, pp. 241–265). The Guilford Press.

Cross, I. (2006). Four issues in the study of music in evolution. *The World of Music*, *48*(3), 55–63.

Cross, I. (2014). Music and communication in music psychology. *Psychology of Music*, *42* (6), 809–819.

Feldman, R. (2017). The neurobiology of human attachments. *Trends in Cognitive Sciences*, *21*(2), 80–99.

Finnegan, R. (1989). *The hidden musicians: Music-making in an English town*. CUP.

Gorisch, J., Wells, B., & Brown, G. J. (2012). Pitch contour matching and interactional alignment across turns: An acoustic investigation. *Language and Speech*, *55*(1), 57–76.

Gratier, M. (2008). Grounding in musical interaction: Evidence from jazz performances. *Musicae Scientiae*, *12*(1 Suppl.), 71–110.

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369* (1658), 1–12.

Kytö, M. (2011). "We are the rebellious voice of the terraces, we are Çarşı": Constructing a football supporter group through sound. *Soccer & Society*, *12*(1), 77–93.

Lewis, J. (2009). As well as words: Congo Pygmy hunting, mimicry, and play. In R. Botha & C. Knight (Eds.), *The cradle of language* (Vol. *2*, pp. 381–413). Oxford University Press.

Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. Ogden & I. A. Richards (Eds.), *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism*. (pp. 296–336) Routledge.

Malloch, S., & Trevarthen, C. (2009). *Communicative musicality: Exploring the basis of human companionship*. Oxford University Press.

McLeod, N. (1974). Ethnomusicological research and anthropology. *Annual Review of Anthropology*, *3*, 99–115.

Robledo Del Canto, J. P., Hawkins, S., Cross, I., & Ogden, R. (2016). Pitch-interval analysis periodic and aperiodic Question+Answer pairs. Speech Prosody, Boston [10.21437/SpeechProsody.2016].

Senft, G. (2018). Theory meets practice – H. Paul Grice's maxims of quality and manner and the Trobriand Islanders' language use. In A. Capone, M. Carapezza & F. Lo Piparo (Eds.), *Further advances in pragmatics and philosophy: Part 1 from theory to practice* (pp. 203–220). Springer International Publishing.

Slobin, M. (1992). Micromusics of the West: A comparative approach. *Ethnomusicology*, *36*(1), 1–87.

Turino, T. (2008). *Music as social life: The politics of participation*. University of Chicago Press.

Music as a social bond in patients with amnesia

Maria Chiara Del Mastro[a], Maria Rosaria Strollo[a] and Mohamad El Haj[b,c,d]

[a]Department of Humanities, University of Naples "Federico II", Via Porta di Massa 1, 00133 Naples, Italy; [b]Nantes Université, Univ Angers, Laboratoire de Psychologie des Pays de la Loire (LPPL – EA 4638), F-44000 Nantes, France; [c]Unité de Gériatrie, Centre Hospitalier de Tourcoing, Tourcoing, France and [d]Institut Universitaire de France, Paris, France
mariachiara.delmastro@unina.it
mariarosaria.strollo@unina.it
mohamad.elhaj@univ-nantes.fr

doi:10.1017/S0140525X20000758, e67

Abstract

The music and social bonding (MSB) hypothesis proposes that human musicality has evolved as mechanisms supporting social bonding. We consider the MSB hypothesis under the lens of amnesia by arguing how patients with amnesia, especially those with Alzheimer's disease, can benefit from music, not only to retrieve personal memories, but also to use them for social bonding.

According to the music and social bonding (MSB) hypothesis, human musicality has evolved as a mechanism supporting social bonding. The MSB hypothesis assumes that musicality relies on multiple neurocognitive components, which likely evolved to facilitate social bonding, although some components may also be used for non-social purposes such as individual mood regulation. The social function of music, or social bonding, as proposed by the MSB hypothesis, can be better understood under the lens of amnesia. Here, we argue that patients with amnesia, especially those with Alzheimer's disease, can benefit from music to retrieve personal (i.e., autobiographical) memories and, critically, to use these memories and music in general as a tool for social bonding.

Autobiographical memory is negatively affected by Alzheimer's disease and patients typically suffer difficulties in retrieving personal memories, especially specific personal memories situated in time and space (El Haj, Antoine, Nandrino, & Kapogiannis, 2015; El Haj, Roche, Gallouj, & Gandolphe, 2017). However, research has demonstrated that autobiographical memory decline can be somewhat alleviated, at least at the mild stage of the disease, by exposure to music. Research has demonstrated that music significantly increases the quantity and quality of autobiographical retrieval in patients with Alzheimer's disease (El Haj, Clément, Fasotti, & Allain, 2013; El Haj, Postal, & Allain, 2012; Irish et al., 2006). The positive effects of music exposure on the autobiographical memory of patients with Alzheimer's disease can be attributed to the involuntary nature of music-evoked autobiographical memories, that is, they are retrieved with little cognitive control and are thus automatic and bypass strategic processes during retrieval (El Haj, Fasotti, & Allain, 2012; Strollo & Romano, 2016).

We believe that music-evoked autobiographical memories can help patients with Alzheimer's disease to share their personal experience with the social environment. Recollecting personal memories from one's life thanks to music exposure helps patients to think, talk, or teach about past experiences (El Haj & Antoine, 2016; Strollo & Romano, 2015). This assumption is supported by reminiscence theories according to which autobiographical retrieval has a social function, because sharing personal memories may (1) promote bonding between older adults, and (2) help in teaching others, especially the younger generation, about past experiences (Cappeliez, Rivard, & Guindon, 2007; Webster, 1993; Westerhof & Bohlmeijer, 2014). Thus, music-evoked autobiographical memories may be used by patients with Alzheimer's disease to connect to others and to transmit their personal experiences and life lessons to others, including to younger adults. Music exposure may also help patients with Alzheimer's disease to share self-defining memories, that is, those vivid, emotionally intense, and repeatedly retrieved memories that reflect enduring concerns in a patient's life. For example, music exposure was shown to cue self-defining memories in patients with Alzheimer's disease (El Haj, Antoine, Nandrino, Gely-Nargeot, & Raffard, 2015). Thus, music exposure can promote retrieval of memories that provide patients with Alzheimer's disease with a sense of purpose, continuity, and meaning in the situations they encounter, as well as with a better understanding of both the self and the world.

Autobiographical retrieval can, therefore, help patients make sense of the past and present, and then to share it with others to create a social bond. The sharing of narratives serves a dual purpose: on the one hand, it promotes self-expression by means of personal stories, and on the other, it increases social bonding. Ultimately, social bonding acts as a kind of "social glue," enabling patients with Alzheimer's disease to form social bonds with other patients and with younger people that they might otherwise not have created. These intra- and inter-generational bonds can be facilitated by the positive effects of music exposure on language, that is, it may enhance the ability of patients to produce complex linguistic structures, such as well-formed grammatical sentences, to describe their autobiographical memories.

To summarize, exposure to music can increase the quantity and quality of autobiographical retrieval in patients with Alzheimer's disease, thereby fueling their narrative with others to strengthen and maintain bonds. As suggested by the MSB hypothesis, music has a key social role to play, that is, to forge and reinforce affiliative inter-individual relationships. We believe that this is also the case for patients with Alzheimer's disease, who can use music-evoked autobiographical memories to share, tell, and/or teach others about these personal experiences. Music-evoked autobiographical memories can help patients with Alzheimer's disease to cement their social identity and reinforce group membership. We, thus, suggest that the MSB hypothesis could be extended to include music-evoked autobiographical memories in patients with Alzheimer's disease.

Financial support. Mohamad El Haj was supported by LABEX (excellence laboratory, program investment for the future) DISTALZ (Development of Innovative Strategies for a Transdisciplinary approach to Alzheimer's disease).

Conflict of interest. None.

References

Cappeliez, P., Rivard, V., & Guindon, S. (2007). Functions of reminiscence in later life: Proposition of a model and applications. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *57*(3), 151–156. doi: http://dx.doi.org/10.1016/j.erap.2005.02.014.

El Haj, M. A., & Antoine, P. (2016). Death preparation and boredom reduction as functions of reminiscence in Alzheimer's disease. *Journal of Alzheimer's Disease: JAD*, *54* (2), 515–523. doi: 10.3233/jad-160497.

El Haj, M., Antoine, P., Nandrino, J. L., Gely-Nargeot, M. C., & Raffard, S. (2015). Self-defining memories during exposure to music in Alzheimer's disease. *International Psychogeriatrics*, *27*(10), 1719–1730. doi: 10.1017/S1041610215000812.

El Haj, M., Antoine, P., Nandrino, J. L., & Kapogiannis, D. (2015). Autobiographical memory decline in Alzheimer's disease, a theoretical and clinical overview. *Ageing Research Reviews*, *23*(Pt B), 183–192. doi: 10.1016/j.arr.2015.07.001.

El Haj, M., Clément, S., Fasotti, L., & Allain, P. (2013). Effects of music on autobiographical verbal narration in Alzheimer's disease. *Journal of Neurolinguistics*, *26*(6), 691–700. doi: http://dx.doi.org/10.1016/j.jneuroling.2013.06.001.

El Haj, M., Fasotti, L., & Allain, P. (2012). The involuntary nature of music-evoked autobiographical memories in Alzheimer's disease. *Consciousness and Cognition*, *21*(1), 238–246. doi: 10.1016/j.concog.2011.12.005.

El Haj, M., Postal, V., & Allain, P. (2012). Music enhances autobiographical memory in mild Alzheimer's disease. *Educational Gerontology*, *38*(1), 30–41.

El Haj, M., Roche, J., Gallouj, K., & Gandolphe, M. C. (2017). Autobiographical memory compromise in Alzheimer's disease: A cognitive and clinical overview. *Geriatrie Et Psychologie Neuropsychiatrie Du Vieillissement*, *15*(4), 443–451. doi: 10.1684/pnv.2017.0704.

Irish, M., Cunningham, C. J., Walsh, J. B., Coakley, D., Lawlor, B. A., Robertson, I. H., & Coen, R. F. (2006). Investigating the enhancing effect of music on autobiographical memory in mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *22*(1), 108–120. doi: 10.1159/000093487.

Strollo, M. R., & Romano, A. (2015). *Music, emotions, autobiographical memory: An interdisciplinary approach*. Liguori.

Strollo, M. R., & Romano, A. (2016). L'autobiografia musicale come dispositivo narrativo con i pazienti affetti da Alzheimer: una ricerca empirica. In F. Marone (Ed.), *Pedagogia della cura. Famiglie, comunità, legami sociali*. Pensa Multimedia.

Webster, J. D. (1993). Construction and validation of the Reminiscence Functions Scale. *Journal of Gerontology*, *48*(5), P256–P262.

Westerhof, G. J., & Bohlmeijer, E. T. (2014). Celebrating fifty years of research and applications in reminiscence and life review: State of the art and new directions. *Journal of Aging Studies*, *29*, 107–114. doi: 10.1016/j.jaging.2014.02.003.

Ancestral human mother–infant interaction was an adaptation that gave rise to music and dance

Ellen Dissanayake

School of Music, University of Washington, Seattle, WA 98195-3450, USA.
ed3@uw.edu

doi:10.1017/S0140525X20001144, e68

Abstract

Human infants are born ready to respond to affiliative signals of a caretaker's face, body, and voice. This ritualized behavior in ancestral mothers and infants was an adaptation that gave rise to music and dance as exaptations for promoting group ritual and other social bonding behaviors, arguing for an evolutionary relationship between mother and infant bonding and both music and dance.

Hypotheses about the origin of music frequently mention the importance of maternal singing, as in soothing lullabies, which is virtually universal. However, studies by developmental psychologists of playful face-to-face interactions between mothers and infants of 4 weeks to 6 months of age show that a mother's vocalizations are more complex than "singing." They are embedded in a multimodal package that includes her facial expressions and head and body movements, as well as singsong vocalizations.

Microanalyses of these early interactions reveal that the behavior and affect of both partners are temporally organized, thereby emotionally coordinating or "attuning" the pair (Beebe, 1982; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001, pp. 13–14; Stern, 1985; Trevarthen, 1999). Infants as young as 4–8 weeks expect social contingency. That is revealed by dual-video recordings of face-to-face interactions on television screens in separate rooms. After 2 min of normal ongoing play, the mother's film is desynchronized (replayed) so that the infant now sees her earlier behavior that had just been engaging and delightful. However, it quickly realizes that something is wrong and shows signs of psychological distress such as averted gaze, closed mouth, frowning, fingering of clothing, and the displacement activity of yawning (Murray & Trevarthen, 1985; Nadel, Carchon, Kervella, Marcelli, & Reserbet-Planty, 1999).

This emotional/behavioral coordination is more than "social." It is *relational*, and has developmental benefits and adaptive implications. Neurobiologists describe the pathological effects to infants of deficient early interactive exchanges (Aitken & Trevarthen, 1997; Koulomzin *et al*., 2002; Schore, 1994; Trevarthen & Aitken, 1994), corroborating psychologists' findings about the importance of positive intimate multimodal early interactions to beneficial outcomes in later (post-infant) life.

These early interaction findings support the music as social bonding (MSB) hypothesis by adding to it (a) the psychobiological (and presumably evolutionary) importance of maternal facial, head, and body *movements* (not only "singing") in the earliest mother–infant interactions; (b) the creation and reinforcement of *emotional bonding* (not only "infant mood regulation") by the temporal coordination achieved and reinforced by the attendant neurochemical mechanisms the authors describe; and (c) the infant's contribution to (*participation in*) the multimodal exchange, as its own behavior affects the mother's changing vocal, facial, and gestural dynamics of faster–slower, louder–softer, larger–smaller, wider–narrower and so forth, which she produces according to her perception of her infant's emotional/behavioral state.

The authors describe "neurobiological proximate mechanisms," as "underpinning musicality's social effects" and occurring "between infant and mother, mates, and among members of groups." Because oxytocin and other endogenous opioids are released during parturition, lactation, and maternal caring behavior in mammals, it is plausible that these emotionally rewarding prosocial mechanisms would have been used, and augmented, to address an adaptive problem in ancestral humans: ensuring maternal care of highly altricial (helpless) infants who need months and years of care. The coevolved creation of and response to ritualized facial, bodily, and vocal signals of affiliative intent by ancestral mother–infant pairs physically coordinated and emotionally bonded them. A successful bond would have helped to assure both infant survival and maternal reproductive success.

Looked at in this way, it is mother–infant mutuality that is the adaptation and musicality can be considered an *exaptation* (not a byproduct) that, as the "set of capabilities that can be utilized in different ways to support multiple functions, all involving social effects," appears universally in all societies in culturally varied ways. Musicality as a set of capacities that was prefigured in early interactions need not have been only, as the MSB authors suggest, a "cultural 'invention'" or "byproduct" of other adaptations but an evolved predisposition to use these capacities according to the different lifeways of varied cultural groups.

This scenario also challenges some of the assumptions and assertions of the authors of the credible signaling hypothesis (MCS) as applied to infant care – that is, that infant-directed song is primarily a coevolved system for credibly signaling parental attention to secondarily altricial infants. The authors of the hypothesis mention affective prosody of mothers to infants – calling it "manipulative" – but do not take account of the whole reciprocal package in early mother–infant interactions that emotionally bonds the pair by means of facial and gestural as well as vocal components (and also includes touch). These early interactions do not only "communicate information" but express and exchange signals of shared positive emotional intent and accord, apart from whatever overt or covert information the mother might also communicate about what is in her mind.

The MCS hypothesis rests upon *competition and territoriality*, and *parent–infant conflict*. The authors emphasize that in ancestral societies (and presumably, later, in other small-scale societies), mothers were required to simultaneously care for multiple dependent offspring and that infants had to compete with multiple juveniles for the attention of multiple caregivers. Accounts of infancy in small-scale societies (e.g., Eibl-Eibesfeldt, 1989, p. 215; Field, Sostek, Vietze, & Leidermann, 1981; Hrdy, 2009; Leiderman, Tuckerman, & Rosenfeld, 1977) suggest that crying and care-eliciting distress vocalizations by infants are rare, as babies are usually carried or otherwise attached to the mother's body or handed to another person – older siblings, cousins, grandparents.

The authors of MCS are rightly concerned that hypotheses about a behavior's adaptive value present evidence of design, distinguishing between proximate and ultimate benefits.

As I describe, the original adaptive problem was to insure maternal care of unprecedently altricial human infants, thereby assuring the ultimate benefits of infant survival and maternal reproductive success. The solution was to co-opt the dopaminergic reward system and produce oxytocin and other opioids (that already underlay mammalian maternal behavior) for use in proximately rewarding face-to-face interactions.

In addition to the mother's ritualized (enhanced, emphasized) facial, bodily, and vocal signals of accord, the interactions became more temporally coordinated and more emotionally rewarding, ultimately bonding the pair.

Let us assume that in small bands of hominins it would be adaptive for individuals in a fractious group to become periodically bonded – physically and emotionally coordinated – and thereby predisposed to cooperate in common cause. Mother–infant early interactions already provided rewarding neurophysiological mechanisms that accompanied ritualized and temporally coordinated vocal and movement modalities to the end of bonding. These could have been developed in group rituals, such as feasting, with its multimodal components of music, dance, costume, ornament, and so forth, as the authors of MCS describe (see Dissanayake, 2018, pp. 91–93, for detailed discussion of the development of ritual and ceremonial behavior in ancestral humans).

With this scenario – music(ality)'s origins in biobehavioral predispositions that originated in adaptive ancestral mother–infant interactions – other suggested evolutionary hypotheses of music's function can fit under one umbrella. One can emphasize either a cooperative or competitive outcome – social bonding, sexual or social or territorial display, credible signaling, or other possibilities – any one or all of which may exist within one group. Considering music in this way as an exaptation supports the emphasis in both articles on cultural evolution as explaining music's many different appearances around the world. All retain the original features of early interactions, as described above.

References

Aitken K., & Trevarthen, C. (1997). Self-other organization in human psychological development. *Development and Psychopathology 9*, 653–677.

Beebe, B. (1982). Micro-timing in mother–infant communication. In M. R. Key (Ed.), *Nonverbal communication today* (pp. 168–195). Mouton.

Dissanayake, E. (2018). From play and ritualization to ritual and its arts: Sources of Upper Pleistocene ritual practices in Lower Middle Pleistocene ritualized and play behaviors in ancestral hominins. In C. Renfrew, I. Morley & M. Boyd (Eds.), *Ritual, play and belief, in evolution and early human societies* (pp. 87–98). Cambridge University Press.

Eibl-Eibesfeldt, I. (1989). *Human ethology*. Aldine de Gruyter.

Field, T. M., Sostek, A. M., Vietze, & Leidermann, P. H. (Eds.) (1981). *Culture and early interactions*. Erlbaum.

Hrdy, S. B. (2009). *Mothers and others: The evolutionary origin of mutual understanding*. Harvard.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). *Rhythms of dialogue in infancy*. Blackwell.

Koulomzin, M., Beebe, B., Anderson, S., Jaffe, J., Feldstein, A., & Crown, C. (2002). Infant gaze, head, face and self-touch at 4 months differentiate secure vs. avoidant attachment at 1 year: A microanalytic approach. *Attachment and Human Development*, *4*(1), 3–24.

Leiderman, P. H., Tuckerman, S. R., & Rosenfeld, A. (Eds.) (1977) *Culture and infancy*. Academic Press.

Murray, L., & Trevarthen, C. (1985) Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177–197). Ablex.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Reserbet-Planty, D. (1999) Expectancies for social contingency in 2-month-olds. *Developmental Science*, *2*, 164–173.

Schore, A. N. (1994) *Affect regulation and the origin of the self: The neurobiology of emotional development*. Erlbaum.

Stern, D. (1985) *The interpersonal world of the infant*. Basic Books.

Trevarthen, C. (1999) Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicæ Scientiæ* [Special Issue 1999–2000], 155–215.

Trevarthen, C., & Aitken, K. (1994) Brain development, infant communication, and empathy disorders: Intrinsic factors in child mental health. *Developmental Psychopathology*, *6*, 597–633.

The music and social bonding hypothesis does require multilevel selection

Dustin Eirdosh[a] and Susan Hanisch[a,b]

[a]Department of Comparative Cultural Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany and [b]Faculty of Education, Department of Primary School Science Education, University of Leipzig, Leipzig, Germany.
dustin.eirdosh@eva.mpg.de
https://www.eva.mpg.de/comparative-cultural-psychology/staff/dustin-eirdosh/
susan.hanisch@eva.mpg.de
https://www.uni-leipzig.de/en/profile/mitarbeiter/dr-susan-hanisch/

doi:10.1017/S0140525X20000679, e69

Abstract

Is musicality an individual level adaptation? The authors of this target article reject the need for group selection within their model, yet their arguments do not fulfill the conceptual requirements for justifying such a rejection. Further analysis can highlight the explanatory value of embracing multilevel selection theory as a foundational element of the music and social bonding (MSB) hypothesis.

In developing the *music and social bonding* (MSB) hypothesis, Savage and colleagues have created a highly valuable synthesis spanning disciplines and levels of analysis that surely represents a much needed clarification on a line of thinking well over a century old. We agree with them in their approach and argumentation in nearly all respects; however, we suggest that their contextualization of the MSB hypothesis within the multilevel selection debates could benefit from further conceptual exposition. We argue that, in fact, multilevel selection is required to make MSB coherent, and embedding the hypothesis within this theoretical framework will extend the explanatory power of the model as a whole.

In section 6.2 on group selection, the authors briefly summarize a slice of the long-standing debates over group selection and multilevel selection, both in biological and cultural evolution, and suggest that group selection is not needed because "the key fitness advantages accrue to individuals."

The invocation of individual level fitness benefits is a common strategy to argue against the need for multilevel selection theory, yet it is a claim that can obscure the explanatory potential of multilevel selection if not carefully unpacked.

First, as Eldakar and Wilson (2011) make clear, it is key to specify if the fitness advantages accrued for individuals are being framed in *relative* or *absolute* terms. This is not made clear by the authors with regard to their framing of MSB. Given that the focus of the hypothesis is on group level social bonding functions, one would be hard pressed to argue that such functions of musicality increase the *relative* fitness of individuals compared to their (presumably equally socially bonding) group members. As Eldakar and Wilson describe, the invocation of *absolute* fitness advantages for individuals is no argument against the value of multilevel selection.

Second, to reject group selection in favor of individual selection, it would have to be shown, or at least argued, that individual level selection (i.e., relative fitness differentials) within groups has

Table 1. (Eirdosh & Hanisch) Payoff matrix for musicality and social bonding

	(A) Socially bonded through music	(B) Not socially bonded through music
(A) Socially bonded through music	1, 1	−1, 0
(B) Not socially bonded through music	0, −1	0, 0

been stronger than between group selection pressures. Given that the foundational claim of the MSB hypothesis is the social bonding function of musicality, this is a theoretical impossibility. Functions such as improved group coordination, by definition, cannot be achieved by selection acting on competing individuals.

We offer a simple payoff matrix (Table 1) to illustrate these points. We imagine groups of interacting individuals in which two phenotypes A and B exist, where A represents individuals with the kinds of neurobiological and psychological mechanisms that Savage et al. describe and that lead to increased social bonding, prosociality, helping behavior, and so forth in response to participatory music making.

Individuals with phenotype A who benefit from the social-bonding functions of musicality will *only* benefit within the context of structured groups of *likewise* socially-bonded musical individuals. Furthermore, it is thinkable that individuals with phenotype A who are in groups of predominantly phenotypes B might, in fact, have a relative fitness disadvantage because the social bonding mechanism is costly or might make the individual vulnerable to exploitation by non-reciprocating others in the group. Musicality and the social bonding effects could, therefore, not evolve through selection acting on competing individuals. However, it could evolve through multilevel selection if different groups with different compositions of phenotypes A and B exist in a population. In groups of more socially-bonded musical individuals, individuals will have a fitness advantage *relative to individuals in other groups* with less socially-bonded musical individuals. This payoff dynamic makes the situation one that requisitely invokes multilevel selection on a group-structured population in evolutionary game theory (e.g., Bowles & Gintis, 2011).

We can take this analysis a step further and use the analytic framework for cultural multilevel selection developed by Kline, Waring, and Salerno (2018) to determine if a trait can be said to be a group level cultural adaptation (i.e., emergent from group-level cultural selection pressures). This framework requires addressing five core criteria, and we argue that musicality, as framed within the MSB hypothesis, very likely does meet all of these requirements:

(1) *Indicators of group adaptation*: MSB suggests group-level adaptive functions of musicality including enhanced within group bonds, improved group coordination, and group membership cues.
(2) *Group structure*: participatory music making is likely a group structured trait, and the social bonding functions are definitionally group structured outcomes.
(3) *Selection mechanism*: the social bonding functions of musicality would plausibly facilitate survival and expansion of groups, as well as cultural transmission and trait-based migration between groups.
(4) *Group-level cultural selection*: because of the plausible probability of cultural selection mechanisms occurring at the group level, group-level cultural selection can be said to be occurring.
(5) *Group-level cultural adaptation*: given the group structured fitness benefits and selection processes resulting from the social-bonding outcomes of participatory music making, it can be seen as a group-level cultural adaptation.

Even in recent years, many leading scholars in the evolutionary social sciences have framed their hypotheses as not requiring group selection, perhaps because of perceptions of persistent controversy, only to be shown the conceptual issues that emerge from this lack of embrace (see discussions of Michael Tomasello's study in Wilson, 2018; and Richard Wrangham's study in Wilson, 2019). Given that the core criticism of group selection offered in the target article is a 2012 popular press piece from Steven Pinker that has since been well refuted (see Richerson, 2012), with no known formal response from Pinker himself, and given the theoretical value offered by multilevel selection theory in providing a framework for understanding the complex of socially structured fitness benefits likely emergent from evolving musicality, we suggest the authors and future scholars would benefit from re-thinking this aversion to group selection. Situating the MSB hypothesis within cultural multilevel selection theory can allow the hypothesis to achieve the full explanatory power the authors rightly ascribe its potential to be.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Bowles, S., & Gintis, H. (2011). *A cooperative species. Human reciprocity and its evolution*. Princeton University Press.
Eldakar, O. T., & Wilson, D. S. (2011). Eight criticisms not to make about group selection. *Evolution: International Journal of Organic Evolution*, *65*(6), 1523–1526.
Kline, M. A., Waring, T. M., & Salerno, J. (2018). Designing cultural multilevel selection research for sustainability science. *Sustainability Science*, *13*(1), 9–19.
Richerson, P. (2012). Comment on Steven Pinker's Edge Essay. This View of Life. The Evolution Institute. https://evolution-institute.org/commentary/peter-j-richerson-comment-on-steven-pinkers-edge-essay/.
Wilson, D. S. (2018). Multilevel selection and Tomasello's a natural history of human morality: A translation manual. *Philosophical Psychology*, *31*(5), 669–679.
Wilson, D. S. (2019) Group Selection in Every Way Except Using the Words: A Critique of "The Goodness Paradox" by Richard Wrangham. This View of Life. The Evolution Institute. https://evolution-institute.org/group-selection-in-every-way-except-using-the-words-a-critique-of-the-goodness-paradox-by-richard-wrangham/.

Is the MSB hypothesis (music as a coevolved system for social bonding) testable in the Popperian sense?

Jonathan B. Fritz[a,b]

[a]Center for Neural Systems, New York University, New York, NY 10003, USA and [b]Division of Behavioral and Cognitive Science, National Science Foundation, Alexandria, VA 22314, USA.
Jonathan.b.fritz@gmail.com

doi:10.1017/S0140525X20001739, e70

Abstract

"Music As a Coevolved System for Social Bonding" (MSB) is a brilliant synthesis and appealing hypothesis offering insights into the evolution and social bonding of musicality, but is so broad and sweeping it will be challenging to test, prove or falsify in the Popperian sense (Popper, 1959). After general comments, I focus my critique on underlying neurobiological mechanisms, and offer some suggestions for experimental tests of MSB.

"Music As a Coevolved System for Social Bonding" (MSB) is a unifying hypothesis making a broad set of interwoven claims and predictions. However, too many predictions are vaguely defined, such as: "social bonding design features operate at multiple levels simultaneously, in the same way that a couple dancing can intensify their own relationship, and their relationship with the broader social group." How do these different types of dance interactions influence social bonding? Specifically, would MSB predict greater social bonding after solo dancing at a rock concert versus couples ballroom dancing versus partner rotation in swing dancing or Greek Kalamatianos folkdance in a circle holding hands? Or predict higher levels of social bonding for experienced, habitual dancers versus occasional dancers? What's the time course of these effects? More importantly, how would MSB measure this and disentangle social bonding arising from music, or dancing, or simply being physically active in a congenial social setting (e.g., doing group calisthenics) with like-minded people? The authors conjecture "singing in large choirs should produce greater bonding than singing in small choirs" and mention studies suggesting oxytocin levels are elevated after singing, but admit "evidence linking oxytocin with music remains limited." What are the best quantitative psychological measures of social bonding and what are its best measurable and accurate neurochemical and neurophysiological markers? Oxytocin, dopamine or EOS levels? Cortisol levels to measure reduced stress? Or the neuronal coherence of dancers or choir singers? The authors sidestep detailed predictions and experimental measures, crucial to test and validate MSB. An example of one possible approach measures prosociality following interpersonal synchrony (Cirelli et al., 2014) or monitors performers and audience for multiple behavioral and neural measures during musical performance (Chang et al., 2017; Swarbrick et al., 2019). Critical tests of MSB require measurements and analysis of psychological measures of social bonding and associated neuromarkers before, during, following live performance or participatory musical or dance events and in longitudinal studies.

Despite "ubiquitous use of music in communal ceremonies and rituals" and the assertion that "dance is a core part of music-making and not a separate domain," there is simply no obligatory nor causal link between social bonding, dance and musicality at an individual level. Anecdotally, Beethoven was asocial and "never learned to dance in time with music" (Cooper, 2010). His greatest pleasure was solitary walks in the countryside where he was inspired to compose. Dancers in Merce Cunningham's company didn't dance to a beat, rather Merce created dances based on chance juxtaposition of movement and music. Would MSB predict less social bonding at Cunningham concerts or among dancers in his company? One of the most popular radio programs, Desert Island Disks, asked what music you would bring if you were a solitary castaway on a desert isle. Such solitary enjoyment of music doesn't jibe with the claim that musicality goes hand-in-hand with social interactions. In fact, much music making is solitary (professional musicians may spend 50–100 times more time alone practicing than performing in a group). However, MSB asserts solitary music making or listening is irrelevant from an evolutionary perspective, emphasizing participatory music. Although MSB's suggested testbeds (drumming circles, campfire singalongs, and folk dances) are certainly valid and important, this restricted realm of explanatory and predictive power limits the generality and applicability of MSB. One avenue for future neuroimaging studies might be to explore "neural homophily" (a functional similarity in brain activity in friends – Parkinson et al., 2018) in relation to shared music preferences.

A key aspect of MSB is a neurobiological model for interactions between music perception, production, and social bonding, which gives a central role for dopamine in music prediction and reward (Ferreri et al., 2019; but see Goupil & Aucouturier, 2019 and Castro et al., 2020 for other music-evoked emotions). However, although a good beginning, the MSB neural model (Fig. 3 of the Savage et al. target article) is overly simplified – and impoverished because it does not incorporate brain structures known to play important roles in music perception and production: (a) cerebellum contributes to timing, rhythm perception, and generation (Teki & Griffiths, 2016; Teki et al., 2011), (b) other frontal areas such as frontopolar cortex (Medalla & Barbas, 2014), and dorsolateral frontal cortex (Mas-Herrero et al., 2018), (c) higher auditory association cortex, in anterior STG (Hackett, 2015; Moerel et al., 2014), and (d) insula. Their model doesn't mention evolution of hemispheric lateralization for musical processing (Albouy et al., 2020; Zatorre, 2001). Although simplification can be useful, it limits precise and well-formulated predictions.

Nor does MSB draw a compelling link between neural music and reward pathways with social cognition networks, focusing on the arcuate fasciculus and research indicating emotionally empathic people have higher arcuate microstructural integrity compared to autistic individuals with impaired social bonding, and less arcuate connectivity. However, music is a strength of autistic individuals, who often have excellent musical memory, pitch perception, and music-evoked emotions (Wenhart et al., 2019). If people with autism have good musical perceptual abilities, but poor integrity of the arcuate and poor social bonding, this is an argument *against* MSB's assumption of a strong link between musicality and social bonding.

MSB raises fascinating evolutionary questions, and opportunities to use animal models to help understand the evolution of "proto-musical" abilities and mechanisms for social bonding – and underlying neural networks and connections. Duetting, in which pair partners combine their songs into temporally coordinated joint displays, helps strengthen pair bonds (Baldassarre et al., 2016). Hoffmann et al. (2019) demonstrated synchronization of premotor brain activity in duetting pairs of sparrow weavers in the wild, consistent with MSB view that rhythmic cooperative behavior leads to "neural resonance" (synchronous brain activity across individuals) that facilitates social bonding. Is this resonance stronger in longer established pairs?

Does duetting in gibbons and turn-taking (Pika et al., 2018) in other NHPs, and coordinated pant-hooting in chimpanzees promote group bonding? Chimpanzees engage in rhythmic group "rain dances" in the wild. Hattori and Tomonaga (2020) showed chimp rhythmic swaying could be sound-induced, providing a model for "proto-dance" studies to measure whether rhythmic group movement enhances social bonding in chimpanzees. Although monkeys don't entrain to rhythm, they're sensitive to harmonic structure. However, macaques lack relative pitch ability to recognize non-octave transpositions of melodies (Wright et al., 2000), perhaps related to the lack of differentiation of a "cortical music pathway" (Norman-Haignere et al., 2019). It would be valuable to confirm whether this result generalizes to other monkeys, such as more vocal and "musical" duetting species (marmosets, geladas, and gibbons; Geissmann, 2000), or to great apes. Research investigating the neural connections between auditory, vocalization, reward, and social cognition networks (Freiwald, 2020; Shepherd et al., 2018; Sliwa & Freiwald, 2017) will illuminate neural mechanisms and the role of vocalizations, reward pathways, and social bonding in primates.

Conflict of interest. None.

References

Albouy, P., Benjamin, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science (New York, N.Y.)*, *367*, 1043–1047.

Baldassarre, D. T., Grieg, E. I., & Webster, M. S. (2016). The couple that sings together stays together: Duetting, aggression and extra-pair paternity in a promiscuous bird species. *Biology Letters*, *12*(2), 20151025.

Castro, M., L'heritier, F., Plailly, J., Saive, A. L., Corneyllie, A., Tillmann, B., & Perrin, F. (2020). Personal familiarity of music and its cerebral effect on subsequent speech processing. *Science Reports*, *10*(1), 14854.

Chang, A., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences USA*, *114*(21), E4134–E4141.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Development Science*, *17*(6), 1003–1011.

Cooper, B. (Ed.) (2010). *The Beethoven compendium*. Thames and Hudson.

Ferreri, L., Mas-Herrero, E., Zatorre, R. J., Ripolles, P., Gomez-Andres, A., Alicart, A., … & Rodriguez-Fornelis, A. (2019). Dopamine modulates the reward experiences elicited by music. *Proceedings of the National Academy of Sciences USA*, *116*(9), 3793–3798.

Freiwald, W. A. (2020). Social interaction networks in the primate brain. *Current Opinion in Neurobiology*, *65*, 49–58.

Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 103–123). MIT Press.

Goupil, L., & Aucouturier, J.-J. (2019). Musical pleasure and musical emotions. *Proceedings of the National Academy of Sciences*, *116*(9), 3364–3366.

Hackett, T. A. (2015). Anatomic organization of the auditory cortex. *Handbook Clinical Neurology*, *129*, 27–53.

Hattori, Y., & Tomonaga, M. (2020). Rhythmic swaying induced by sound in chimpanzees. *Proceedings of the National Academy of Sciences USA*, *117*(2), 936–942.

Hoffmann, S., Trost, L., Voigt, C., Leitner, S., Lemazina, A., Sagunsky, H., … & Gahr, M. (2019). Duets recorded in the wild reveal that interindividually coordinated motor control enables cooperative behavior. *Nature Communications*, *10*(1), 2577.

Mas-Herrero, E., Dagher, A., & Zatorre, R. J. (2018). Modulating musical reward sensitivity up and down with transcranial magnetic stimulation. *Nature Human Behavior*, *2*, 27–32.

Medalla, M., & Barbas, H. (2014). Specialized prefrontal "auditory fields": Organization of primate prefrontal-temporal pathways. *Frontiers in Neuroscience*, *8*, 77.

Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, *8*, 225.

Norman-Haignere, S. V., Kanwisher, N., McDermott, J. H., & Conway, B. R. (2019). Divergence in the functional organization of human and macaque auditory cortex revealed by fMRI responses to harmonic tones. *Nature Neuroscience*, *22*(7), 1057–1060.

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2018). Similar neural responses predict friendship. *Nature Communications*, *9*(1), 332.

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: Bridging the gap between human and animal communication. *Proceedings of the Royal Society B*, *285*(1880), 20180598.

Popper, K. (1959). *The logic of scientific discovery*. Basic Books.

Shepherd, S. V., & Freiwald, W. A. (2018). Functional networks for social communication in the macaque monkey. *Neuron*, *99*(2), 413–420.

Sliwa, J., & Freiwald, W. A. (2017). A dedicated network for social interaction processing in the primate brain. *Science (New York, N.Y.)*, *356*, 745–749.

Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H., & Trainor, L. J. (2019). How live music moves us: Head movement differences in audiences to live vs recorded music. *Frontiers in Psychology*, *9*, 2682.

Teki, S., & Griffiths, T. D. (2016). Brain bases of working memory for time intervals in rhythmic sequences. *Frontiers in Neuroscience*, *10*, 239.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *Journal of Neuroscience*, *31*(10), 3805–3812.

Wenhart, T., Bethlehem, R. A. I., Baron-Cohen, S., & Altenmuller, E. (2019). Autistic traits, resting-state connectivity, and absolute pitch in professional musicians: Shared and distinct neural features. *Molecular Autism*, *10*, 20.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology General*, *129*(3), 291–307.

Zatorre, R. J. (2001). Neural specializations for tonal processing. *Annals of the New York Academy of Sciences*, *930*, 193–210.

If music be the food of love, play on: Four ways that music may lead to social connection

Shira Gabriel[a] and Elaine Paravati[b]

[a]State University of New York, University at Buffalo, Buffalo, NY 14260, USA and [b]Department of Psychology, Hamilton College, Clinton, NY 13323, USA.
sgabriel@buffalo.edu
esparava@buffalo.edu

doi:10.1017/S0140525X20001430, e71

Abstract

The hypothesis that music is well suited to facilitate social bonding (Savage et al., target article) is highly consistent with social psychological research on the need to belong. We explore how music is uniquely placed to increase feelings of connections to large collectives by increasing collective effervescence, providing narratives, reminding one of others, and providing social surrogates.

The hypothesis that music is well suited to facilitate social bonding (Savage et al., target article) is highly consistent with social psychological research on the need to belong (for a review, see Leary & Gabriel, in press). For decades, social psychologists have argued that human beings evolved complex mechanisms that drive them toward social interaction and away from isolation (Baumeister & Leary, 1995). Detecting which activities increase social connections is tricky because the social motivational system is a primitive one and often operates outside of conscious awareness (Gabriel et al., 2016). Based on the compelling arguments laid out by Savage and colleagues and recent research on nontraditional means of maintaining social bonds (e.g., Paravati, Naidu, & Gabriel, 2020), we propose that music may play a key and integral role in helping people feel connected via collective effervescence, narratives, reminders of others, and social surrogates.

1. Collective effervescence

Social psychological research suggests that collective bonds form surprisingly easily. People automatically, quickly, and easily assimilate collective identities, even on the basis of the most minimal criteria (Tajfel, 1970). One of the key ways connections are felt to large groups is through the sense of collective effervescence (Gabriel, Naidu, Paravati, Morrison, & Gainey, 2020). Collective effervescence is the sensation of sacredness and feeling of connection to others that people sometimes experience when in a group activity such as a religious gathering, concert, or wedding. Collective effervescence mediates the effects of group participation on a plethora of positive outcomes (Páez, Rimé, Basabe, Wlodarczyk, & Zumeta, 2015). Experiencing collective effervescence is what leads groups to decrease loneliness, increase positive feelings, and increase spiritual transcendence (Gabriel, Read, Young, Bachrach, & Troisi, 2017; Gabriel et al., 2020).

We propose that music is likely to play a key role in facilitating collective effervescence. Our research suggests that the kinds of events that people describe as causing collective effervescence tend to be ones that involve music, dancing, and rhythmic chanting (Gabriel et al., 2017). Based on this research and the arguments of Savage and colleagues, we propose that music increases the collective effervescence experienced in group gatherings, which then leads to an increased feeling of connection and affiliation.

2. Narratives

Music can also increase feelings of belonging through the narratives that are presented in the lyrics of some songs. Mar and Oatley (2008) argue that one core function of narratives is to mentally simulate social interactions, facilitating subsequent social behavior. In addition, identifying with characters while reading a narrative leads to a merging of self with characters (Sestir & Green, 2010; Shedlosky-Shoemaker, Costabile, & Arkin, 2014), which has the potential to provide social benefits.

Narratives can also serve more immediate demands of the need to belong by filling belongingness needs and protecting against the harmful effects of rejection, social isolation, and loneliness (Derrick, Gabriel, & Hugenberg, 2009). Narratives lead individuals to psychologically become a part of the collective described within the narrative (Gabriel & Young, 2011). In other words, when people engage with a narrative, they feel like a member of the group described within the narrative. That feeling increases wellbeing via increased social connections (Gabriel & Young, 2011). Thus, providing narratives is another way that music can increase feelings of social connection.

3. Reminders of others

Music may also help fill belongingness needs by evoking shared memories (Kornhaber, 2020). Nonhuman reminders of actual human relationships, including photographs and letters (e.g., Gardner, Pickett, & Knowles, 2005), foods associated with loved ones (Troisi & Gabriel, 2011), and pictures of others on Facebook or other social media sites (Nadkarnia & Hofmann, 2012), can reduce feelings of loneliness and isolation (Sherman, 1991). Based on existing research and the arguments of Savage and colleagues, we argue that music can remind people of actual relationship partners – people with whom one shared musical experiences – and can thus increase feelings of social connection.

4. Social surrogates

Music may also fill the need to belong through parasocial relationships – one sided bonds with celebrities (Giles, 2002; Isotalus, 1995; Perse & Rubin, 1989; Rubin & McHugh, 1987). People may develop parasocial bonds with the musicians who produce their music. Although people consciously understand that parasocial relationships are not “real” relationships, they are nevertheless experienced as real and often elicit cognitive, affective, and behavioral responses that are similar to those of real relationships (Cohen, 2003; Kanazawa, 2002). Bonds with parasocial relationship partners contribute to a sense of social connection and wellbeing (Gabriel et al., 2016). Based on existing research and the arguments of Savage and colleagues, we argue that music can increase feelings of belonging even when used in isolation via listeners’ parasocial bonds with musicians.

5. Conclusion

Savage and colleague’s thesis that music evolved to communicate and strengthen social connections is highly consistent with social psychological research on belonging. Music is uniquely placed to increase feelings of connections to large collectives by increasing collective effervescence, providing narratives, reminding one of others, and providing social surrogates. As Mehr and colleagues point out, any evolutionary theory requires a certain amount of guess work and conjecture. However, their argument that “music does not *directly* cause social cohesion: rather, it signals existing social cohesion that was obtained by other means” (sect. 4.2.1, para. 14) is not consistent with existing research. For example, there is evidence that rhythmic synchronization (an important part of music) promotes increased prosocial behavior that comes from an increased sense of social connection (Mogan, Fischer, & Bulbulia, 2017; Rennung & Göritz, 2016), and that people who dance with others feel more connected to them, similar to them, and like them more (Tarr, Launay, Cohen, & Dunbar, 2015; Tarr, Launay, & Dunbar, 2016). In other words, key components of music seem to lead to social connection and not just signal that connection to others. Nonetheless, much more data are necessary to examine the causal link from music to social connection, and we hope that the ideas in this commentary suggest some possible avenues for exploration.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Cohen, J. (2003). Parasocial breakups: Measuring individual differences in responses to the dissolution of parasocial relationships. Mass *Communication & Society*, *6*, 191–202.

Derrick, J. L., Gabriel, S., & Hugenberg, K. (2009). Social surrogacy: How favored television programs provide the experience of belonging. *Journal of Experimental Social Psychology*, *45*(2), 352–362.

Gabriel, S., Naidu, E., Paravati, E., Morrison, C. D., & Gainey, K. (2020). Creating the sacred from the profane: Collective effervescence and everyday activities. *Journal of Positive Psychology*, *15*(1), 129–154.

Gabriel, S., Read, J. P., Young, A. F., Bachrach, R. L., & Troisi, J. D. (2017). Social surrogate use in those exposed to trauma: I get by with a little help from my (fictional) friends. *Journal of Social and Clinical Psychology*, *36*(1), 41–63.

Gabriel, S., Valenti, J., & Young, A. F. (2016). Watching, reading, and eating your way to belonging: Symbolic social relationships and the social self. *Advances in Experimental Social Psychology*, *53*, 189–243.

Gabriel, S., & Young, A. F. (2011). Becoming a vampire without being bitten: The narrative collective-assimilation hypothesis. *Psychological Science*, *22*(8), 990–994.

Gardner, W. L., Pickett, C. L., & Knowles, M. (2005). Social snacking and shielding: Using social symbols, selves, and surrogates in the service of belonging needs. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 227–241). Psychology Press.

Giles, D. C. (2002). Parasocial interaction: A review of the literature and a model for future research. *Media Psychology*, *4*, 279–305.

Isotalus, P. (1995). Friendship through screen: Review of parasocial relationship. *Nordicom Review*, *1*(1995), 59–64.

Kanazawa, S. (2002). Bowling with our imaginary friends. *Evolution and Human Behavior*, *23*, 167–171.

Kornhaber, S. (2020). The coronavirus has forced a repurposing of music. *The Atlantic*. 1238 https://www.theatlantic.com/culture/archive/2020/04/coronavirus-has-forced-repurposing1239music/609601/.

Leary, M. R., & Gabriel, S. (in press). The relentless pursuit of acceptance and belonging. *Advances in Motivational Science*.

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*, 173–192.

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony’s 1350 effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*, 13–20.

Nadkarnia, A., & Hofmann, S. G. (2012). Why do people use Facebook? *Personality and Individual Differences*, *52*, 243–249.
Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*, 711–729.
Paravati, E., Naidu, E., & Gabriel, S. (2020). From "love actually" to love, actually: The sociometer takes every kind of fuel. *Self and Identity*, *20*(3), 1–19.
Perse, E. M., & Rubin, R. B. (1989). Attribution in social and parasocial relationships. *Communication Research*, *16*, 59–77.
Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A meta-analysis. *Zeitschrift für Psychologie 224*(3): 168–189.
Rubin, R. B., & McHugh, M. P. (1987). Development of parasocial interaction relationships. *Journal of Broadcasting & Electronic Media*, *31*, 279–292.
Sestir, M., & Green, M. C. (2010). You are who you watch: Identification and transportation effects on temporary self-concept. *Social Influence*, *5*(4), 272–288.
Shedlosky-Shoemaker, R., Costabile, K. A., & Arkin, R. M. (2014). Self-expansion through fictional characters. *Self and Identity*, *13*(5), 556–578.
Sherman, E. (1991). Reminiscentia: Cherished objects as memorabilia in late-life reminiscence. *The International Journal of Aging & Human Development*, *33*, 89–100.
Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*, 96–102.
Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, *11*(10), 20150767.
Tarr, B., Launay, J., & Dunbar, R. I. (2016). Silent disco: Dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior*, *37*(5), 343–349.
Troisi, J. D., & Gabriel, S. (2011). Chicken soup really is good for the soul: "comfort food" fulfills the need to belong. *Psychological Science*, *22*(6), 747–753.

Music's putative adaptive function hinges on a combination of distinct mechanisms

Bruno Gingras

Department of Psychology, University of Innsbruck, 52f Innrain, Innsbruck
A-6020, Austria.
brunogingras@gmail.com

doi:10.1017/S0140525X20001752, e72

Abstract

Music's efficacy as a credible signal and/or as a tool for social bonding piggybacks on a diverse set of biological and cognitive processes, implying different proximate mechanisms. It is likely this multiplicity of mechanisms that explains why it is so difficult to account for music's putative biological role(s), as well as its possible origins, by proposing a single adaptive function.

In this volume, Savage et al. and Mehr et al. propose two theories of the evolution of musicality. Both theories stress the prosocial function of musicality, but whereas Savage et al. emphasize the role of music as a social bonding facilitator (the MSB hypothesis), Mehr et al. argue that music evolved as a credible signal in at least two distinct contexts, namely coalitional interactions and infant care.

Before considering these theories, let us backtrack slightly. Synchronization to an isochronous beat, which may be construed as a prerequisite for both Savage et al.'s (sect. 2.2) and Mehr et al.'s (sect. 5.1) hypotheses, can be viewed as a specific instantiation of more general self-organization phenomena widely seen in social animals from a wide range of taxa (O'Keeffe, Hong, & Strogatz, 2017). Along the same lines, mimicry, as well as emotional entrainment and/or movement synchronization, occur spontaneously in pairs or larger groups of humans (Néda, Ravasz, Brechet, Vicsek, & Barabási, 2000; Páez, Rimé, Basabe, Wlodarczyk, & Zumeta, 2015; Zivotofsky, Gruendlinger, & Hausdorff, 2012). It is worth noting that gait synchronization is facilitated much more by tactile and auditory feedback than by visual feedback (Zivotofsky et al., 2012), which may explain why auditory signals, and more specifically music, are preferentially used to facilitate synchronization, particularly in large groups where tactile feedback would be impractical (Savage et al., this volume).

Thus, many, if not most, prolonged interactions between conspecifics lead to spontaneous interpersonal synchrony, particularly with respect to mood and movement. Over time, this interpersonal synchrony could conceivably have coalesced into prototypical dance or musical forms, especially when reinforced by movement-generated acoustic feedback such as audible steps or clapping. At the same time, enhancing this naturally occurring interpersonal synchrony using external acoustic stimulation (which could be anything from a basic isochronous beat to more complex rhythmic and musical structures) would also yield behaviors akin to dance or music and would presumably facilitate bonding and/or signal greater group cohesion, as suggested by a large literature (see Rennung & Göritz, 2016, for a review). Both mechanisms outlined in this paragraph may in fact co-exist, making it difficult to establish a cause-and-effect relationship, or, indeed, an ontogenetic pathway for musicality in the context of social bonding.

The challenge of pinpointing an adaptive function for musicality is compounded by the fact that most musical genres and traditions combine distinct components that can be found in isolation or in non-musical contexts, and may have different evolutionary histories. Thus, rhythm constitutes an efficient tool for synchronization (and by extension social bonding) even in the absence of melody, for instance in drumming. Similarly, melody can convey emotions and affects in the absence of rhythm, for example in the case of the lament. The model proposed by Mehr et al. acknowledges this diversity of components by postulating separate selection pressures for rhythm and melody (sect. 5.1).

Along these lines, the perception and recognition of musical affect is based on several distinct mechanisms ranging from brainstem reflex to episodic memory (Juslin, 2013). Musical expectancy, or more generally the ability to predict upcoming musical events, is but one of these mechanisms (Juslin, Barradas, & Eerola, 2015). Basic acoustic parameters such as sound intensity, rate of change, or frequency spectrum play a major role in conveying emotions induced by music (Gingras, Marin, & Fitch, 2014) and indeed by a wide range of environmental sounds, including non-biological ones (Ma & Thompson, 2015). On the contrary, the predictive rewards associated with musical expectancy (Huron, 2006) only apply to certain auditory stimuli and furthermore involve culture-specific features such as tonality. From an evolutionary perspective, it may be more sensible to focus primarily on emotion-inducing mechanisms that have a broader purview, and are presumably phylogenetically more ancient, than to emphasize more specialized ones such as prediction.

Besides providing a credible explanation for the adaptive purposes of musicality, a plausible theory for the origins of musicality should account both for music's shared features across cultures and for the remarkable variability and complexity of musical styles. The signal elaboration and cultural ritualization aspects mentioned by Mehr et al. are, in my view, critical in this regard. Indeed, a simple rhythmic structure, augmented with just enough variation in repetitive melodic formulas to be distinguishable from the productions of other groups or cultures, would suffice to evoke most of the social bonding effects (including outgroup exclusion) predicted by the MSB, thus rendering any additional complexity and diversity superfluous, except as increasingly sophisticated indicators of group identity (Savage et al., sects. 2.2.3 and 2.2.4). On the contrary, the mechanisms suggested by Mehr et al. may help explain the diversity observed in many musical cultures.

In conclusion, music's efficacy as a credible signal and/or as a tool for social bonding appears to piggyback on a diverse set of biological and cognitive processes, implying different proximate mechanisms. Rhythmic synchronization, which stems from the basic tendency to mimic and imitate our associates, is tied to group cohesion and sociality. Melody, or more generally the systematic use of pitch variation, is related to prosody and is tied to emotion communication and mood regulation. Finally, basic acoustic attributes such as intensity, rate of change, and frequency range broadly convey information about power, speed, size, and distance, and lead to appropriate responses (e.g., fear). It is likely this multiplicity of mechanisms that explains why is it so difficult to account for music's putative biological role(s), as well as its possible origins, by proposing a single adaptive function.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Gingras, B., Marin, M. M., & Fitch, W. T. (2014). Beyond intensity: Spectral features effectively predict music-induced subjective arousal. *Quarterly Journal of Experimental Psychology*, *67*(7), 1428–1446.
Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.
Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266.
Juslin, P. N., Barradas, G., & Eerola, T. (2015). From sound to significance: Exploring the mechanisms underlying emotional reactions to music. *The American Journal of Psychology*, *128*(3), 281–304.
Ma, W., & Thompson, W. F. (2015). Human emotions track changes in the acoustic environment. *Proceedings of the National Academy of Sciences*, *112*(47), 14563–14568.
Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A. L. (2000). The sound of many hands clapping. *Nature*, *403*(6772), 849–850.
O'Keeffe, K. P., Hong, H., & Strogatz, S. H. (2017). Oscillators that sync and swarm. *Nature Communications*, *8*(1), 1–13.
Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*(5), 711–729.
Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A meta- analysis. *Zeitschrift für Psychologie*, *224*(3), 168–189.
Zivotofsky, A. Z., Gruendlinger, L., & Hausdorff, J. M. (2012). Modality-specific communication enabling gait synchronization during over-ground side-by-side walking. *Human Movement Science*, *31*(5), 1268–1285.

Is neural entrainment to rhythms the basis of social bonding through music?

Jessica A. Grahn[a], Anna-Katharina R. Bauer[b,c] and Anna Zamm[d]

[a]Department of Psychology, Brain and Mind Institute, Western University, London, ON N6A 3K7, Canada; [b]Department of Experimental Psychology, University of Oxford, Oxford OX2 6GG, UK; [c]Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Oxford OX3 7JX, UK and [d]Department of Cognitive Science, Central European University, Quellenstrasse 51, Vienna 1100, Austria.
jgrahn@uwo.ca; www.jessicagrahn.com
anna-katharina.matke-bauer@psy.ox.ac.uk; http://www.annakatharinabauer.com/
zamma@ceu.edu; https://annapzamm.github.io/

doi:10.1017/S0140525X20001296, e73

Abstract

Music uses the evolutionarily unique temporal sensitivity of the auditory system and its tight coupling to the motor system to create a common neurophysiological clock between individuals that facilitates action coordination. We propose that this shared common clock arises from entrainment to musical rhythms, the process by which partners' brains and bodies become temporally aligned to the same rhythmic pulse.

Many human biological processes – from breathing to walking – are rhythmic. Savage et al. propose entrainment as one neurobiological underpinning of music's social effects. Here, we clarify the concept of entrainment and its neurobiological mechanisms, arguing that entrainment should take the central role in the discussion of music's social origins.

Entrainment is the process by which two – or more – oscillatory (fluctuating) processes become coupled via phase or frequency adjustment (Pikovsky, Rosenblum, & Kurths, 2003). When two oscillators are entrained, they are temporally aligned, such as when musicians synchronize to another's tone onsets (Demos, Layeghi, Wanderley, & Palmer, 2019). This synchronization is sustained beyond direct physical coupling, which differentiates entrainment from *resonance*, a term that is often (and incorrectly) used interchangeably with entrainment (Helfrich, Breska, & Knight, 2019). In fact, the intention and attention required to maintain behavioral entrainment (Leow, Waclawik, & Grahn, 2018) may create the "social" nature of the signal: entrainment in musical group contexts is not simply reflexive or the result of unintentional mirroring.

When an individual entrains their actions with an external oscillation – such as the music of a partner – they achieve *behavioral* synchrony. Growing evidence suggests that behavioral entrainment is governed by *neural entrainment*, the alignment of rhythmic brain activity to external rhythmic stimuli (for a review, see Lakatos, Gross, & Thut, 2019). For music, neural entrainment may alter auditory perception by guiding attention toward rhythmically salient acoustic events (Henry & Herrmann, 2014; Jones, Moynihan, MacKenzie, & Puente, 2002). Moreover, neural entrainment to perceived rhythms facilitates detection of auditory stimulus features (Bauer, Bleichner, Jaeger, Thorne, & Debener, 2018; Henry & Obleser, 2012), suggesting that entrainment facilitates more than simply moving along to music.

Merely hearing an auditory rhythm activates motor brain areas (Grahn & Brett, 2007; Grahn & Rowe, 2009; Zatorre, Chen, & Penhune, 2007), particularly when the rhythm features a salient *beat* (Grahn & Rowe, 2013); a regular, psychologically salient, recurring event in the rhythm. The beat is central to entrainment, as the beat is what people generally synchronize to, and it enables synchronization even to novel music. Perhaps unsurprisingly then, entrainment of neural oscillations to the beat occurs not only in auditory, but also in motor regions (Fujioka, Trainor, Large, & Ross, 2009, 2015; Morillon & Baillet, 2017). Auditory-motor entrainment may be the neural driver behind moving to music; entrainment of beat-related auditory-motor activity influences behavioral synchronization (Mathias, Zamm, Gianferrara, Ross, & Palmer, 2020; Nozaradan, Zerouali, Peretz, & Mouraux, 2015). We propose that this tight auditory-motor coupling, translating musical rhythms into action, is critical for group music-making.

Musical partners mutually synchronize auditory and motor brain activity to the sounds that they produce (Müller, Sänger, & Lindenberger, 2013; Sänger et al., 2012, 2013; Zamm et al., 2018). In fact, when two brains are synchronized through electrical stimulation, spontaneous (Pan et al., 2021) and intentional (Novembre, Knoblich, Dunne, & Keller, 2017) bodily movements also become more synchronized. Moreover, physiological activity is coupled during synchronous musical behavior (Gordon et al., 2020; Mueller & Lindenberger, 2011). Thus, *interpersonal synchrony* operates at behavioral, neural, and physiological levels, all of which may support the social effects of synchronous music-making.

Interpersonal synchrony has been suggested to play an evolutionary role in promoting social cohesion (Launay, Tarr, & Dunbar, 2016) by facilitating trust (Launay, Dean, & Bailes, 2013), affiliation (Hove & Risen, 2009), and prosocial behavior (Cirelli, Einarson, & Trainor, 2014). Consistent with this view, interpersonal behavioral, physiological, and neural synchrony occur between mothers and infants (for a review, see Wass, Whitehorn, Haresign, Phillips, & Leong, 2020), between certain animal species' movements and computer-generated or conspecific rhythms (Cook, Rouse, Wilson, & Reichmuth, 2013; Lameira, Eerola, & Ravignani, 2019; Patel et al., 2009). These links between entrainment and social cohesion across adults, children, and certain animals, suggest that entrainment may be the musical feature evolutionarily selected to facilitate bonding.

Interestingly, both infants and animals are poorer at synchronization than adult humans, raising the question of whether precise synchronization and well-honed prediction mechanisms are critical for experiencing the social benefits of interpersonal synchrony, the latter being a central argument of Savage et al. Even adults show a wide range of synchronization abilities (Grahn & Schuit, 2012), yet there is little evidence that poor synchronizers – adults or children – do not enjoy dancing or moving to music, nor that they experience less social affiliation afterward. In fact, anecdotally, the social affiliation within an amateur, less synchronized musical group may surpass that of a professional, highly synchronized group. Therefore, bonding can arise in social entrainment contexts regardless of individuals' ability to synchronize accurately. Thus, the ability to experience the social benefits of entrainment in group music-making is not necessarily dependent on accurate entrainment ability.

Overall, auditory entrainment to musical rhythms activates the motor system. During music-making, synchronization of actions, brain rhythms, and physiological activity occurs between musical partners. However, many non-musical behaviors involving *entrained* actions, such as walking side-by-side (Nessler & Gilliland, 2010), rowing (Cohen, Ejsmond-Frey, Knight, & Dunbar, 2010), and dance (Chauvigné, Walton, Richardson, & Brown, 2019), facilitate interpersonal synchrony of behavior and likely also neural and physiological rhythms. We argue that entrainment – of bodies and minds – is the key evolutionary mechanism underlying the social bonding that arises in music. The highly rhythmic structure of music – which arises from both well-defined temporal and spectral patterns – makes music a superior facilitator of interpersonal synchrony and bonding to other entrained activities; however, this hypothesis remains to be definitively proven. We hope that future research inspired by this special issue will test how group music-making specifically enhances social bonding beyond other forms of interpersonal entrainment.

Acknowledgments. A.-K.R.B. receives funding from the German Research Foundation (DFG: MA 8554/1-1). A.Z. receives funding from European Commission H2020 Marie Skodowska-Curie Fellowship JAL-843722.

Conflict of interest. None.

References

Bauer, A.-K. R., Bleichner, M. G., Jaeger, M., Thorne, J. D., & Debener, S. (2018). Dynamic phase alignment of ongoing auditory cortex oscillations. *NeuroImage*, *167*, 396–407.

Chauvigné, L. A., Walton, A., Richardson, M. J., & Brown, S. (2019). Multi-person and multisensory synchronization during group dancing. *Human Movement Science*, *63*, 199–208.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011.

Cohen, E. E. A., Ejsmond-Frey, R., Knight, N., & Dunbar, R. I. M. (2010). Rowers' high: Behavioural synchrony is correlated with elevated pain thresholds. *Biological Letters*, *6*, 106–108.

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non-vocal mimic. *Journal of Comparative Psychology*, *127*(4), 412.

Demos, A. P., Layeghi, H., Wanderley, M. M., & Palmer, C. (2019). Staying together: A bidirectional delay–coupled approach to joint action. *Cognitive Science*, *43*(8), e12766.

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.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, *1169*, 89–92.

Gordon, I., Gilboa, A., Cohen, S., Milstein, N., Haimovich, N., Pinhasi, S., & Siegman, S. (2020). Physiological and behavioral synchrony predict group cohesion and performance. *Scientific Reports*, *10*(1), 1–12.

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.

Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, *23*(4), 913–921.

Grahn, J. A., & Schuit, D. (2012). Individual differences in rhythmic ability: Behavioral and neuroimaging investigations. *Psychomusicology: Music, Mind, and Brain*, *22*(2), 105.

Grahn, J., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Helfrich, R. F., Breska, A., & Knight, R. T. (2019). Neural entrainment and network resonance in support of top-down guided attention. *Current Opinion in Psychology*, *29*, 82–89.

Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception*, *2*, 62–86.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(49), 20095–20100.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social cognition*, *27*(6), 949–960.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319.

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*, *29*(18), R890–R905.

Lameira, A. R., Eerola, T., & Ravignani, A. (2019). Coupled whole-body rhythmic entrainment between two chimpanzees. *Scientific Reports*, *9*(1), 1–8.

Launay, J., Dean, R. T., & Bailes, F. (2013). Synchronization can influence trust following virtual interaction. *Experimental Psychology*, *60*, 53–63.

Launay, J., Tarr, B., & Dunbar, R. I. (2016). Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology*, *122*(10), 779–789.

Leow, L. A., Waclawik, K., & Grahn, J. A. (2018). The role of attention and intention in synchronization to music: Effects on gait. *Experimental Brain Research*, *236*(1), 99–115.

Mathias, B., Zamm, A., Gianferrara, P. G., Ross, B., & Palmer, C. (2020). Rhythm complexity modulates behavioral and neural dynamics during auditory–motor synchronization. *Journal of Cognitive Neuroscience*, *32*(10), 1864–1880.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, *114*(42), E8913–E8921.

Müller, V., & Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE*, *6*(9), e24893. doi: 10.1371/journal.pone.0024893

Müller, V., Sänger, J., & Lindenberger, U. (2013). Intra- and inter-brain synchronization during musical improvisation on the guitar. *PLoS ONE*, *8*(9), e73852.

Nessler, J. A., & Gilliland, S. J. (2010). Kinematic analysis of side-by-side stepping with intentional and unintentional synchronization. *Gait & Posture*, *31*(4), 527–529.

Novembre, G., Knoblich, G., Dunne, L., & Keller, P. E. (2017). Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Social Cognitive and Affective Neuroscience*, *12*(4), 662–670.

Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex*, *25*(3), 736–747.

Pan, Y., Novembre, G., Song, B., Zhu, Y., & Hu, Y. (2021). Dual brain stimulation enhances interpersonal learning through spontaneous movement synchrony. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 210–221.

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*(10), 827–830.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2003). *Synchronization: A universal concept in nonlinear sciences* (Vol. *12*). Cambridge University Press.

Sänger, J., Müller, V., & Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Frontiers in Human Neuroscience*, *6*, 312.

Sänger, J., Müller, V., & Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Frontiers in Human Neuroscience*, *7*, 234.

Wass, S. V., Whitehorn, M., Haresign, I. M., Phillips, E., & Leong, V. (2020). Interpersonal neural entrainment during early social interaction. *Trends in Cognitive Sciences*, *24*(4), 329–342.

Zamm, A., Debener, S., Bauer, A-K.R., Bleichner, M.G., Demos, A.P., & Palmer, C. (2018). Amplitude envelope correlations measure synchronous cortical oscillations in performing musicians. *Annals of the New York Academy of Sciences*, *1423*(1), 251–263.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews. Neuroscience*, *8*(7), 547–558.

An evolutionary theory of music needs to care about developmental timing

Erin E. Hannon[a], Alyssa N. Crittenden[b], Joel S. Snyder[a] and Karli M. Nave[a]

Departments of [a]Psychology and [b]Anthropology, University of Nevada, Las Vegas, NV 89154-5030, USA.
erin.hannon@unlv.edu
joel.snyder@unlv.edu
alyssa.crittenden@unlv.edu
nave@unlv.nevada.edu

doi:10.1017/S0140525X20001168, e74

Abstract

Both target papers cite evidence from infancy and early childhood to support the notion of human musicality as a somewhat static suite of capacities; however, in our view they do not adequately acknowledge the critical role of developmental timing, the acquisition process, or the dynamics of social learning, especially during later periods of development such as middle childhood.

Both Savage et al.'s and Mehr et al.'s articles draw on evidence from infancy and early childhood to strengthen their arguments, yet in our view they do not adequately acknowledge the critical importance of developmental timing especially during middle childhood. Both papers take a somewhat static approach to describing musical abilities during childhood, essentially equating the very presence of any sensitivity to a musical structure during early childhood with the adult form of that capacity. Coordinated group behaviors are at the center of both papers, and both assume that because the newborn brain responds to missing events in a drum pattern, the capacity for drumming and dance are present at birth. Although they acknowledge the slower developmental trajectory of precise entrainment, they nevertheless conclude that "the capacity to perceive and move to a beat/dance is a mode of participation that is accessible to large numbers of individuals regardless of age, familiarity with the music, or instrumental/singing virtuosity." We argue that these abilities are not necessarily accessible regardless of age or familiarity, and that a more nuanced account of *when* and *how* musical abilities and behaviors develop may hold the key to understanding the co-evolution of musicality and music.

Framing musicality as a gradually gated developmental process rather than a static suite of capacities could reveal why some cultural variants are favored over others, and what functions are served by particular musical features and behaviors. Similar to language, musical systems are transmitted across generations, so knowing something about who is learning and when learning occurs may be important for characterizing the pressures that act upon those systems. Because musical and linguistic systems only survive if children can learn them, systems must maximize transmissibility by adapting to the capacities of the learner and the acquisition process. In language evolution, children's memory limitations and learning biases may give rise to increased grammatical complexity in languages that are learned or created primarily by children (Christiansen & Chater, 2008; Senghas, Kita, & Özyürek, 2004), and reduced complexity in populations with more adult learners (Lupyan & Dale, 2010). In music, a recent study suggests rhythms also evolve as they are learned, reproduced, and transmitted across multiple "generations" of adult participants, and initially random sequences become increasingly structured, learnable, and consistent with observed musical universals (Ravignani, Delgado, & Kirby, 2016). If children learn differently than adults, then understanding these developmental constraints could explain how and why certain features exist across diverse musical systems.

Drumming and dancing are emphasized by both target papers, but if these abilities are assumed to be drivers of evolutionary processes, understanding their developmental timing is as important as systematic examination of variation across cultures. In order to dance to music, a listener must infer the beat and meter from cues in the musical surface, internally sustain that structure, and coordinate movements with the music and others in culturally appropriate ways. Although we agree that detecting changes, discriminating rhythms, and moving to music are precursors to mature adult behaviors, a growing body of evidence suggests that childhood capacities are starkly different than those observed during adulthood. Although some studies describe preschoolers performing "synchronous drumming" whenever drumming is not random (e.g., Kirschner and Tomasello, 2009), other evidence suggests children's synchronization to music does not approach adult-like levels until late childhood (McAuley, Jones, Holub, Johnston, & Miller, 2006; see Hannon, Nave-Blodgett, & Nave, 2018). Similarly, although children can tell when a metronome matches the music, they are less accurate than adults and show no sensitivity to simultaneous metrical levels until early adolescence (Einarson & Trainor, 2016; Nave-Blodgett, Hannon, & Snyder, 2020). These findings parallel recent evidence that language learning may have a more protracted developmental time course than previously assumed (Hartshorne, Tenenbaum, & Pinker, 2018; McMurray, Danelz, Rigler, & Seedorff, 2018).

Although we agree that musical behaviors play an important role during infancy and early childhood, middle childhood is also a critical period of development. During this time, children exhibit an increase in the ability to learn and transmit knowledge across myriad domains (e.g., cognitive, behavioral, and social) as they begin to physically and hormonally mature (Del Giudice, 2009; Lancy & Grove, 2011) and better understand their role within a larger social network (House et al., 2020). The transition to middle childhood is characterized by increased motor control and complex reasoning skills (Weisner, 1996) that emerge alongside increases in prosociality (Crittenden & Zes, 2015; Fehr, Glätzle-Rützler, & Sutter, 2013). Evolutionary models of childhood have proposed that middle childhood is a "developmental switch point" where genetic and environmental inputs converge to act on phenotypes (Del Giudice, 2009). This stage is characterized by heightened sensitivity to the environment, potentially mediated by the expression of new genetic or epigenetic factors (Del Giudice, 2018), and is accompanied by the acquisition of large amounts of knowledge through imitation, teaching, and other forms of cultural learning (Hewlett, Fouts, Boyette, & Hewlett, 2011; Kline, Boyd, & Henrich, 2013).

In music, this learning may entail increased awareness of metrical hierarchies and a stronger grasp of the relationship between the musical surface, meter, and rhythmic movements in one's culture. For example, in Western music, events are statistically more likely to occur at strong metrical positions, but recent study suggests that this is not the case for some non-Western corpora (Holzapfel, 2015; London, Polak, & Jacoby, 2016). If the ability to perceive meter does not emerge until middle childhood or adolescence (Nave-Blodgett et al., 2020), this means that the capacity to participate in more complex forms of musical coordination (such as group or partner dances) is only just emerging at a time when individuals are refining their social skills and paying greater attention to social models. Importantly, this suggests that capacity to dance varies with age and familiarity, and it undermines the notion that displays of group synchrony are perceptible to all on-lookers. We advocate for a more thorough treatment of existing developmental evidence, including incorporation of middle childhood as a key developmental period, and further research addressing how and when individuals learn and master specific musical structures and behaviors across cultures.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Christiansen, M. C., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, 489–558. https://doi.org/10.1017/S0140525X08004998.

Crittenden, A. N., & Zes, D. A. (2015). Food sharing among Hadza hunter-gatherer children. *PLoS ONE*, *10*(7), e0131996. https://doi.org/10.1371/journal.pone.0131996.

Del Giudice, M. (2009). Sex, attachment, and the development of reproductive strategies. *Behavioral and Brain Sciences*, *32*, 1–21. https://doi.org/10.1017/S0140525X09000181.

Del Giudice, M. (2018). Middle childhood: An evolutionary-developmental synthesis. In Halfon, N., Forrest, C. B., Lerner, R. M. & Faustman, E. M. (Eds.) *Handbook of life course health development* (pp. 95–107). Springer. https://doi.org/10.1007/978-3-319-47143-3.

Einarson, K. M., & Trainor, L. J. (2016). Hearing the beat: Young children's perceptual sensitivity to beat alignment varies according to metric structure. *Music Perception*, *34*, 56–70. https://doi.org/10.1525/mp.2016.34.1.56.

Fehr, E., Glätzle-Rützler, D., & Sutter, M. (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, *64*, 369–383. https://doi.org/10.1016/j.euroecorev.2013.09.006.

Hannon, E. E., Nave-Blodgett, J. E., & Nave, K. M. (2018). The developmental origins of the perception and production of musical rhythm. *Child Development Perspectives*, *12*, 194–198. https://doi.org/10.1111/cdep.12285.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277. https://doi.org/10.1016/j.cognition.2018.04.007.

Holzapfel, A. (2015). Relation between surface rhythm and rhythmic modes in Turkish Makam music. *Journal of New Music Research*, *44*, 25–38. https://doi.org/10.1080/09298215.2014.939661.

Hewlett, B. S., Fouts, H. N., Boyette, A. H., & Hewlett, B. L. (2011). Social learning among Congo Basin hunter–gatherers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1168–1178. https://doi.org/10.1098/rstb.2010.0373.

House, B. R., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebioglu, S., Crittenden, A. N., … Yilmaz, S. (2020). Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour*, *4*(1), 36–44. https://doi.org/10.1038/s41562-019-0734-z.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*, 299–314. https://doi.org/10.1016/j.jecp.2008.07.005.

Kline, M. A., Boyd, R., & Henrich, J. (2013). Teaching and the life history of cultural transmission in Fijian villages. *Human Nature*, *24*(4), 351–374. https://doi.org/10.1007/s12110-013-9180-1.

Lancy, D. F., & Grove, M. A. (2011). Getting noticed. *Human Nature*, *22*(3), 281. https://doi.org/10.1007/s12110-011-9117-5.

London, J., Polak, R., & Jacoby, N. (2016). Rhythm histograms and musical meter: A corpus study of Malian percussion music. *Psychonomic Bulletin & Review*, *24*, 474–480. https://doi.org/10.3758/s13423-016-1093-7.

Lupyan, G. & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559. https://doi.org/10.1371/journal.pone.0008559.

McAuley, J., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367. https://doi.org/10.1037/0096-3445.135.3.348.

McMurray, B., Danelz, A., Rigler, H., & Seedorff, M. (2018). Speech categorization develops slowly through adolescence. *Developmental Psychology*, *54*, 1472–1491. https://doi.org/10.1037/dev0000542.

Nave-Blodgett, J. E., Hannon, E. E., & Snyder, J. S. (2020). Hierarchical beat perception develops throughout childhood and adolescence and is enhanced in those with musical training. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000903.

Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, *1*(1), 1–7. http://doi.org/10.1038/s41562-016-0007.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*, 1779–1782. https://doi.org/10.1126/science.1100199.

Weisner, T. S. (1996). The 5 to 7 transition as an ecocultural project. In A. J. Sameroff & M. M. Haith (Eds.), *The John D. and Catherine T. MacArthur Foundation series on mental health and development. The five to seven year shift: The age of reason and responsibility* (pp. 295–326). University of Chicago Press.

Oxytocin as an allostatic agent in the social bonding effects of music

Niels Chr. Hansen[a,b] and Peter E. Keller[c]

[a]Aarhus Institute of Advanced Studies, Aarhus University, Høegh-Guldbergs Gade 6B, 8000 Aarhus C, Denmark; [b]Center for Music in the Brain, Aarhus University & Royal Academy of Music Aarhus/Aalborg, Aarhus, Denmark and [c]The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Penrith, NSW 2751, Australia.
nchansen@aias.au.dk; http://www.nielschrhansen.com/;
p.keller@westernsydney.edu.au; http://pkpublications.weebly.com/

doi:10.1017/S0140525X20001235, e75

Abstract

Despite acknowledging that musicality evolved to serve multiple adaptive functions in human evolution, Savage et al. promote social bonding to an overarching super-function. Yet, no unifying neurobiological framework is offered. We propose that oxytocin constitutes a socio-allostatic agent whose modulation of sensing, learning, prediction, and behavioral responses with reference to the physical and social environment facilitates music's social bonding effects.

Both target articles acknowledge that human musicality evolved to serve more than one adaptive function. Although Mehr, Krasnow, Bryant, and Hagen exalt credible signaling to a unitary mechanism, they eventually promote two ultimate-level explanations in terms of credible signaling of coalition strength and parental attention, manifesting in divergent forms of music (i.e., loud drumming/chanting vs. lullabies). In this way, one of their arguments opposing Savage et al.'s *music and social bonding* (MSB) hypothesis – namely that (some types of) musics are poorly designed to coordinate groups – can be turned against themselves in that lullabies are poorly designed to signal coalition strength and war chants are poorly designed to signal parental attention. This functional multiplicity leads Savage and colleagues to search for an overarching super-function through synthesis and extension of previous proposals.

Savage et al.'s MSB account is convincing in many aspects. Yet, their overview of candidate neurobiological mechanisms exposes the somewhat rudimentary state of research in this area. For example, their Figure 3 treats the oxytocinergic and endogenous opioid systems under the same umbrella despite comprising distinct – albeit interconnected (Kovatsi & Nikolaou, 2019) – systems. In light of oxytocin's well-established role in social learning and memory (Chini, Leonzino, Braida, & Sala, 2014), direct links of oxytocin to learning/knowledge (e.g., through laryngeal motor cortex, ventral tegmental area, or Broca's area; Theofanopoulou, Boeckx, & Jarvis, 2017) could also be considered in addition to indirect links via the dopaminergic system (Baskerville & Douglas, 2010; Love, 2014). Although the detailed interactions of these three systems lie beyond the scope of a brief commentary, it suffices to say that the overarching super-function envisioned by the target authors indeed requires a unifying neurobiological framework.

Given its manifold roles in a multitude of peripheral and central nervous processes with direct relevance for the proposed adaptive functions of music (as depicted in Fig. 1), we suggest that the nonapeptide oxytocin may contribute crucially to such a unifying neurobiological framework. Specifically, oxytocin-mediated increases in generosity (Zak, Stanton, & Ahmadi, 2007), emotional face recognition (Shahrestani, Kemp, & Guastella, 2013), and (potentially) trust (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; but see Declerck, Boone, Pauwels, Vogt, & Fehr, 2020) may increase music's effects on group cohesion resulting in social coping benefits. Changes in eye gaze (Eckstein et al., 2019), selective sociality (Carter, 2017), and sexual arousal and orgasm (Alley & Diamond, 2020) mediated by oxytocin may enhance musical effects on courtship behavior and exclusionary pair formation. Oxytocin-mediated empathy (Wu, Li, & Su, 2012), playful behavior (Szymanska, Schneider, Chateau-Smith, Nezelof, & Vulliez-Coady, 2017), and peripheral effects on parturition and lactation (Carter, 2014) could underlie musical infant–caregiver bonding. Moreover, oxytocin-related increases in synchrony (Gebauer et al., 2016; Josef, Goldstein, Mayseless, Ayalon, & Shamay-Tsoory, 2019) along with out-group gloating (Shamay-Tsoory et al., 2009), derogation (De Dreu, Greer, Van Kleef, Shalvi, & Handgraaf, 2011), and threat perception (Egito, Nevat, Shamay-Tsoory, & Osório, 2020) may promote competitive musical coalition signaling. Oxytocin's analgesic (González-Hernández, Rojas-Piloni, & Condés-Lara, 2014) and anxiolytic (Lancaster et al., 2018) effects may, in turn, enhance music's facilitation of coordinated physical labor. Finally, oxytocin-enhanced social learning and memory (Graustella & MacLeod, 2012) could increase music's capacity for preserving civilizational knowledge and forming shared cultural identities.

Importantly, the two target theories propose opposite causal relationships between oxytocin and musical synchronization. Although MSB predicts that joint music making would increase oxytocin levels, the credible signaling theory predicts that higher oxytocin levels achieved through pre-existing social cohesion would increase musical synchronization. Burgeoning evidence for both of these effects (*former*: Grape, Sandgren, Hansson, Ericson, & Theorell, 2002; Keeler et al., 2015; *latter*: Gebauer et al., 2016; Josef et al., 2019) suggests that both hypotheses may ultimately be viable. Yet, findings are sometimes contradictory (e.g., Schladt et al., 2017), sample sizes are typically low (Walum, Waldman, & Young, 2016), and intranasal administration and salivary assays of oxytocin have faced methodological criticisms (Leng & Ludwig, 2016; McCullough, Churchland, & Mendez, 2013). Theorizing, therefore, converges on oxytocin effects being highly individualistic/context-dependent (Bartz, Zaki, Bolger, & Ochsner, 2011) and modulatory/interactive rather than primary (Marsh, Marsh, Lee, & Hurlemann, 2020).

Consistent with this view, an allostatic theory of oxytocin has recently gained traction (Quintana & Guastella, 2020). This theory posits that oxytocin facilitates the adjustment of sensing and response set-points and assists learning and prediction in ways that proactively optimize systemic adaptation with reference to a constantly changing environment. The resulting criticality in brain states and behavioral flexibility are beneficial under many circumstances – but perhaps especially so in complex social interactions such as music and dance where ongoing anticipation of environmental consequences is essential (D'Ausilio, Novembre, Fadiga, & Keller, 2015). Beyond the individual level, oxytocin's allostatic effects may further extend to social allostasis in that behavioral flexibility can resolve vicious circles of rigid interaction patterns and inflexible group dynamics (Saxbe, Beckes, Stoycos, & Coan, 2020; Schulkin, 2011). A socio-allostatic view of oxytocin, moreover, connects well with intellectual currents such as enactive cognition and biocultural evolution where biological and cultural dimensions constitute factors in the same evolving system of living agents embedded in an environment (van der Schyff & Schiavio, 2017).

Savage et al. (2020) propose that music and musicality coevolved through an iterated Baldwin process whereby musical behaviors were culturally invented to tackle the environmental challenges of social living which gradually manifested as adaptive musicality traits (Podlipniak, 2017). The neurochemical feedback loop depicted in Figure 1 is consistent with such a process of gene-culture coevolution in that genetic mutations leading to tighter connections between music making and oxytocin release would accelerate the extent to which the adaptive functions of music can capitalize upon oxytocin-mediated sensing, learning, prediction, and response processes. The proposed feedback mechanism whereby oxytocin in itself stimulates more oxytocin release (Grippo et al., 2012) may further accelerate this process.

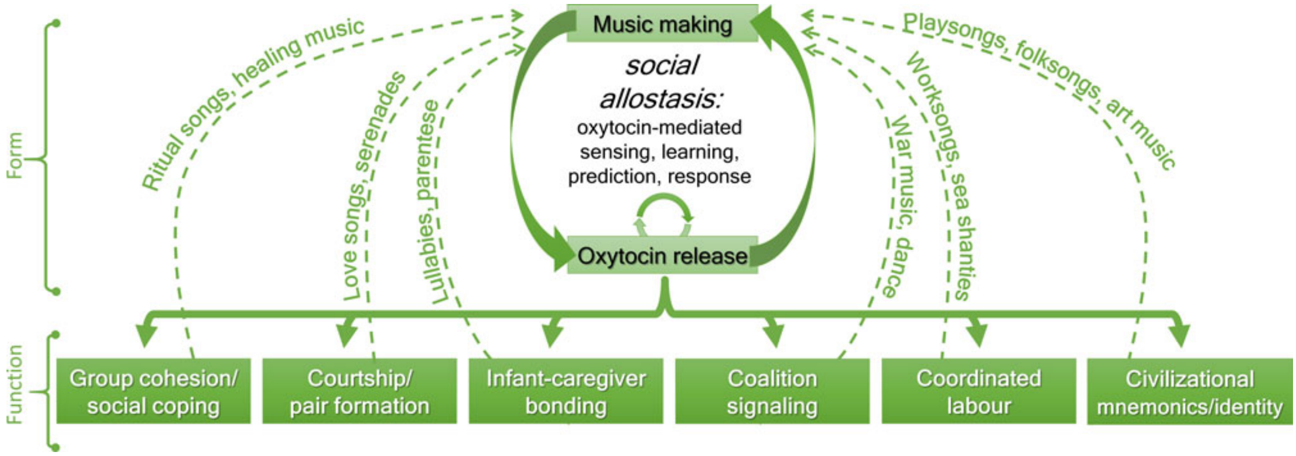


Figure 1. (Hansen & Keller) Overview of how the socio-allostatic effects of oxytocin may contribute to a unifying neurobiological framework for music's adaptive functions. Specifically, music making – which often occurs in social contexts – may lead to oxytocin release that promotes internal adjustment of sensing and response set-points as well as assists learning and prediction processes with reference to the external physical and social environment. These allostatic optimizations may, in turn, increase the efficacy of music's adaptive functions as a facilitator of social cohesion, pair formation, parental bonding, coalition signaling, group coordination, and preservation of communal knowledge and identity. These adaptive functions of music are each associated with specific musical forms employing distinct acoustic features. Such function-to-form associations are, however, statistical (rather than deterministic) as well as highly interconnected in that dance, for example, can contribute to multiple adaptive functions such as cohesion, courtship, coalition, and coordination, depending on the context in which it appears.

When oxytocin-mediated sensing, learning, prediction, and response capacities are sharpened in the encounter with the in-group, phenomena such as social cohesion arise, which in turn promote coordinated labor as well as civilizational mnemonics and identity formation (Fig. 1). When faced with an out-group, heightened social allostasis may manifest in coalition signaling. Finally, when presented with offspring or potential mates, infant–caregiver bonding and dyadic courtship result, respectively. These effects could be mediated by the manifold empirically substantiated effects of oxytocin on human behavior summarized above.

Financial support. NCH received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement (No. 754513), The Aarhus University Research Foundation, Carlsberg Foundation (CF18-0668), and Lundbeck Foundation (R266-2017-3339).

Conflict of interest. None.

References

Alley, J. C., & Diamond, L. M. (2020). Oxytocin and human sexuality: Recent developments. *Current Sexual Health Reports*, *12*, 182–185. doi: 10.1007/s11930-020-00274-4.

Bartz, J. A., Zaki, J., Bolger, N., & Ochsner, K. N. (2011). Social effects of oxytocin in humans: Context and person matter. *Trends in Cognitive Sciences*, *15*(7), 301–309. doi: 10.1016/j.tics.2011.05.002.

Baskerville, T. A., & Douglas, A. J. (2010). Dopamine and oxytocin interactions underlying behaviors: Potential contributions to behavioral disorders. *CNS Neuroscience & Therapeutics*, *16*(3), e92–e123. doi: 10.1111/j.1755-5949.2010.00154.x.

Carter, C. S. (2014). Oxytocin pathways and the evolution of human behavior. *Annual Review of Psychology*, *65*, 17–39. doi: 10.1146/annurev-psych-010213-115110.

Carter, C. S. (2017). The role of oxytocin and vasopressin in attachment. *Psychodynamic Psychiatry*, *45*(4), 499–517. doi: 10.1521/pdps.2017.45.4.499.

Chini, B., Leonzino, M., Braida, D., & Sala, M. (2014). Learning about oxytocin: Pharmacologic and behavioral issues. *Biological Psychiatry*, *76*(5), 360–366. doi: 10.1016/j.biopsych.2013.08.029.

D'Ausilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, *19*(3), 111–114. doi: 10.1016/j.tics.2015.01.005.

De Dreu, C. K., Greer, L. L., Van Kleef, G. A., Shalvi, S., & Handgraaf, M. J. (2011). Oxytocin promotes human ethnocentrism. *Proceedings of the National Academy of Sciences*, *108*(4), 1262–1266. doi: 10.1073/pnas.1015316108.

Declerck, C. H., Boone, C., Pauwels, L., Vogt, B., & Fehr, E. (2020). A registered replication study on oxytocin and trust. *Nature Human Behaviour*, *4*, 646–655. doi: 10.1038/s41562-020-0878-x.

Eckstein, M., Bamert, V., Stephens, S., Wallen, K., Young, L. J., Ehlert, U., & Ditzen, B. (2019). Oxytocin increases eye-gaze towards novel social and non-social stimuli. *Social Neuroscience*, *14*(5), 594–607. doi: 10.1080/17470919.2018.1542341.

Egito, J. H., Nevat, M., Shamay-Tsoory, S. G., & Osório, A. A. C. (2020). Oxytocin increases the social salience of the outgroup in potential threat contexts. *Hormones and Behavior*, *122*, 104733. doi: 10.1016/j.yhbeh.2020.104733.

Gebauer, L., Witek, M. A., Hansen, N. C., Thomas, J., Konvalinka, I., & Vuust, P. (2016). Oxytocin improves synchronisation in leader-follower interaction. *Scientific Reports*, *6*, 38416. doi: 10.1038/srep38416.

González-Hernández, A., Rojas-Piloni, G., & Condés-Lara, M. (2014). Oxytocin and analgesia: Future trends. *Trends in Pharmacological Sciences*, *35*(11), 549–551. doi: 10.1016/j.tips.2014.09.004.

Grape, C., Sandgren, M., Hansson, L. O., Ericson, M., & Theorell, T. (2002). 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. doi: 10.1007/BF02734261.

Graustella, A. J., & MacLeod, C. (2012). A critical review of the influence of oxytocin nasal spray on social cognition in humans: Evidence and future directions. *Hormones and Behavior*, *61*(3), 410–418. doi: 10.1016/j.yhbeh.2012.01.002.

Grippo, A. J., Pournajafi-Nazarloo, H., Sanzenbacher, L., Trahanas, D. M., McNeal, N., Clarke, D. A., … Sue Carter, C. (2012). Peripheral oxytocin administration buffers autonomic but not behavioral responses to environmental stressors in isolated prairie voles. *Stress*, *15*(2), 149–161. doi: 10.3109/10253890.2011.605486.

Josef, L., Goldstein, P., Mayseless, N., Ayalon, L., & Shamay-Tsoory, S. G. (2019). The oxytocinergic system mediates synchronized interpersonal movement during dance. *Scientific Reports*, *9*(1), 1–8. doi: 10.1038/s41598-018-37141-1.

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*, 518. doi: 10.3389/fnhum.2015.00518.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*(7042), 673–676. doi: 10.1038/nature03701.

Kovatsi, L., & Nikolaou, K. (2019). Opioids and the hormone oxytocin. *Vitamins and Hormones*, *111*, 195–225. doi: 10.1016/bs.vh.2019.05.003.

Lancaster, K., Goldbeck, L., Pournajafi-Nazarloo, H., Connelly, J. J., Carter, C. S., & Morris, J. P. (2018). The role of endogenous oxytocin in anxiolysis: Structural and functional correlates. *Biological Psychiatry*, *3*(7), 618–625. doi: 10.1016/j.bpsc.2017.10.003.

Leng, G., & Ludwig, M. (2016). Intranasal oxytocin: Myths and delusions. *Biological Psychiatry*, *79*(3), 243–250. doi: 10.1016/j.biopsych.2015.05.003.

Love, T. M. (2014). Oxytocin, motivation and the role of dopamine. *Pharmacology Biochemistry and Behavior*, *119*, 49–60. doi: 10.1016/j.pbb.2013.06.011.

Marsh, N., Marsh, A. A., Lee, M. R., & Hurlemann, R. (2020). Oxytocin and the neurobiology of prosocial behavior. *The Neuroscientist* *15*(2), 469–482. doi: 10.1177/1745691619876528.

McCullough, M. E., Churchland, P. S., & Mendez, A. J. (2013). Problems with measuring peripheral oxytocin: Can the data on oxytocin and human behavior be trusted? *Neuroscience & Biobehavioral Reviews*, *37*(8), 1485–1492. doi: 10.1016/j.neubiorev.2013.04.018.

Podlipniak, P. (2017). The role of the Baldwin effect in the evolution of human musicality. *Frontiers in Neuroscience*, *11*, 542. doi: 10.3389/fnins.2017.00542.

Quintana, D. S., & Guastella, A. J. (2020). An allostatic theory of oxytocin. *Trends in Cognitive Sciences*, *24*(7), 515–528. doi: 10.1016/j.tics.2020.03.008.

Saxbe, D. E., Beckes, L., Stoycos, S. A., & Coan, J. A. (2020). Social allostasis and social allostatic load: A new model for research in social dynamics, stress, and health. *Perspectives on Psychological Science*, *15*(2), 469–482. doi: 10.1177/1745691619876528.

Schladt, T. M., Nordmann, G. C., Emilius, R., Kudielka, B. M., de Jong, T. R., & Neumann, I. D. (2017). Choir versus solo singing: Effects on mood, and salivary oxytocin and cortisol concentrations. *Frontiers in Human Neuroscience*, *11*, 430. doi: 10.3389/fnhum.2017.00430.

Schulkin, J. (2011). Social allostasis: Anticipatory regulation of the internal milieu. *Frontiers in Evolutionary Neuroscience*, *2*, 111. doi: 10.3389/fnevo.2010.00111.

Shahrestani, S., Kemp, A. H., & Guastella, A. J. (2013). The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: A meta-analysis. *Neuropsychopharmacology*, *38*(10), 1929–1936. doi: 10.1038/npp.2013.86.

Shamay-Tsoory, S. G., Fischer, M., Dvash, J., Harari, H., Perach-Bloom, N., & Levkovitz, Y. (2009). Intranasal administration of oxytocin increases envy and schadenfreude (gloating). *Biological Psychiatry*, *66*(9), 864–870. doi: 10.1016/j.biopsych.2009.06.009.

Szymanska, M., Schneider, M., Chateau-Smith, C., Nezelof, S., & Vulliez-Coady, L. (2017). Psychophysiological effects of oxytocin on parent-child interactions: A literature review on oxytocin and parent–child interactions. *Psychiatry and Clinical Neurosciences*, *71*(10), 690–705. doi: 10.1111/pcn.12544.

Theofanopoulou, C., Boeckx, C., & Jarvis, E. D. (2017). A hypothesis on a role of oxytocin in the social mechanisms of speech and vocal learning. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1861), 20170988. doi: 10.1098/rspb.2017.0988.

van der Schyff, D., & Schiavio, A. (2017). Evolutionary musicology meets embodied cognition: Biocultural coevolution and the enactive origins of human musicality. *Frontiers in Neuroscience*, *11*, 519. doi: 0.3389/fnins.2017.00519.

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. doi: 10.1016/j.biopsych.2015.06.016.

Wu, N., Li, Z., & Su, Y. (2012). The association between oxytocin receptor gene polymorphism (OXTR) and trait empathy. *Journal of Affective Disorders*, *138*(3), 468–472. doi: 10.1016/j.jad.2012.01.009.

Zak, P. J., Stanton, A. A., & Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS ONE*, *2*(11), e1128. doi: 10.1371/journal.pone.0001128.

Against unitary theories of music evolution

Peter M. C. Harrison[a] and Madeleine Seale[b]

[a]Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany and [b]Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK.
peter.harrison@ae.mpg.de; https://www.pmcharrison.com/
maddy.seale@plants.ox.ac.uk;
https://www.plants.ox.ac.uk/people/madeleine-seale

doi:10.1017/S0140525X20001314, e76

Abstract

Savage et al. and Mehr et al. provide well-substantiated arguments that the evolution of musicality was shaped by adaptive functions of social bonding and credible signalling. However, they are too quick to dismiss byproduct explanations of music evolution, and to present their theories as complete unitary accounts of the phenomenon.

The primary purpose of studying music's evolution is to understand the selective pressures that caused different musical traits to evolve. When studying the evolution of a given musical trait, the first question to ask is whether the trait can be parsimoniously explained as a byproduct of non-musical adaptive functions. An example might be ears: Ears are indisputably useful for appreciating music, yet in all likelihood the reason that we possess ears is for general auditory perception rather than music listening. If no byproduct explanation is forthcoming, the way forward is clear for an alternative hypothesis explaining the musical trait in terms of a particular adaptive function of music.

Although the two target articles by Mehr et al. and Savage et al. ultimately come to contradictory explanations for music's evolution, their approaches share certain commonalities. First, both articles minimize the sense in which musical traits may be explained as byproducts of non-musical adaptive functions. Second, both articles present unified theories purporting to explain the evolution of musicality in terms of music's own adaptive functions.

Both articles provide substantive positive arguments for their respective theories, making it clear that both social bonding and credible signalling are plausible adaptive functions of music-making with broad cross-cultural and historical relevance. However, we would push back against the claims that (a) byproduct explanations have little relevance for music evolution, and that (b) the articles' theories provide complete unified accounts of music evolution.

Savage et al. address byproduct explanations towards the end of their article, writing that "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." However, they conclude against this possibility, writing that the cross-cultural universality of music and language implies that music and language fulfil *independent* functions. This argument holds little water; if one believed that music were a byproduct of language, then music's universality would simply be a byproduct of language's universality.

Mehr et al.'s arguments against byproduct explanations are numerous but weak. Music's ancient nature, cross-cultural universality, early developmental manifestation, and complex language-like syntax could surely be explained by music being the byproduct of ancient, cross-culturally universal, and early manifesting capacities such as language cognition. The convergent evolution of music-like behaviours in other species could likewise be the consequence of music being a byproduct of other adaptive traits; for example, Patel (2014) argues that cross-species rhythmic capacities may be a byproduct of vocal learning abilities. Finally, much of music's neural specialization might be explained by neural plasticity (e.g., Münte, Altenmüller, & Jäncke, 2002).

Meanwhile, having minimized the role of byproduct explanations, Savage et al. conclude that their theory of music-induced social bonding provides a "unified" and "comprehensive theory" that "synthesize[s] and extend[s] previous proposals into a new, parsimonious framework." The implication is that there is no space for alternative explanations for music's evolution, such as credible signalling or sexual selection; however, the authors do not present any arguments against these alternative explanations. Section 6 rather muddies the waters, with credible signalling and sexual selection being described as both "complementary" and "competing" with respect to the authors' own hypothesis. It seems as if the authors want to claim that their theory provides a complete account of music evolution, but simultaneously want to avoid the obligation of disproving competing theories.

In contrast, Mehr et al. do systematically attempt to disprove competing explanations of music evolution, including the byproduct hypothesis as discussed above. Their critique of Darwin's sexual selection theory is particularly convincing, but their critique of Savage et al.'s social-bonding hypothesis is problematic. The authors interpret social bonding as altruism, and make the important observation that music-induced altruism would be counteracted by free-rider mutations. However, social bonding does not have to be altruistic in nature. If music-making has the byproduct of enhancing my ability to physically coordinate and nonverbally communicate with my fellow tribespeople, with the consequence that I'm less likely to die on our next hunting exhibition, then my participation is not altruistic, because I benefit myself as well as the group. The authors also claim that music offers no clear benefits for inducing social bonding over language; this might be the case if social bonding is operationalized as altruism, but it is not true if we consider more complex aspects of social bonding, such as developing joint coordination and communication skills, for which music seems to be particularly well adapted (see e.g., Cross & Morley, 2008).

We are left wondering why both articles are so keen to downplay byproduct explanations for musicality, and to promote their own theories as complete unitary explanations for musicality. On the one hand, Occam's razor tells us to prefer simple theories, and as music researchers we have a vested interest in discounting byproduct explanations that might be seen to downgrade music's evolutionary importance. On the other hand, a broader biological perspective provides countless examples of evolved traits that were shaped by a variety of selection pressures, some with only tangential relevance to the trait's ultimate function. For example, spider silk initially evolved to line burrows, but in most extant spider species webs are formed for other functions such as prey trapping, pheromone release, water capture, and dispersal parachuting (Duffey, 1998; Nentwig & Heimer, 1987; Vollrath & Edmonds, 1989). It seems quite plausible that music's evolution could have similarly complex explanations.

Financial support. Madeleine Seale was funded by a Leverhulme Early Career Research fellowship.

Conflict of interest. None.

References

Cross, I., & Morley, I. (2008). The evolution of music: Theories, definitions and the nature of the evidence. In Malloch, S. & Trevarthen C. (Eds.), *Communicative musicality* (pp. 61–81). Oxford University Press.

Duffey, E. (1998). Aerial dispersal in spiders. In P. A. Selden (ed.). *Proceedings of the 17th European Colloquium of Arachnology* (pp. 187–191) British Arachnological Society.

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, 3(6), 473–478.

Nentwig, W., & Heimer, S. (1987). *Ecological aspects of spider webs in ecophysiology of spiders*. Springer-Verlag.

Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was Darwin wrong? *PLoS Biology*, *12*(3). https://doi.org/10.1371/journal.pbio.1001873.

Vollrath, F., & Edmonds, D. T. (1989). Modulation of the mechanical properties of spider silk by coating with water. *Nature*, *340*, 305–307.

Bonding system in nonhuman primates and biological roots of musicality

Yuko Hattori

Primate Research Institute, Kyoto University, Center for International Collaboration and Advanced Studies in Primatology (CICASP), Kanrin, Inuyama, Aichi 484-8506, Japan.
hattori.yuko.8w@kyoto-u.ac.jp;
http://www.cicasp.pri.kyoto-u.ac.jp/people/yuko-hattori

doi:10.1017/S0140525X2000148X, e77

Abstract

Comparative studies of primates indicate that humans have evolved unique motivations and cognitive skills for sharing emotions, experiences, and collaborative actions. Given the characteristics of music, the music and social bonding (MSB) hypothesis by Savage et al. fits this view. Within a cross-species approach, predispositions not observed in current communication system may contribute to a better understanding of the biological roots of human musicality.

Living in groups involves a balance of costs and benefits and, especially for primates, this inevitably entails competition and cooperation with the other group members. If food resources are scarce and all individuals are competitive, splitting into smaller groups will be favored. However, if group members cooperate with each other, for example by hunting together, larger groups become advantageous. Despite the high costs of investing in bonding relationships with non-kin, one prominent characteristic of primate society is the large and complex group memberships including non-kin group members (Byrne & Whiten, 1988; Harcourt & de Waal, 1992), and conducive social skills and bonding mechanisms have, therefore, evolved among primates (Van Schaik, 2016).

Social play, for example, generally follows a bell-shaped curve through ontogeny with a peak in the juvenile phase (Fagen, 1993); however, primates in particular continue to play as adults (Palagi, 2018). Some species such as great apes play frequently even when they become adults, and researchers have proposed that social play functions to share positive emotions (e.g., laughing) and reduce stress or tension (Palagi et al., 2006a, 2006b), effects that have parallels with music and dance. It is reported that the relative striatal volume is correlated with the rate of social (but not non-social) play behavior (Graham, 2011). The striatum structure is intimately connected to dopaminergic pathways, and thus to the experience and anticipation of pleasure and reward, as well as to sequential behavior, motor control, and cognitive flexibility (Ashby, Turner, & Horvitz, 2010; Elliott, Newman, Longe, & Deakin, 2003; Erikson et al., 2010; Vink et al., 2005). Given that social play, along with grooming, plays a crucial role in social bonding, which in turn may have fitness benefits (e.g., in chimpanzees, Muller & Mitani, 2005), selection may have favored a link between play and an enhanced ability to make sequential behaviors for rhythmic and repetitive joint actions with others, and experiencing pleasure and reward through those activities.

However, among primates, humans are extremely highly motivated to share psychological states with others (shared intentionality, Tomasello & Carpenter, 2007). As "togetherness" or "jointness" is what researchers consider to distinguish cooperative or collaborative interactions between humans and nonhuman primates (Tomasello & Moll, 2010), a bonding system such as music might have been selected for such purposes. Some research has led to the suggestion that coalition signaling did not evolve in most nonhuman primate societies because their bonding relationships are not strong enough, and so they succumb to the collective action problem (CAP) or social dilemma, when all individuals would benefit by cooperating but fail to do so because of conflicting interests that discourage joint action (Willems & Van Schaik, 2015). This suggests that credible signaling of coalition strength, size, and cooperation ability mentioned in Mehr et al. can evolve after group members have strongly bonded relationships.

Concerning the cross-species approach, the perception/production dissociation mentioned by Savage et al. should be possible to explore in nonhuman primates. Indeed, although imitation is relatively rare in nonhuman primates, they recognize when their actions are being imitated (Haun & Call, 2008; Paukner, Anderson, Borelli, Visalberghi, & Ferrari, 2005) and show more affiliative responses to those who recently imitated them (Paukner, Suomi, Visalberghi, & Ferrari, 2009). In vocal communication, some species such as Japanese macaques match acoustic features (i.e., frequency range) when they call back to other group members (Sugiura, 1998). However, whether acoustic similarity increases bonding in relationships has not been rigorously tested. Additionally, I suggest that abilities or predispositions related to music-making but which are not components of the current communication system, should also be considered for understanding the biological roots of musicality. For example, chimpanzees, but not monkeys, seem to be intrinsically motivated to match their body movement (i.e., tapping) to auditory rhythms (Hattori, Tomonaga & Matsuzawa, 2013; 2015; Zarco, Merchant, Prado, & Mendez, 2009), although it is a rudimentary capacity and entrainment occurs only when the auditory tempo is close to the individual's spontaneous motor tempo. We also recently reported that rhythmic body movements such as swaying can be induced in chimpanzees (Hattori, in press; Hattori & Tomonaga, 2020a, in press). A similar effect has been reported in the wild, in the form of chimpanzees' rhythmic displays and vigorous charges in reaction to rain or waterfalls (Goodall, 1986). Although neither the rain dance nor waterfall dance appears to be a joint activity involving multiple individuals, components of musicality may have arisen based on those predispositions in chimpanzees' and also humans' ancestors. As nonhuman animals do not themselves make music, any proto-musical behaviors have their own species-specific variations. Understanding such variations in relation to bonding systems is a laudable aim of further research in the comparative approach to the biology of music.

Financial support. This study was funded by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (B) 20H04490.

Conflict of interest. None.

References

Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in the Cognitive Sciences*, *14*, 208–215.

Byrne, R. W., & Whiten, A. (1988). *Machiavellian intelligence: Uniquely expertise and the evolution of intellect in monkeys, apes, and humans*. Clarendon Press.

Elliott, R., Newman, J. L., Longe, O. A., & Deakin, J. F. W. (2003). Differential response patterns in the striatum and orbitofrontal cortex in financial reward in humans: A parametric functional magnetic resonance imaging study. *Journal of Neuroscience*, *23*, 303–307.

Erikson, K. I., Boot, W. R., Basak, C., Neider, M. B., Prakesh, R. S., Voss, M. W., Graybiel, A. M., Simons, D. J., Fabiani, M., Gratton, G., & Kramer, A. F. (2010). Striatal volume predicts level of video game skill acquisition. *Cerebral Cortex*, *20*, 45–57.

Fagen, R. (1993). Primate juvenile and primate play. In M. E. Pereira & L. A. Fairbanks (Eds.), *Juvenile primates* (pp. 182–196). Oxford University Press.

Goodall, J. (1986). The chimpanzees of Gombe. In *Pattern of behavior* (p. 335). The Belknap Press of Harvard University Press.

Graham, K. L. (2011). Coevolutionary relationship between striatum size and social play in nonhuman primates. *American Journal of Primatology*, *73*(4), 314–322.

Harcourt, A. H., & de Waal, F. B. M. (1992). *Coalitions and alliances in humans and other animals*. Oxford University Press.

Hattori, Y. (in press). Behavioral coordination and synchronization in non-human primates. In J.R. Anderson & H. Kuroshima (Eds.), *Comparative cognition: Commonalities and diversity*. Springer.

Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2015). Distractor effect of auditory rhythms on self-paced tapping in chimpanzees and humans. *PLoS ONE*, *10*, e0130682.

Hattori, Y., & Tomonaga, M. (2020a). Rhythmic swaying induced by sound in chimpanzees (*Pan troglodytes*). *Proceeding of the National Academy of Sciences*, *117*(2), 936–942.

Hattori, Y., & Tomonaga, M. (in press). *Reply to Bertolo et al.: Rhythmic swaying in chimpanzees has implications for understanding the biological roots of human music and dance*. Proceeding of the National Academy of Sciences.

Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Scientific Report*, *3*, 1566.

Haun, D. B. M., & Call, J. (2008). Imitation recognition in great apes. *Current Biology*, *18*(7), R288–R290.

Muller, M. N., & Mitani, J. C. (2005). Conflict and cooperation in wild chimpanzees. *Advances in the Study of Behavior*, *35*, 275–331.

Palagi, E. (2018). Not just for fun! Social play as a springboard for adult social competence in human and non-human primates. *Behavioral Ecology and Sociobiology*, *72*(6), 90.

Palagi, E., Burghardt, G. M., Smuts, B., Cordoni, G., Dall'Olio, S., Fouts, H. N., Řeháková-Petrů, M., Siviy, S. M., & Pellis, S. M. (2016a). Rough- and tumble play as a window on animal communication. *Biological Reviews*, *91*, 311–327.

Palagi, E., Cordoni, G., Demuru, E., & Bekoff, M. (2016b). Fair play and its connection with social tolerance, reciprocity and the ethology of peace. *Behaviour*, *153*, 1195–1216.

Paukner, A., Anderson, J. R., Borelli, E., VOsalberghi, E., & Ferrari, P. F. (2005). Macaques (*Macaca nemestrina*) recognize when they are being imitated. *Biology Letters*, *1*(2), 219–222.

Paukner, A., Suomi, S. J., Visalberghi, E., & Ferrari, P. F. (2009). Capuchin monkeys display affiliation toward humans who imitate them. *Science*, *325*(5942), 880–883.

Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, *55*, 673–687.

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*(1), 121–125.

Tomasello, M. & Moll, H. (2010). *The gap is social: Human shared intentionality and culture*. Springer.

Van Schaik, CP. (2016). *The primate origins of human nature*. John Wiley & Sons.

Vink, M., Kahn, R. S., Raemaekers, M., van den Heuvel, M., Boersma, M., & Ramsey, N. F. (2005). Function of striatum beyond inhibition and execution of motor responses. *Human Brain Mapping*, *25*, 336–344.

Willems, E. P., & Van Schaik, C. P. (2015). Collective action and the intensity of between-group competition in non-human primates. *Behavioral Ecology*, *26*(2), 625–631.

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, *102*, 3191–3202.

Unravelling the origins of musicality: Beyond music as an epiphenomenon of language

Henkjan Honing

Amsterdam Brain & Cognition, Institute for Logic, Language and Computation,
University of Amsterdam, 1090 GE Amsterdam, The Netherlands.
honing@uva.nl
https://www.mcg.uva.nl/hh

doi:10.1017/S0140525X20001211, e78

Abstract

The two target articles address the origins of music in complementary ways. However, both proposals focus on overt musical behaviour, largely ignoring the role of perception and cognition, and they blur the boundaries between the potential origins of language and music. To resolve this, an alternative research strategy is proposed that focuses on the core cognitive components of musicality.

Savage et al. start with making the important distinction between *musicality* as a set of traits that allow us to perceive, produce, and appreciate music, and *music* as a social and cultural construct "generated by and for music making" (sect. 1, para. 3). This distinction seems trivial, but it adds a powerful level of explanation to the study of the origins of musicality as a phenomenon with both a cultural and a biological basis (Fitch, 2018; Honing, 2018b). As such, a good starting point for the search for these origins would be musicality, rather than music.

Although the cross-cultural study of the structure of *music* (melodic patterns, scales, tonality, etc.) has offered exciting insights (Mehr et al. 2019; Savage, Brown, Sakai, & Currie, 2015), the approach used in these studies is indirect: the object of study here is music – the result of musicality – rather than musicality itself. Hence, it is difficult to distinguish between the individual contributions of culture and biology. For example, it is not clear whether the division of an octave into small and unequal intervals in a particular musical culture results from a widespread theoretical doctrine or from a music perception ability or preference.

All this is an important motivation to study the structure of *musicality* (the capacity *for* music), its constituent components (see Table 1), and how these might be shared with other animals, aiming to disentangle the biological and cultural contributions to the human capacity for music (Honing, ten Cate, Peretz, & Trehub, 2015).

The two target articles address the origins of music and musicality in complementary ways. Savage et al. aim for an overarching theory that proposes music to be a relatively recent cultural invention that then further evolved through gene-culture coevolution. In contrast, Mehr et al. present a single hypothesis capturing the biological origins of musical behaviour, suggesting a long evolutionary history.

The first thing to note is that both articles base their arguments on overt musical behaviour (i.e., music production), with little or no attention to the perception and appreciation of music. This is surprising because there is quite a body of research that aims to identify the core constituent components of musicality by focusing on the perception and cognition of music (see Table 1). This in support of the important realization that we all share a predisposition for music: Even those of us who can't play a musical instrument or claim to lack a sense of rhythm can perceive and enjoy music (Margulis, 2019; Rentfrow & Levitin, 2019). As such, music production is not necessarily a good proxy for the perception and appreciation of music (cf. Trehub, Weiss, & Cirelli, 2019).

Second, both target articles struggle with identifying the role of language in how musical behaviour evolved.

Savage et al. decided to avoid a strict focus on musicality ("We make no claim that the mechanisms discussed here are entirely specific to music," sect. 6.1, para. 3). This could be considered a methodological weakness. Of course, music and musicality have been and will continue to be influenced by a variety of non-musical factors. Nevertheless, to be able to pinpoint what is essential to musicality, what components we share with other nonhuman animals, and what its potential evolutionary history is, it seems a more fruitful strategy to restrict oneself to those components of musicality for which it can be argued that they are *not* linked or useful to language (for instance, beat perception or tonality; see Table 1).

Table 1. (Honing) Potential candidates for a multicomponent model of musicality (cf. Honing, 2018a)

Domain	Constituent component
Melody	Relative pitch (pitch interval)
	Absolute pitch (fundamental frequency)
	Melodic contour (change of pitch direction)
	Tonal encoding of pitch (*tonality*)
Rhythm	Isochrony perception
	Beat perception
	Rhythmic contour (grouping)
	Metrical encoding of rhythm (*metricality*)
Timbre	Timbre perception
	Consonance perception (*harmonicity*)
	Spectral contour (change of timbre)

Mehr et al. are less explicit in what their hypothesis says about language versus music. A large proportion of the arguments appear to be equally applicable to the origins of speech. In fact, the credible signalling hypothesis is resonating with the idea of a "musical protolanguage" (Darwin, 1871; cf. Fitch, 2013) in interesting ways. Charles Darwin argued that language does not depend on the skill of being able to articulate sounds, but "obviously depends on the development of the mental faculties" (Darwin, 1871, p. 54). Hence, the credible signalling hypothesis could be improved by making explicit which core musical building blocks (i.e., those of lesser or no use to language) are linked to which music-specific mental faculties. As such making precise what is special about music.

Although the relation between language and music has been a topic of much debate (Arbib, 2013), in the current context it appears to be less relevant what these two domains have in common. What should be the focus is what makes the capacity for music *distinct* from that of language: the study of musicality is in need of its own research agenda.

Although both target articles note that some components of musicality overlap with non-musical cognitive features, this is in itself no evidence against musicality as a separately evolved biological set of traits. Theories that suggest musicality to be an epiphenomenon of language (Pinker, 1997) have to demonstrate that the components of musicality are *not* domain specific, but each cognitively linked to some non-musical mental ability.

As in language, musicality could have evolved from existing elements that are brought together in unique ways, and that system may still have emerged as a biological product through evolutionary processes. As such there is no need for musicality to show a modular structure (Fodor, 1983). Alternatively, converging evidence suggests music-specific responses along specific neural pathways (Albouy, Benjamin, Morillon, & Zatorre, 2020) and it could be that brain networks that support musicality are partly recycled for language (Peretz, Vuvan, Armony, Lagrois, &

Armony, 2018). This could imply that both language and music originate from musicality. In fact, this is one possible route to test the Darwin-inspired conjecture that musicality precedes music and language (Honing, 2018a).

Acknowledgments. I thank the members of the Music Cognition Reading Group for sharing their views on this topic, and Bas Cornelissen and Andrea Ravignani for their feedback on the manuscript improving its presentation.

Conflict of interest. None.

References

Albouy, P., Benjamin, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science*, *367*(6481), 16. https://doi.org/10.1126/science.1252826.

Arbib, M. A. (Ed.). (2013). *Language, music, and the brain: A mysterious relationship*. MIT Press.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. John Murray.

Fitch, W. T. (2013). Musical protolanguage: Darwin's theory of language evolution revisited. In J. J. Bolhuis & M. Everaert (Eds.), *Birdsong, speech, and language: Exploring the evolution of mind and brain* (pp. 489–503). MIT Press.

Fitch, W. T. (2018). Four principles of biomusicology. In H. Honing (Ed.), *The origins of musicality* (pp. 23–48). MIT Press.

Fodor, J. A. (1983). *The modularity of mind*. MIT Press. https://doi.org/10.2307/2184717.

Honing, H. (2018a). Musicality as an upbeat to music: Introduction and research agenda. In H. Honing (Ed.), *The origins of musicality* (pp. 3–20). MIT Press.

Honing, H. (Ed.). (2018b). *The origins of musicality*. The MIT Press.

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *370*(1664), 20140088. https://doi.org/10.1098/rstb.2014.0088.

Margulis, E. H. (2019). *The psychology of music: A very short introduction*. Oxford University Press.

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*, *366*(6468), 957–970. https://doi.org/10.1126/science.aax0868.

Peretz, I., Vuvan, D. T., Armony, J. L., Lagrois, M.-É., & Armony, J. L. (2018). Neural overlap in processing music and speech. In H. Honing (Ed.), *The origins of musicality* (Vol. *370*, pp. 205–220). The MIT Press. https://doi.org/http://dx.doi.org/10.1098/rstb.2014.0090.

Pinker, S. (1997). *How the mind works*. Norton.

Rentfrow, P. J., & Levitin, D. J. (Eds.). (2019). *Foundations in music psychology: Theory and research*. The MIT Press.

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*, *112*(29), 8987–8992. https://doi.org/10.1073/pnas.1414495112.

Trehub, S. E., Weiss, M. W., & Cirelli, L. K. (2019). Musicality across the lifespan. In P. J. Rentfrow & D. J. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 265–303). The MIT Press.

What's not music, but feels like music to you?

Vijay Iyer

Department of Music & Department of African and African American Studies,
Harvard University, Cambridge, MA 02138, USA.
vijayiyer@fas.harvard.edu; www.vijay-iyer.com

doi:10.1017/S0140525X20001740, e79

Abstract

The category "music" as used in this area of science is inconsistent and unstable, and its logical relationship to the word "musicality" – used by scientists to denote the human capacity for music – is circular. Therefore, rather than pursue the question, "Why did music(ality) evolve?" let us ask more inclusively, "What experiences in humankind's deep past might have *felt like music*?"

I am a music-maker who likes to explore the limits of what I think music is (Iyer, 2017, 2021; Iyer & Smith, 2016; Iyer & Taborn, 2019). I am also a music scholar, and I have offered critical perspectives on the music cognition since the 1990s (Iyer, 1998, 2002, 2004, 2016, 2020a, 2020b). I have long argued for an embodied understanding of music as, first and foremost, human action: the sounds of bodies in motion. As composer-pianist Cecil Taylor once put it, music is "everything that you do" (Mann, 1981).

I'm not here to assess the relative scientific merits of these papers. I am not an evolutionary scientist, nor am I invested in the research question of why music might have evolved. Instead, I wish to scrutinize some assumptions underlying both articles. I submit that (1) the category "music" as used in this area of science is inconsistent and unstable, (2) its logical relationship to the word "musicality" – used by scientists to denote the "human capacity for music" – is circular (Iyer & Born, 2020), and (3) this circularity conceals a problem at the heart of this scientific enterprise, namely that its very object of study does not strictly *exist*, but rather is constructed and managed through scientific discourse.

An influential study by Mehr et al. (2019), cited in both papers, drew from a massive corpus of "human song": field recordings described with tags such as "song," "voice," "vocal," and so forth. Although its purview is immense, it expressly does not include purely instrumental music. Yet, the domain of this study has been routinely conflated with the larger category of "music," as is done in Table 1 of the current study by Savage et al. This category called music, however, apparently does not include poetic or rhythmic speech: in their current study, for example, Mehr et al. assert that a coxswain (whose rhythmic chanting coordinates rowers) "does not sing" and should, therefore, be taken as an example of "an alternative to music," namely, language. Meanwhile, Savage et al. posit that we should include dance as "a core part of music-making."

The pattern here is that common human behaviors are freely added to or subtracted from the category of music by scientists as a matter of course. These unpredictable inclusions and exclusions are presented as if normal, even commonsensical. Such arbitrary assertions cause a humanist's ears to perk up, for they are a telltale sign of what Michel Foucault (1972) called a *discursive formation*: an accumulation of statements insistently held together, less by facts than by language, desire, and power. By extension, if *musicality* is defined as the species-wide capacity for making music, and *music* is then defined as (an arbitrarily delimited subset of) the diverse range of outcomes of that capacity known as musicality, then these two definitions orbit, reinforce, and validate each other, but do not resolve onto a definable, consistent, or stable object of study.

We can historicize our contemporary use of the word "music," analyzing it as a contingent, historically specific concept. Our understanding of music today is bound up with our technologically mediated, market-dependent listening experiences, which condition us to think of music as *an object or substance* that is nameable and consumable (see, e.g., Sterne, 2003). Our conception of music additionally depends on western epistemologies of aurality (Erlmann, 2010; Ochoa Gautier, 2014), instrumentality (Dolan, 2013), and humanity/animality (Jackson, 2020; Mundy, 2018; Wynter, 2003) that emerged in the last half-millennium of European imperialism, "exploration," "enlightenment," enslavement, genocide, and extraction. As European "explorers" encountered their racialized others, they believed indigenous people to be at the edge of the human category, and took their musical behaviors as noise; they, like today's scientists, discursively delimited what human music must and must not be (Mundy, 2018; Ochoa Gautier, 2014).

Instead, we must listen beyond our last century of sound mediation, beyond the last five centuries of racial capitalism that undergird western science and reason (Robinson, 1983/2000), and reexamine what music might have been on the timescale of 100,000 years – the era of human sociality, as evidenced in the remains of Blombos Cave (Henshilwood et al., 2011; Wynter & McKittrick, 2015).

Recently, I posed a question on Twitter (Iyer, 2020b): "What's not music, but feels like music to you?" This post received an avalanche of responses: "Cooking." "A really good conversation." "Longing." "Dancing, weather, grandparents." "People being insulted in Yoruba." "Construction sites." "Water droplets, door creaks, astrological planet alignments, white noise, difference tones." "My youngest son's gentle snore, my older son's footsteps, my daughter's laughter." "The wind through desiccating leaves." "Dogs barking." "The unfolding of an exquisite meal." "Shafts of sunlight through slow-moving clouds." "Astrology, Astrophysics, Poetry, the crunch of leaves on a crisp Chicago autumn day, looking deeply in the eyes of a lover, traffic." "The groove a washing machine or dishwasher gets into." "The ocean." "Love." "Memories." "Rollercoasters." "Touch." "Thought." "Everything." Hundreds of people from around the planet offered examples of what "feels like music," illuminating the contested edges of the category, tracing out a larger space of *musical mattering*: affectively charged attending to phenomenological experience. These included human and non-human actions, complex sensations, emotions, thoughts, and meaningful social relations among people. The category of music encompasses many behaviors, but it is also surrounded by a vast sea of experiences that "feel like" they belong in the category too.

Therefore, rather than pursue the question, "Why did music (-ality) evolve?" let us ask more inclusively, "What experiences in humankind's deep past might have *felt like music*, and how, and why?" Feelings are not scientifically trivial; they are ongoing bodily activation at the intersection of biology and culture. As the tweets revealed, virtually any experience can "feel like music" to *somebody*; and they cannot be proven wrong, for *music is felt into being*. We must, therefore, treat the category of music(ality) not as one that coheres as the direct effect of a specific cause (be it social bonding, credible signaling, or anything else), but instead as a sphere of experience, opening out endlessly to human possibility.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Dolan, E. (2013). *The orchestral revolution: Haydn and the technologies of timbre*. Cambridge University Press.

Erlmann, V. (2010). *Reason and resonance: A history of modern aurality*. Zone Books.

Foucault, M. (1972). *The archaeology of knowledge and The discourse on language*. Translated by A.M.S. Smith. Pantheon Books.

Henshilwood, C., d'Errico, F., van Niekerk, K. L., Coquinot, Y., Jacobs, Z., Lauritzen, S.-E., Menu, M., & García-Moreno, R., (2011). A 100,000-year-old ochre-processing workshop at Blombos Cave, South Africa. *Science (New York, N.Y.)*, *334*(6053), 219–222.

Iyer, V. (1998). *Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and African-American Musics. (Publication No. 9922889)*. Ph.D. Dissertation, University of California, Berkeley. ProQuest Dissertations & Theses Global.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, *19*(3), 387–414.

Iyer, V. (2004). Improvisation, temporality, and embodied experience. *Journal of Consciousness Studies*, *11*(3–4), 159–173.

Iyer, V. (2016). Improvisation, action understanding, and music cognition with and without bodies. In G. E. Lewis & B. Piekut (Eds.), *The Oxford handbook of critical improvisation studies* (pp. 74–90). Oxford University Press.

Iyer, V. (2017). *Far from over* (sound recording). ECM Records.

Iyer, V. (2020a). Beneath improvisation. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory* (pp. 761–774). Oxford University Press.

Iyer, V. (2020b). *What's not music but feels like music to you* [Tweet]. Twitter. Retrieved from https://twitter.com/vijayiyer/status/1327294519457554432.

Iyer, V. (2021). *Uneasy* (sound recording). ECM Records.

Iyer, V., & Born, G. (2020). Of musicalities and musical experience: Vijay Iyer and Georgina Born in conversation. Wigmore Hall Podcasts. Retrieved from https://wigmore-hall.org.uk/podcasts/of-musicalities-and-musical-experience-vijay-iyer-and-georgina-born-in-conversation.

Iyer, V., & Smith, W. L. (2016). *A cosmic rhythm with each stroke (sound recording)*. ECM Records.

Iyer, V., & Taborn, C. (2019). *The transitory poems (sound recording)*. ECM Records.

Jackson, Z. I. (2020). *Becoming human: Matter and meaning in an antiblack world*. NYU Press.

Mann, R. (1981). *Imagine the sound* (documentary feature film). Janus Films.

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*, *366*(6468), 957–970.

Mundy, R. (2018). *Animal musicalities: Birds, beasts, and evolutionary listening*. Wesleyan University Press.

Ochoa Gautier, A. M. (2014). *Aurality: Listening and knowledge in nineteenth-century Colombia*. Duke University Press.

Robinson, C. (with Kelley, R. D. G.) (2000). *Black marxism*. University of North Carolina Press. (Original work published 1983).

Sterne, J. (2003). *The audible past: Cultural origins of sound reproduction*. Duke University Press.

Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom: Towards the human, after man, its overrepresentation – an argument. *CR: The New Centennial Review*, *3*(3), 257–337. Coloniality's persistence (fall 2003).

Wynter, S., & McKittrick, K. (2015). Unparalleled catastrophe for our species? Or, to give humanness a different future: Conversations. In K. McKittrick (Ed.), *Sylvia Wynter: On being human as praxis* (pp. 9–89). Duke University Press.

Mind the gap: The mediating role of emotion mechanisms in social bonding through musical activities

Patrik N. Juslin

Department of Psychology, Uppsala University, SE-751 42, Uppsala, Sweden.
patrik.juslin@psyk.uu.se;
https://www.psyk.uu.se/forskning/forskargrupper/the-music-psychology-group/

doi:10.1017/S0140525X2000120X, e80

Abstract

I support the music and social bonding (MSB) framework, but submit that the authors' predictions lack discriminative power, and that they do not engage sufficiently with the emotion mechanisms that mediate between musical features and social bonding. I elaborate on how various mechanisms may contribute, in unique ways, to social bonding at various levels to help account for the socio-emotional effects of music.

Table 1. (Juslin) Emotion mechanisms and their possible bonding effects and neural correlates

Mechanism	Simplified description	Possible bonding effects	Neural correlates[a]
Brain stem reflex	Hardwired attention response to subjectively "extreme" values of simple acoustic features (speed, volume, or sensory dissonance), which exceed an evolutionarily evolved cut-off value	Simple features (e.g., loudness) may facilitate bonding by creating joint attention; their use in performances yields shared feelings of power vis-à-vis intimidated outgroup members; startle pranks aid bonding via humor	The inferior colliculus, the reticulospinal tract of the reticular formation, and the intralaminar nuclei of the thalamus
Rhythmic entrainment early	A gradual synchronization of an internal body rhythm (e.g., heart rate) with an external rhythm in the music, which affects feelings through proprioceptive feedback	Facilitates cooperation and intimacy, building trust; enables shared trance-like states; instills group confidence (e.g., military drill); serves as test of social compatibility (e.g., in dancing)	Multiple oscillators in the early auditory areas, the cerebellum, the sensorimotor cortex, the supplementary motor area, and the caudate nucleus
Evaluative conditioning	The regular pairing of a song with other positive or negative stimuli, gradually leading to an emotional association which may be evoked implicitly or explicitly	Alerts musicians/listeners to the wider social context; enables them to evoke shared associations because of a common learning history, thus enhancing their cultural identity and social belonging	The lateral nucleus of the amygdala, the interpositus nucleus of the cerebellum (hippocampal activation may occur if the context is crucial)
Contagion	Immediate emotional convergence via internal "mimicry" of voice-like emotional expressions in the music (e.g., happiness, tenderness, pride, sadness, and joy)	Harmonizes the mood during social ceremonies; evokes empathy, liking, rapport, and sympathy; enhances the mother–infant attunement; mitigates feelings of loneliness via voice sounds	Right-lateralized inferior frontal areas (e.g., the frontal gyrus), the basal ganglia, "mirror neurons" in the pre-motor regions (especially, those related to vocal expression)
Visual imagery	Mental images conjured up by the listener (e.g., a romantic meeting) via metaphorical mappings of the musical structure, which serve as "internal triggers" of emotions	Enables listeners to visualize and feel close to absent people, and to fantasize about future encounters; serves bonding by first simulating social interactions internally	Various spatially mapped regions of the occipital cortex, the visual association cortex, and (for the process of image generation) the left temporo-occipital regions
Episodic memory	A conscious recollection of a specific event from the listener's past, which is "cued" by the music (e.g., a theme) and re-evokes the feeling experienced during the original event	Enabling listeners to re-visit, re-experience, and reflect upon past social events; facilitates the nostalgic remembrance of loved ones, binding them to one's self	The medial temporal lobe (e.g., the hippocampus), the medial pre-frontal cortex, the precuneus, the entorhinal cortex (applies primarily to memory retrieval)
Musical expectancy	A response to the gradual unfolding of the syntactic structure of the music and its expected and unexpected events (in both musicians and listeners)	Successful predictions of others' musical patterns yield a sense of shared goals and understandings, indicating similar musical history	The left perisylvian cortex, "Broca's area" (BA 44), the dorsal region of the anterior cingulate cortex, and the left lateral orbitofrontal cortex
Esthetic judgment	A subjective evaluation of the music as "art," based on an individual set of weighted criteria (e.g., novelty, skill, beauty, expressivity, and message) which are linearly combined	Esthetic experiences produce self-transcendent emotions like wonder, awe, admiration and gratitude, with pro-social effects on behavior, such as goodwill, helping, and generosity	Depends on the esthetic criterion upon which the judgment is made (e.g., *beauty*: medial orbito-frontal cortex; *novelty*: e.g., the prefrontal and posterior association cortices)

Note: For further explanations of mechanisms and empirical evidence, see Juslin (2019, sects. 3 and 4). For further information about neural correlates, see review in Juslin and Sakka (2019, pp. 301–304). The eight mechanisms are jointly referred to as the BRECVEMA framework.

[a] The predictions focus on networks that distinguish among mechanisms, rather than those that are common (auditory analysis, felt emotion).

I applaud the main aspects of the music and social bonding (MSB) framework proposed by Savage et al. (social bonding as an overarching evolutionary function; gene-culture coevolution; musicality deriving from a set of "proto-musical" components; an inter-disciplinary approach). However, I believe that their case can be bolstered if a few weaknesses are addressed. Thus, I offer some constructive criticism and elaboration here, focusing on two aspects: predictions and mechanisms.

First, I argue that although much of the evidence reviewed in the target article is (more or less) consistent with the MSB hypothesis, it does not provide anything like *unequivocal support* for the notion. As Savage et al. note, there is no shortage of hypotheses about why and how music evolved – the difficulty lies in distinguishing their merits. One function of a framework is thus to offer predictions that are specific, novel, and falsifiable.

However, the predictions provided by Savage et al. are in many cases either too trivial or too vague to distinguish between rival hypotheses. For instance, Savage et al. predict that "social bonding functions should be distributed widely in space and time." What, exactly, counts as "widely"? And what is novel about the prediction? We have abundant evidence that music performs bonding functions across cultures, but does this by itself offer unique support to the MSB hypothesis about the *origin* of music? To illustrate, nearly all musical events across the world involve movement; does this uniquely support the hypothesis that music originally evolved to stimulate movement?

Similarly, Savage et al. predict that bonding via nonmusical methods like language, ritual and sports should be enhanced by the addition of musical components. To the extent that we really need further evidence to show that this is the case, wouldn't positive findings merely be proof of the emotional power of music, as opposed to specifically supporting the MSB hypothesis?

Savage et al. further argue that the MSB hypothesis may be tested by disrupting the "reward systems" that appear to be involved in social bonding. However, if bonding is reduced when such systems are disrupted, it merely confirms that emotion is a key aspect of social bonding – it hardly supports the MSB hypothesis per se. (Moreover, elevated levels of oxytocin found during social bonding may be regarded as the *outcome* of the emotion-induction process that confirms that bonding is taking place, rather than the psychological cause of this effect.)

The lack of precision extends to the overall statement that "the MSB hypothesis proposes that human musicality has been shaped by biological and cultural evolution." I find it hard to think of how it could have been otherwise.

My second line of criticism is that despite the authors' acknowledgment that emotion is a key aspect of social bonding, they do not engage sufficiently with the psychological – as opposed to neurological – mechanisms underlying emotions. They outline direct links between design features of music and neural substrates, but neglect the functional level of analysis needed to link the two. Only a functional account can truly explain *why* and *how* music enhances social bonding, and thus brings people together across barriers of language and culture (Juslin, Barradas, Ovsiannikow, Limmo, & Thompson, 2016).

Moreover, fairly course reference to "reward systems" of the brain fails to do justice to the *multiple* ways such rewards (emotions is perhaps a better term) may be produced. The case can be made that the "mechanistic cycle" shown in Figure 2 of the Savage et al. target article is underspecified or incomplete by suggesting a single route – all bonding effects are not tied to "action-perception circuits." (Elsewhere, the authors themselves mention memory, albeit briefly.)

Perhaps, in their pursuit of a "mechanistic" approach, searching for answers "in engineering terms" (e.g., "what is it for?"), the authors are, unwillingly and temporarily, falling victim to what Nesse (2020) refers to as "tacit creationism": viewing organisms as if they are products of design, without attributing the design to a deity.

Machines serve purposes envisioned by a designer. Bodies are shaped by natural selection. Machines have distinct parts, which serve specific functions. Most parts of the body serve multiple functions, and many functions (e.g., emotions) are distributed among many parts. The resulting complex causal connections frustrate attempts to frame a simple description, such as the "mechanistic cycle" proposed by Savage et al.

I argue that social-bonding effects of music are mediated by multiple and partly redundant emotion mechanisms (Juslin, 2019) that involve additional (partly distinct) brain networks (Juslin & Sakka, 2019), both sensori-motoric mechanisms reflecting concurrent interaction and mechanisms involving detached mental representations, which enable mental travel in time and space. All of the mechanisms may contribute, in their own unique ways, to social bonding – some more than others (see Table 1 for some preliminary examples).

Because these mechanisms can evoke emotions at multiple levels of the brain, social bonding may occur at several levels simultaneously. The iterative process through which the proto-musical components arose initially is mirrored by the gradual evolutionary development of emotion mechanisms, which enabled novel forms of social bonding. Savage et al. focus on participatory musical activities, but as noted by the authors, even solo listening can support social bonding goals.

Perhaps the apparent risk of circularity of the MSB hypothesis (Mehr & Krasnow, 2017) can be addressed by arguing that already existing emotion mechanisms enhanced social bonding, and that the subsequent invention of music built upon, and extended, that foundation in ways that because of its effects on social bonding became subject to natural selection.

Music is social, through and through. And I agree with Savage et al. that the social bonding and coalition signaling hypotheses are not mutually exclusive. Social bonding effects might contribute to coalition formation and signaling, whereas coalition signaling might contribute to social bonding through its contagious effects. However, if the MSB framework truly aspires to be the most comprehensive theory to date of the biological and cultural evolution of music, it needs closer consideration of the psychological mediation between musical design features and social bonding effects.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Juslin, P. N. (2019). *Musical emotions explained: Unlocking the secrets of musical affect*. Oxford University Press.

Juslin, P. N., Barradas, G. T., Ovsiannikow, M., Limmo, J., & Thompson, W. F. (2016). Prevalence of emotions, mechanisms, and motives in music listening: A comparison of individualist and collectivist cultures. *Psychomusicology: Music, Mind, and Brain*, *26*, 293–326.

Juslin, P. N., & Sakka, L. S. (2019). Neural correlates of music and emotion. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 285–332). Oxford 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.

Nesse, R. M. (2020). Tacit creationism in emotion research. *Emotion Researcher*, June 2020, pp. 21–30. Retrieved from: http://emotionresearcher.com/tacit-creationism-in-emotion-research/.

Musicality as a predictive process

Nils Kraus[a,b] and Guido Hesselmann[a]

[a]Department of General and Biological Psychology, Psychologische Hochschule Berlin, 10179 Berlin, Germany and [b]Department of Experimental Psychology and Neuropsychology, Freie Universität Berlin, 14195 Berlin, Germany.
n.kraus@phb.de
g.hesselmann@phb.de

doi:10.1017/S0140525X20000746, e81

Abstract

Savage et al. argue for musicality as having evolved for the overarching purpose of social bonding. By way of contrast, we highlight contemporary predictive processing models of human cognitive functioning in which the production and enjoyment of music follows directly from the principle of prediction error minimization.

In their target article, Savage et al. gather numerous cross-disciplinary findings supporting the hypothesis that musicality evolved for the overarching purpose of social bonding. While acknowledging the large interconnectedness between musicality and events of social bonding, we propose a different conceptualization of the evolutionary mechanisms that led humans to develop a preference for music. Given the high opportunity costs and resources necessary for the evolution and practice of a complex trait-like musicality, one would not only have to argue that music is an effective tool for social bonding, but that its specific features are necessary and most efficient in building cohesive bonds, in the presence of feasible and less resource intensive alternatives. We argue that natural selection favored individuals who tend toward building social bonds in general, and that, therefore, the formation of social groups has revolved around numerous trivial aspects of human life. This becomes apparent in looking at studies of the so-called minimal group paradigm (Tajfel, 1978), in which such arbitrary rules as a coin toss determining group affiliation were sufficient to create in-group favoritism (Tajfel, 1978). Following that line of reasoning, the question of why musicality evolved broadens to the question of why there is esthetic appreciation at all. As Savage et al. point out, reactions in the dopaminergic reward system are elicited by sensory information of every domain allowing for the build-up, validation, and violation of expectations, thereby referring to predictive processing models of human perception and action (Clark, 2013).

Being able to interact with, adapt to and predict future changes in the environment will have constituted the crucial advantage of running and maintaining a perceptual system. For that, a cognitive agent has to acquire an internal model of one's surroundings that is constantly generating new predictions and matching those predictions with current sensory evidence. In order to achieve that in an energy efficient fashion, stimuli that allow for confirmations of prior predictions as well as deviations from them are necessary for optimal learning (Sarasso, Neppi-Modona, Sacco, & Ronga, 2020), and are, therefore, preferred over highly predictable or unpredictable stimuli (Van de Cruys, 2017). We can find this particular relationship between stimulus complexity and likability across multiple sensory domains (Essick et al., 2010; Gold, Pearce, Mas-Herrero, Dagher, & Zatorre, 2019; Van de Cruys & Wagemans, 2011). Perception in this regard can be considered as a form of statistical inference learning, in which intermediate stimulus complexity allows for optimal learning rates and is, therefore, valued (Erle, Reber, & Topolinski, 2017).

Further support for this conceptualization comes from neuroimaging studies. Here, activity in regions such as the orbitofrontal cortex (OFC) has been associated with the attribution of salience to a presented stimulus, and activity in striatal regions such as the caudate, putamen, and nucleus accumbens has been associated with the discrepancy between expected and observed outcomes (reward prediction error). Elevated levels of activity in both those regions can be found during probabilistic reward learning tasks as well as during listening to pleasurable music (Koelsch, 2020; Werlen et al., 2020). Furthermore, temporally more fine-grained measures of neural activity such as the electroencephalogram (EEG) suggest that the process of listening to music is characterized by a constant build-up of musical expectations based on recent harmonic context and integration of currently perceived acoustic inputs into this context (Koelsch, Gunter, Friederici, & Schröger, 2000). During this process, the experienced tension and esthetic pleasure derived from music is not only dependent on the validation, but also the violation of formed expectations (Salimpoor, Zald, Zatorre, Dagher, & McIntosh, 2015). The latter would be implausible if the syntactical structure of music had evolved mainly for the purpose of synchronizing and coordinating musical group performances, as stated by Savage et al. in their target article.

Moreover, this conceptualization of musicality is not limited to one's emotional status being informed by current rates of prediction error in an act of passive listening. It is rather embedded in an action–perception cycle in which not only the generation of, but also the consumption of music is an enactive process. This process is characterized by the constant generation of new predictions, attentional shifts selecting for the currently most informative aspects of music as well as the attempt to reduce prediction error rates via actions such as dancing (Koelsch, Vuust, & Friston, 2019). Furthermore, the emotional status derived from current prediction error rates could in turn modulate subsequent perceptual processing, that is, by how much weight is ascribed to prior predictions and current sensory inputs, respectively (Kraus, Niedeggen, & Hesselmann, 2021).

According to this conceptualization of musicality, deriving pleasure from mere perception of sensory stimuli would not be a "useless by-product of other capacities" but a direct implication of a perceptual system that has evolved to run an internal model of its environment which needs to be updated constantly (Friston & Friston, 2013). This principle is also mirrored in phenomena such as the Pandora effect (Hsee & Ruan, 2016), where participants are willing to engage in self-harming behavior solely for the purpose of uncertainty reduction. Note that this framework derives not only a plausible explanation for musicality (i.e., the tendency to intentionally produce and perceive specific auditory stimuli), but also for esthetic appreciation in other sensory domains (e.g., visual art or fine dining) from basic evolutionary demands that were placed on cognitive agents. This conceptualization of why humans evolved to create, consume, and value music is still compatible with all evidence put forward linking social bonding to musicality, but it questions the implied causality.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Erle, T. M., Reber, R., & Topolinski, S. (2017). Affect from mere perception: Illusory contour perception feels good. *Emotion (Washington, D.C.)*, *17*(5), 856.

Essick, G. K., McGlone, F., Dancer, C., Fabricant, D., Ragin, Y., Phillips, N., … Guest, S. (2010). Quantitative assessment of pleasant touch. *Neuroscience & Biobehavioral Reviews*, *34*(2), 192–203.

Friston, K. J., & Friston, D. A. (2013). A free energy formulation of music generation and perception: Helmholtz revisited. In R. Bader (Ed.) *Sound-Perception-Performance* (pp. 43–69). Springer.

Gold, B. P., Pearce, M. T., Mas-Herrero, E., Dagher, A., & Zatorre, R. J. (2019). Predictability and uncertainty in the pleasure of music: A reward for learning? *The Journal of Neuroscience*, *39*(47), 9397–9409. https://doi.org/10.1523/JNEUROSCI.0428-19.2019.

Hsee, C. K., & Ruan, B. (2016). The Pandora effect: The power and peril of curiosity. *Psychological Science*, *27*(5), 659–666.

Koelsch, S. (2020). A coordinate-based meta-analysis of music-evoked emotions. *NeuroImage*, *223*, 117350. https://doi.org/10.1016/j.neuroimage.2020.117350.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, *23*(1), 63–77.

Kraus, N., Niedeggen, M., & Hesselmann, G. (2021). Trait anxiety is linked to increased usage of priors in a perceptual decision making task. *Cognition*, *206*, 104474.

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.

Sarasso P., Neppi-Modona M., Sacco K., Ronga I. (2020a). Stopping for knowledge: The sense of beauty in the perception-action cycle. *Neuroscience & Biobehavioral Reviews*, *118*, 723–738. https://doi.org/10.1016/j.neubiorev.2020.09.004.

Tajfel, H. E. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press.
Van de Cruys, S. (2017). *Affective value in the predictive mind*. MIND Group.
Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *I-Perception*, *2*(9), 1035–1062.
Werlen, E., Shin, S.-L., Gastambide, F., Francois, J., Tricklebank, M. D., Marston, H. M., … Walton, M. E. (2020). Amphetamine disrupts haemodynamic correlates of prediction errors in nucleus accumbens and orbitofrontal cortex. *Neuropsychopharmacology*, *45* (5), 793–803. https://doi.org/10.1038/s41386-019-0564-8.

Pluralism provides the best chance for addressing big questions about music

Elizabeth Hellmuth Margulis

Woolworth Center for Musical Studies, Princeton University, Princeton, NJ 08544, USA.
margulis@princeton.edu; www.elizabethmargulis.com

doi:10.1017/S0140525X20000965, e82

Abstract

Studying a complex cultural phenomenon like music requires many kinds of expertise. Savage et al. adopt a pluralistic approach, considering multiple forms of evidence and perspectives from multiple fields. This commentary argues that a similar scholarly ecumenicism should be embraced by more studies of music and other cultural phenomena.

Accounts of music's evolutionary origins have a history of ethnocentrism (Cross, 2003; Ochoa Gautier, 2014; Piilonen, 2019; Zon, 2017). They have tended to either overtly position Western classical music as the endpoint of a teleological process or to more tacitly essentialize it by assuming that Western classical music's key attributes define the broader category of "music" which the evolutionary theory seeks to explain. Systemic pressures within individual disciplines can exacerbate this problem. As an object of scholarship, music is squishy; its borders indistinct; its experience subjective; its forms highly variable. Unsurprisingly, it has long been the subject of humanistic inquiry. Science, of course, needs to quantify, operationalize, and measure. But, this process cannot succeed by pretending the squishiness does not exist; for real insights to emerge, quantification and measurements must build sensitively on the contours of actual musical characteristics and behaviors understood in the complexity of their cultural embedding.

Consider the psychometric tests for measuring musical ability devised by Carl Seashore in the early 1900s. One part involved playing pairs of chords that varied in consonance and asking participants which sounded better – Seashore couldn't imagine that enculturation modulates the relationship between sonority and perceived pleasantness (McDermott, Schultz, Undurraga, & Godoy, 2016). Or consider a recent paper which construed time series irreversibility (measured by examining pitch sequences in MIDI representations of melodies from Western classical music) as evidence for the existence of "musical narratives" that influence music's "pleasantness" (González-Espinoza, Martínez-Mekler, & Lacasa, 2020). Fundamental to the paper's logic is the notion that membership in a corpus of Western classical music constitutes incontrovertible evidence of pleasantness, and the notion that sequences of melodic pitches can serve as a proxy for music's fundamental structure. A computational account is only as good as the data and assumptions that are fed into it; overconfidence in quantitative models at the expense of the underlying premises risks generating more "weapons of math destruction" (O'Neill, 2016). To produce musical inventories and statistical measures that are meaningful and capable of answering interesting questions requires expertise not just in psychometrics and statistics, but also in music information retrieval, music and human culture. When knowledge about a relevant subfield is missing, it's that much harder for a study to yield relevant, actionable results, and that much easier for a study to mislead and obfuscate.

Savage et al. expressly "combine cross-disciplinary evidence from archeology, anthropology, biology, musicology, psychology, and neuroscience" (abstract), a "pluralistic approach" that involves "experts from diverse disciplines to synthesize evidence into a single framework" (sect. 1, para. 7). This integration of expertise from different discipline positions allows them to resist the misleading, cartoon-cutout picture of music that can plague scientific inquiry around the subject. In many languages, music and dance are referenced by the same word and conceptualized as a single phenomenon (see Lewis, 2013). Savage et al.'s account acknowledges that dance "is a core part of music-making…and not a separate domain," building their account of music's evolutionary origins around components shared between music and dance, such as sound-action coupling, repetition, prediction, and synchronization. Conversely, accounts of music's origins that center pitch-based syntax and grammar-like structures, or the passive enjoyment of presentational music performed by professional musicians at the expense of participatory accounts (Turino, 2008) risk failing to explain the aspects of music that are most common in music around the world, and thus likelier to be fundamental to the category. Even worse, they risk positioning "Western musical culture … at the top of a racist chain of excellence" (Piilonen, 2019).

When a paper tackles the evolutionary origins of music from the vantage point of a narrow disciplinary perspective, and without consideration for the troubled past of this line of research, it makes scientific progress all but impossible. It would be like trying to understand the evolution of language without domain knowledge in linguistics, or with the assumption that English constituted the sole language from which evidence should be drawn. Music and other products of human culture are no different. Intuitive notions of what music is or how it works are rarely a sufficient basis for the scientific study of it, especially given that decades of deep and nuanced study of the topic already exist in multiple disciplines. Science that does not consider the existing literature on their topic of interest risks making foundational errors that a colleague from across the university or an article on a nearby library shelf could have easily rectified.

Because Savage et al. embrace this idea and involve the scholars and literature of multiple disciplines, they end up with conclusions that are resilient, and do not easily break down when considered in light of other evidence. For example, even though they do not consider Gabrielsson's (2011) study documenting people's peak musical experiences, Savage et al.'s claim about the importance of shared phenomenology, joint intentionality, and self-other merging is supported by the frequency with which descriptions of out-of-body sensations and imagined communion figure in the survey respondents' accounts of their most rewarding musical episodes.

Although the question of music's origins is not one that animates my own research, the fraught character of this line of study makes it a domain in which differences in fundamental approaches are especially apparent. Savage et al.'s paper demonstrates that incorporating expertise from multiple fields, although not sufficient, is necessary for any science that attempts to tackle big questions about cultural phenomena. The bar for quality science in this area should not be pretending away complexity in order to produce a paper that appears solid on the surface, but should rather involve acknowledging that complexity and developing new tools to understand it (Jacoby et al., 2020). The risk of continuing to do science about human culture with blinders on is not just producing faulty conclusions, but – especially given that this study is often covered by the popular press – actually risks doing harm in the real world, as in the case of Carl Seashore's tests of musicality and their relationship to eugenics (Devaney, 2019). Combining perspectives and knowledge from multiple fields is critical to answering questions about human culture.

Financial support. This study is funded by National Science Foundation, Award Number 1734025.

Conflict of interest. None.

References

Cross, I. (2003). Music and biocultural evolution. In M. Clayton, T. Herbert & R. Middleton (Eds.), *The cultural study of music: A critical introduction* (pp. 17–27). Routledge.

Devaney, J. (2019). Eugenics and musical talent: Exploring Carl Seashore's work on talent testing and performance. *American Music Review*, *48*(2), 1–6.

Gabrielsson, A. (2011). *Strong experiences with music: Music is much more than just music* (R. Bradbury, Trans.). Oxford University Press.

González-Espinoza, A., Martínez-Mekler, G., & Lacasa, L. (2020). Arrow of time across five centuries of classical music. *Physical Review Research*, *2*(3), 033166. https://doi.org/10.1103/PhysRevResearch.2.033166.

Jacoby, N., Margulis, E. H., Clayton, M., Hannon, E., Honing, H., Iversen, J., … Wald-Fuhrmann, M. (2020). Cross-cultural work in music cognition: Challenges, insights and recommendations. *Music Perception*, *37*, 185–195.

Lewis, J. (2013). A cross-cultural perspective on the significance of music and dance to culture and society. In M. A. Arbib (Ed.), *Language, music, and the brain: A mysterious relationship* (pp. 45–66). MIT Press. https://doi.org/10.7551/mitpress/9780262018104.003.0002.

McDermott, J., Schultz, A., Undurraga, E., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature 535*, 547–550. https://doi.org/10.1038/nature18635.

Ochoa Gautier, A. M. (2014). *Aurality: Listening and knowledge in nineteenth-century Colombia*. Duke University Press.

O'Neill, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin.

Piilonen, M. (2019). *Resonating subjects: Music and emotion in Victorian evolutionary thought. (Publication No. 22587186).* [Doctoral dissertation, Northwestern University]. ProQuest Dissertations & Theses Global.

Turino, T. (2008). *Music as social life: The politics of participation*. University of Chicago.

Zon, B. (2017). *Evolution and Victorian musical culture*. Cambridge University Press.

Music, bonding, and human evolution: A critique

Bjorn Merker

Independent Scholar, SE-29194 Kristianstad, Sweden.
bjornmerker@gmail.com

doi:10.1017/S0140525X20001429, e83

Abstract

Savage et al. propose that music filled a hypothetical "bonding gap" in human sociality by Baldwinian gene-culture coevolution (or protracted cognitive niche construction). Both these stepping stones to an evolutionary account of the function and origin of music are problematic. They are scrutinized in this commentary, and an alternative is proposed.

Savage et al. follow Dunbar and colleagues in tracing the selection pressure behind human music to a "bonding gap" in human sociality (Dunbar 2012a, 2012b; Launay, Tarr, & Dunbar, 2016; Tarr, Launay, & Dunbar, 2014). Following Podlipniak (2017) and Patel (2018), they further propose that Baldwinian gene-culture coevolution, or protracted cognitive niche construction, supplied the mode of selection for its evolution.

The relationship between grooming time and group size on which the bonding gap is premised is not linear, as Dunbar originally assumed, but starts leveling off for group sizes exceeding a few dozen individuals (Lehmann, Korstjens, & Dunbar, 2007, Fig. 1). This allows some nonhuman primates to sustain groups far larger than the roughly 150 individuals for which a hypothetical human bonding deficit has been invoked (Dunbar, 2012b; Lehmann et al., 2007, Fig. 3b; Snyder-Mackler, Beehner, & Bergman, 2012). The bonding gap construct itself is problematic, in other words.

A principal predictor of group size in primates is the ratio of neocortex to the rest of the brain (Dunbar, 1992, 1998). Neocortex capacity is not needed for grooming, but is essential for tracking social contingencies over time, for the acquisition of familiarity with individuals, and for budgeting the investment of time and effort in fruitful relationships. These are the principal determinants and the substance of personal bonds, whereas their expression (secondarily reinforcing them, in part hormonally, as in grooming) may vary, as the human versus nonhuman primate case shows (Dunbar, 1993). For large-scale human sociality, moreover, personal bonds are joined by major factors such as human language as well as political and religious authority with power to administer, reward and punish (Turchin et al., 2018).

Mode of selection: Savage et al. explicitly treat music as a componential phenomenon based on a set of diverse behavioral capacities, with complex, protracted genesis. Not one of these, we are told, evolved by ordinary natural or sexual selection or the shaping agency of iterated learning by pure cultural transmission. Yet, at least one critical component of musicality cannot evolve by Baldwinian or niche construction mechanisms. Every song we sing became ours through a specialized cerebral mechanism for learning to reproduce vocally, using feedback from one's own voice, novel sound patterns received by ear. We alone among primates have this capacity for vocal production learning (Janik & Slater, 1997), not to be confused (as some primatologists do) with learned modification of innate calls (Janik & Slater, 2000). Without the capacity for vocal learning it is impossible to produce the behavior it enables, leaving Baldwinian or niche construction modes of selection without a starting point, by "invention," happenstance, or otherwise. The stakes for explaining the origin of this human autapomorphy are high, indeed (Nottebohm, 1976).

Whence the esthetic extravagance of music? In behavioral biology terms, music belongs to the class of esthetic displays that forced Darwin to add sexual selection to his theory (Darwin, 1871; Fisher, 1930; Zahavi, 1975). Neither a "solo" nor a "male" performance bias is integral to this mode of selection, as Savage et al. suppose. The key, instead, is manipulative signaling driven by divergence of interests between evolutionary players (Enquist, Arak, Ghirlanda, & Wachtmeister, 2002), something that cannot be assessed apart from social and mating system particulars. Thus, in monogamy sexual selection operates bi-directionally (Miller, 1998, sects. 3.3 and 4; Wachtmeister, 2001; Wachtmeister & Enquist, 2000). The esthetic extravagance of music, which Savage et al. do not attempt to explain, suggests a role for sexual selection in its genesis (Merker, 2019). But, what of the many group-based formats and functions of music, emphasized by Savage et al.?

A unique communal display: On irregular occasions, typically when a foraging subgroup discovers a ripe fruit tree or when two subgroups of the same territory meet after a period of separation, chimpanzees launch an excited bout of loud calls, running, branch shaking, and slapping of tree buttresses. These so-called carnival displays, which feature no coordination between individuals of a musical nature, may last for hours, even a whole night, and induce others on the territory, both males and females, to join the fray (Ghiglieri, 1984; Reynolds & Reynolds, 1965; Sugiyama, 1969, 1972; Wrangham, 1975).

Note the social significance of the occasions for a carnival display: a lucky subgroup invites the rest of its territory mates to share in its windfall, and "celebration" by subgroups reunited after separation. Everyone participate in the commotion, which we may think of as a communal expression of belonging to a territorial group, *a communal bonding display*. The setting for this unique behavior is male communal territoriality combined with female exogamy, a rare combination of traits which on cladistic grounds is attributed to the common ancestor of humans and chimpanzees (Ember, 1978; Ghiglieri, 1987; Lovejoy, 2009; Wrangham, 1979).

It is likely, therefore, that the common ancestor of humans and chimpanzees on occasion engaged in something like the communal bonding display of present-day chimpanzees. A common ancestor population in which display participants – unlike present-day chimpanzees – managed to synchronize their voices to a common, isochronous pulse (most readily provided by in-place locomotor activity), would achieve superposition of their voice amplitudes, resulting in a joint signal with far greater geographical reach than unsynchronized voices. This "supervoice" would carry far across neighboring territories to reach exogamously migrating females, attracting them to the synchronizers *at the reproductive expense of non-synchronizing groups*.

The just described "inter-communal sexual selection" pressure for synchrony would, once migrating females were "on territory," be joined by an "intra-communal sexual selection" pressure for elaboration of the structural content of the display in the direction of added complexity and sophistication, in both pitch and rhythm domains, in keeping with the esthetic logic of sexual selection. The result: an ancestral group display featuring chanting (eventually singing) in synchrony to rhythmic movement "in place," amounting to a form of dancing. Thus arose, I suggest, a major cross-cultural universal of music: our tendency to gather from time to time to sing and dance together in a group, that is, music was born (Merker, 1999, 2000; Merker, Madison, & Eckerdal, 2009; Merker, Morley, & Zuidema, 2018). Crucially, the bonding function and its motivational underpinnings were *already present in the ancestral communal display prior to its elaboration into music by inter- and intra-communal sexual selection.*

This account of the origin and function of music is compatible with the broad outlines Mehr et al. proposed in their target article. The origin of human vocal production learning in the ancestral setting just sketched is the topic of a forthcoming publication of mine.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Darwin, C. (1871). *The descent of man and selection in relation to sex*. D. Appleton & Co.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *20*, 469–493.

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. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *6*, 178–190.

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*, 1837–1846.

Ember, C. R. (1978). Myths about hunter-gatherers. *Ethnology*, *17*, 439–448. doi: 10.2307/3773193.

Enquist, M., Arak, A., Ghirlanda, S., & Wachtmeister, C.-A. (2002). Spectacular phenomena and limits to rationality in genetic and cultural evolution. *Transactions of the Royal Society of London B*, *357*, 1585–1594.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Clarendon Press.

Ghiglieri, M. P. (1984). *The chimpanzees of Kibale forest*. Columbia University Press.

Ghiglieri, M. P. (1987). Sociobiology of the great apes and the hominid ancestor. *Journal of Human Evolution*, *16*, 319–357. doi: 10.1016/0047-2484(87)90065-0.

Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, *26*, 59–99.

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*, 1–11.

Launay, J., Tarr, B., & Dunbar, R. I. M. (2016). Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology*, *122*, 779–789.

Lehmann, J., Korstjens, A. H., & Dunbar, R. I. M. (2007). Group size, grooming and social cohesion in primates. *Animal Behavior*, *74*, 1617–1629. doi: 10.1016/j.anbehav.2006.10.025.

Lovejoy, C. O. (2009). Reexamining human origins in light of *Ardipithecus ramidus*. *Science*, *326*(5949), 74e1–74e8. doi: 10.1126/science.1175834.

Merker, B. (1999). Synchronous chorusing and the origins of music. Musicae Scientiae, Special Issue, Rhythm, Musical narrative, and Origins of Human Communication, pp. 59–74. [these special issues have no volume number!!!].

Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 315–328). The M.I.T. Press.

Merker, B. (2019). When extravagance impresses: Recasting esthetics in evolutionary terms. In M. Thaut & D. Hodges (Eds.), *The Oxford handbook of music and the Brain* (pp. 66–86). Oxford University Press.

Merker, B., Madison, G., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*, 4–17.

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. 49–80). The MIT Press.

Miller, G. F. (1998). How mate choice shaped human nature: A review of sexual selection and human evolution. In C. Crawford & D. Krebs (Eds.), *Handbook of evolutionary psychology: Ideas, issues, and applications* (pp. 87–130). Lawrence Erlbaum.

Nottebohm, F. (1976). Discussion paper. Vocal tract and brain: A search for evolutionary bottlenecks. In S. R. Harnad, H. D. Steklis & J. Lancaster (Eds.), Origins and evolution of language and speech. *Annals of the New York Academy of Sciences 280* (pp. 643–649). New York Academy of Sciences.

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). The MIT Press.

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.

Reynolds, V., & Reynolds, R. (1965). Chimpanzees of the Budongo forest. In I. Devore (Ed.), *Primate behavior: Field studies of monkeys and apes* (pp. 368–424). Holt, Rinehart & Winston.

Snyder-Mackler, N., Beehner, J. C., & Bergman, T. J. (2012). Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *International Journal of Primatology 33*, 1054–1068. https://doi.org/10.1007/s10764-012-9584-5.

Sugiyama, Y. (1969). Social behavior of chimpanzees in the Budongo Forest, Uganda. *Primates*, *9*, 225–258. doi:10.1007/bf01730343.

Sugiyama, Y. (1972). Social characteristics and socialization of wild chimpanzees. In F. E. Poirer (Ed.), *Primate socialization* (pp. 145–163). Random House.

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, 1–10. https://doi.org/10.3389/fpsyg.2014.01096.

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 115*, E144–E151 https://doi.org/10.1073/pnas.1708800115.

Wachtmeister, C.-A. (2001). Display in monogamous pairs: A review of empirical data and evolutionary explanations. *Animal Behaviour*, *61*, 861–868.

Wachtmeister, C.-A., & Enquist, M. (2000). The evolution of courtship rituals in monogamous species. *Behavioral Ecology*, *11*, 405–410.

Wrangham, R. W. (1975). *The behavioural ecology of chimpanzees in Gombe National Park, Tanzania*. Unpublished doctoral dissertation, University of Cambridge, Cambridge, UK.

Wrangham, R. W. (1979). On the evolution of ape social systems. *Social Science International*, *18*, 335–368.

Zahavi, A. (1975). Mate selection – a selection for a handicap. *Journal of Theoretical Biology*, *53*, 205–214.

Bonds and signals underlie the music learning experience

Steven J. Morrison

Center for the Study of Education and the Musical Experience, Henry & Leigh Bienen School of Music, Northwestern University, Evanston, IL 60208, USA.
steven.morrison@northwestern.edu

doi:10.1017/S0140525X20001351, e84

Abstract

The music learning environment is a context in which fundamental forces and values underlying human musicality may be evident. Social bonding within music-making groups is characterized by a high degree of complexity whereas issues of clarity, accuracy, and coordination remain the focus of learning. Physical and cognitive impairments that compromise music learning opportunities offer a critical test of music's link to social bonding.

Although the definitive primary impetus for human musicality remains elusive, both Savage et al. and Mehr et al. speculate that the shadow of that impetus may be discernible within present day music thinking and behaving, albeit in two different yet complementary forms: social bonding and credible signaling. The difficulty in isolating the two theories from each other reflects the deeply entangled relationship among the multiple facets of the music learning experience. It is fruitful to consider the manner in which humans in the here-and-now become musical, actively seek opportunities for musical growth and engagement, and facilitate the musical growth of others as echoes of music's foundational place in the human experience. Taking perspectives from music learning and teaching, the evidence for music as a facilitator for social bonding is compelling among both children and adults who chose to make the ongoing study of music, either formally or informally, a part of their lives. At the same time, the emphasis placed on the clarity, precision, and, in turn, power of musical expressions that lie at the heart of many music education models (Abril & Gault, 2008) is consistent with the demands of what Mehr et al. describe as a credible signal.

Savage et al. argue that musicality (as a "cognitive toolkit") facilitates social bonding among larger groups more effectively and efficiently than alternative methods (e.g., grooming), yet this would arguably have first demanded a demarcation of what music was and was not. Conversely, Mehr et al.'s proposition that music evolved from its value as a credible signal presumes the presence of a signalee and immediately introduces a social component to music's beginning, one that functioned in both out-of-group (territoriality) and in-group (parent/infant dyads) interactions.

Social aspects of music making are central to the music experiences of young people (Ilari, 2016) and adult learners (Dabback, 2008), as well as music specialists (Kokotsaki & Hallam, 2007). As a context in which social engagement occurs, it is notable that music facilitates interpersonal dynamics that reach beyond solely bonding, both in complexity and valence. On this point, Savage et al. posit that musicality "increased the quality (depth and complexity) of existing relationships" (sect. 2.4). Musical play establishes social bonds but also facilitates within-group hierarchies and competition, much to the preference of the children taking part (Roberts, 2016). Groups that interact toward a common musical cause can demonstrate a high degree of social complexity. This is apparent in the intricate array of relationships observed within a music ensemble performance (D'Ausilio et al., 2012; Dineen, 2011) and among large musical organizations (Dabback, 2008; Weren, 2015). It may even extend to broad music-adjacent contexts such as that famously described by Small (1998) in his examination of the symphony orchestra concert (although Small considered all related activities as musical). The value placed on the social cohesion facilitated by music is demonstrated by its absence. Beyond simply reducing the feeling of social closeness, a breakdown in musical synchrony can lead to stress among group members and ill feeling toward the musical leader (Lorenz, 2020). This is in stark contrast to the interpersonal synchronization observed at the neural level between teacher and student in a song-learning setting (Pan, Novembre, Song, Li, & Hu, 2018).

To emphasize the compatibility of music with social bonding, Savage et al. point out that music is characterized by its reproducibility and by its accommodation for successful participation by multiple individuals. Thus, a premium is placed on musical behaviors featuring a high level of predictability (e.g., the use of regular metric patterns) and repetition (such as the employment of a small number of recurring melodic and/or rhythmic figures). Music is arguably built to be a group endeavor. This is congruent with Mehr et al. who argued that there was an evolutionary pressure placed on the clarity of a credible signal, with a coordinated signal being more discernible and stronger (with an emphasis on both synchrony and coordination, see Lee, Launay, & Stewart, 2020). Mehr et al. proposed that "selection should push receivers to better discriminate differences in degrees of coordination, and signalers to produce more complex coordinate signals" (sect. 4.2.1, para. 4). On a much smaller temporal scale, this is an apt description of a music curriculum. Within the active music learning context, such an emphasis on clarity of signal is evident to the point of being a fundamental truth. Achievement in music learning is demonstrated by a person's ability to perform "correctly," in coordination with others, and in adherence to prevailing norms.

Although Mehr et al. states that “music does not *directly* cause social cohesion” (sect. 4.2.1, para. 14), Savage et al. predict the two demonstrate correlational (bonding increases with number of musical components) and additive (bonding greater for participatory than non-participatory experiences) properties. If musicality is inextricably linked to social bonding, a critical question might be posed whether social bonding is compromised among those for whom the learning of coordinated group music activity is difficult. Savage et al. make the case that observation of musical behaviors is sufficient to facilitate social bonding among groups too large to include only performers. Peretz (2016) speculated that the reward system response to group music making may serve as an effective intervention for amusics, suggesting that context rather than musical coordination is more salient. Whether as a participant or as an observer, it may be tested whether degraded or compromised capabilities, physiological or cognitive, related to metric, rhythmic, or melodic perception result in a corresponding decrease in bonding.

Be it Cross’s “floating intentionality” (2008) or Keil’s “participatory discrepancies” (1987), the shared experience of music in which a multiplicity of meanings and interpretations is accommodated might be considered antagonistic to the clarity or interpretability of musical signaling. Alternatively, the proposition of music as fundamentally a means of social bonding might manifest in the present through its ability to establish and support implicit relationships that, although in cases hierarchical, do not compromise the contribution of any given individual to the collective. The earliest development of behaviors and processes supporting music may have appeared through the need to establish credible signals, but perhaps the ultimate utility of these behaviors and processes to establish and sustain social relationships lies at the heart of their distinction as music.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Abril, C. R., & Gault, B. M. (2008). The state of music in secondary schools: The principal’s perspective. *Journal of Research in Music Education*, *56*(1), 68–81. https://doi.org/10.1177/0022429408317516.

Cross, I. (2008). Musicality and the human capacity for culture. *Musicae Scientiae*, *12* (1_Suppl.), 147–167. https://doi.org/10.1177/1029864908012001071.

D’Ausilio, A., Badino, L., Tokay, S., Craighero, L., Canto, R., Aloimonos, Y., & Fadiga, L. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS ONE*, *7*(5), e35757. https://doi.org/10.1371/journal.pone.0035757.

Dabback, W. M. (2008). Identity formation through participation in the Rochester New Horizons Band programme. *International Journal of Community Music*, *1*(2), 267–286. https://doi.org/10.1386/ijcm.1.2.267_1.

Dineen, P. M. (2011). Gestural economies in conducting. In A. Gritten & E. King (Eds.), *New perspectives on music and gesture* (pp. 131–157). Ashgate Publishing Ltd.

Ilari, B. (2016). Music in the early years: Pathways into the social world. *Research Studies in Music Education*, *38*(1), 23–39. https://doi.org/10.1177/1321103X16642631.

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, *2*(3), 275–283. https://www.jstor.org/stable/656427.

Kokotsaki, D., & Hallam, S. (2007). Higher education music students’ perceptions of the benefits of participative music making. *Music Education Research*, *9*(1), 93–109. https://doi.org/10.1080/14613800601127577.

Lee, H., Launay, J., & Stewart, L. (2020). Signals through music and dance: Perceived social bonds and formidability on collective movement. *Acta Psychologica*, *208*, 103093. https://doi.org/10.1016/j.actpsy.2020.103093.

Lorenz, T. (2020). *“I play for togetherness”: Impacts of audio-visual asynchrony on feelings of social closeness in adult community wind band musicians* (Publication No. 28022922). Doctoral dissertation, University of Washington, ProQuest Dissertations & Theses Global.

Pan, Y., Novembre, G., Song, B., Li, X., & Hu, Y. (2018). Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *NeuroImage*, *183*, 280–290. http://doi.org/10.1016/j.neuroimage.2018.08.005.

Peretz, I. (2016). Neurobiology of congenital amusia. *Trends in Cognitive Sciences*, *20*(11), 857–867. http://doi.org/10.1016/j.tics.2016.09.002.

Roberts, J. C. (2016). ‘Wanna race?’: Primary student preference for competitive or non-competitive singing games. *British Journal of Music Education*, *33*(2), 159–174. https://doi.org/10.1017/S0265051715000236.

Small, C. (1998). *Musicking: The meanings of performing and listening*. Wesleyan University Press.

Weren, S. (2015). *Motivational and social network dynamics of ensemble music making: A longitudinal investigation of a collegiate marching band* (Publication No. 3701556). Doctoral dissertation, Arizona State University, ProQuest Dissertations & Theses Global.

Where they sing solo: Accounting for cross-cultural variation in collective music-making in theories of music evolution

Aniruddh D. Patel[a,b] and Chris von Rueden[c]

[a]Department of Psychology, Tufts University, Medford, MA 02445, USA;
[b]Program in Brain, Mind, and Consciousness, Canadian Institute for Advanced Research (CIFAR), Toronto, ON M5G 1M1, Canada and [c]Jepson School of Leadership Studies, University of Richmond, Richmond, VA 23173, USA.
a.patel@tufts.edu; https://ase.tufts.edu/psychology/people/patel/
cvonrued@richmond.edu; https://sites.google.com/site/chrisvonrueden/home

doi:10.1017/S0140525X20001089, e85

Abstract

Collective, synchronous music-making is far from ubiquitous across traditional, small-scale societies. We describe societies that lack collective music and offer hypotheses to help explain this cultural variation. Without identifying the factors that explain variation in collective music-making across these societies, theories of music evolution based on social bonding (Savage et al.) or coalition signaling (Mehr et al.) remain incomplete.

Savage et al. and Mehr, Krasnow, Bryant, & Hagen argue that collective music-making (e.g., group singing and dancing) is part of our evolved human nature because of its adaptive function over human evolution, either because of its role in social bonding (Savage et al., target article) or in signaling coalition strength to other groups (Mehr et al., target article).

The social bonding theory, in particular, has old roots (e.g., Dunbar, 2004; Roederer, 1984), yet both theories face an important, unaddressed challenge: Namely, the existence of significant variation in the extent to which traditional, small-scale societies engage in collective music-making. Simply put, the traditional practices of some small-scale societies include little to no collective music-making, whereas in other indigenous cultures collective music-making is prominent. Documenting and understanding this variation is crucial for collectivist theories of music’s origins, because unless this variation can be explained, such theories remain incomplete.

Consider the Tsimane of the Bolivian Amazon, a population of forager-horticulturalists, with whom one of us (CvR) has worked since 2005. Ethnographic research indicates very little collective music-making in Tsimane society. Between 2002 and 2007, structured observations ($n \simeq 80{,}000$) of hundreds of people in different villages included only nine instances of collective music-making – all pairs of young children singing together in the context of play (Tsimane Health and Life History Project, unpublished). Traditionally, Tsimane music-making was largely solo, including singing by shamans or other older adults whose songs conveyed traditional knowledge, reinforced cultural norms, and propitiated ancestors and the guardian spirits of forest animals (Huanca, 2006; Godoy, personal communication). Missionary influence has contributed to the demise of traditional song, and led to establishment of churches in several villages. Where CvR has observed collective singing among adults, it occurred in the context of a church service, and Tsimane participation in collective singing was fairly reluctant, timid, and uncoordinated. Little experience with collective music-making may help explain unusual features of Tsimane music perception, such as no esthetic preferences related to consonance and dissonance, and a lack of octave equivalence (Jacoby et al., 2019; McDermott, Schultz, Undurraga, & Godoy, 2016; Thompson, Sun, & Fritz, 2019; Zatorre, 2016).

The Tsimane are not unique in having little collective musical behavior. Coordinated group music (vocal and instrumental) is largely absent in traditional cultures in parts of Siberia, including among the Tuvans and Yakuts, both of whom engage in animal husbandry (Levin, 2006; Nikolsky, Alekseyev, Alekseev, & Dyakonova, 2020). Our informal canvassing of a few ethnographers who have carried out years of fieldwork with hunter-gatherers revealed a relative absence of collective music-making in the traditional practices of several such groups. These include the northern Aché of Paraguay and the Agta of the Philippines, both of which are egalitarian cultures (Kim Hill, Bion Griffin, and Thomas Headland, personal communication), and the Ayoreo people of Bolivia and Paraguay (Lucas Bessire, personal communication). We strongly suspect that these societies are far from a complete list of those where traditional music is largely performed solo.

Compiling a list of such small-scale, traditional societies and understanding what drives their tendency toward solo music have implications for the frequency and importance of collective music-making over human evolution. We offer three hypotheses to explain cultural variation in collective music-making:

(1) Music-making is less likely to be collective where there is less collective action in general. In particular, a relative absence of inter-group conflict may predict less collective, synchronous singing and dancing, given less demand for signaling group cohesion (Mehr et al., target article) and lower returns to social bonding and identity fusion (Savage et al., target article). The Tsimane may be an example: historical documents and ethnographic accounts suggest a relative lack of inter-group coalitional violence over the past several centuries (Godoy, 2015). Furthermore, collective action in productive activities outside of the extended family is infrequent in Tsimane society (von Rueden, Gurven, Kaplan, & Stieglitz, 2014). It may be that inter-group conflict or exchange fosters the genesis and spread of collective music making via cultural evolutionary dynamics, because collective music-making facilitates the competitive success of groups.

(2) Music-making is less likely to be collective where it is a principal means of conveying expert knowledge or individual accomplishment. In such cases, individuals may be more reluctant to produce music or be discouraged from doing so unless their status (e.g., shaman) justifies their music-making. Among the Ayoreo, who lack synchronous singing, men (and sometimes women) take turns singing at night to describe notable events, but singers are implicitly ranked on skill and "bad" singers meet with critical commentary from others (Bessire, personal communication, 2006, 2014).

(3) The adoption of certain musical styles can constrain subsequent adoption of collective music-making. For example, timbre-based music, common in Tuva and elsewhere in Siberia (Levin & Süzükei, 2018; Nikolsky et al., 2020), is less conducive to collective, synchronous music-making than is pitch-based music. Timbre-based music instead enables highly individualized forms of expression that are used by Tuvans and Yakuts as a form of personal identification (Nikolsky et al., 2020).

In summary, collective music-making is far from ubiquitous across small-scale, traditional societies, and documenting and seeking to explain this variation should be part of any evolutionary theory which appeals to the adaptive value of collective musical behavior. We do not claim that collectivist theories of music evolution are necessarily wrong. It is possible that the musical capacities that enable solo music-making evolved largely in the context of collective music-making. However, it is also possible that collective musicality arose and spread via purely cultural dynamics and did not play a necessary or sufficient role in genetic selection of traits underlying human musicality. This is an important null hypothesis for gene-culture coevolutionary theories of collective musicality (e.g., Savage et al., target article; Patel, 2018).

Acknowledgments. We thank Kim Hill, Bion Griffin, Thomas Headland, Lucas Bessire, Ricardo Godoy, and Tomás Huanca for their valuable input.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Bessire, L. (2006). *From honey to ashes* (film). Documentary Educational Resources.

Bessire, L. (2014). *Behold the Black Caiman: A chronicle of Ayoreo life*. University of Chicago Press.

Dunbar, R. I. M. (2004). Language, music, and laughter in evolutionary perspective. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 257–274). MIT Press.

Godoy, R. (2015). *A natural experiment in westernization: Of Faustian deals and the primitive*. Center for Global Development and Sustainability Working Paper. Brandeis University.

Huanca, T. (2006). *Tsimane' oral tradition, landscape, and identity in tropical forest*. SEPHIS. South-South Exchange Programme for Research on the History of Development.

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.

Levin, T. (2006). *Where rivers and mountains sing*. Indiana University Press.

Levin, T., & Süzükei, V. (2018). Timbre-centered listening in the soundscape of Tuva. In E. I. Dolan & A. Rehding (online) (Eds.), *The Oxford handbook of Timbre*. Oxford University Press. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190637224.001.0001/oxfordhb-9780190637224-e-15.

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*(7613), 547–550.

Nikolsky, A., Alekseyev, E., Alekseev, I., & Dyakonova, V. (2020). The overlooked tradition of "personal music" and its place in the evolution of music. *Frontiers in Psychology*, *10*, 3051.

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.

Roederer, J. G. (1984). The search for a survival value of music. *Music Perception*, *1*, 350–356.

Thompson, W. F., Sun, Y., & Fritz, T. (2019). Music across cultures. In P. J. Rentfrow & D. J. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 503–541). MIT Press.

Tsimane Health and Life History Project (n.d.). Unpublished database. http://tsimane.anth.ucsb.edu/.

von Rueden, C., Gurven, M., Kaplan, H., & Stieglitz, J. (2014). Leadership in an egalitarian society. *Human Nature*, *25*, 538–566.

Zatorre, R. (2016). Human perception: Amazon music. *Nature*, *535*(7613), 496–497.

Music production deficits and social bonding: The case of poor-pitch singing

Peter Q. Pfordresher

Department of Psychology, Center for Cognitive Science, University at Buffalo, State University of New York, Buffalo, NY 14260, USA.
pqp@buffalo.edu;
http://arts-sciences.buffalo.edu/psychology/faculty/faculty-directory/pfordresher.html

doi:10.1017/S0140525X20001247, e86

Abstract

Both of the companion target articles place considerable performance on music performance ability, with specific attention paid to singing in harmony for the music and social bonding (MSB) hypothesis proposed by Savage and colleagues. In this commentary, I evaluate results from recent research on singing accuracy in light of their implications for the MSB hypothesis.

Both target articles in this dual treatment focus on the importance of music performance for engaging in collective behaviors. Although the target papers did not place great emphasis on individual differences in performance ability, these differences are clearly relevant. In particular, the music and social bonding (MSB) hypothesis proposed by Savage and colleagues proposes that collective musical performance promotes social bonding, which in turn may have yielded survival benefits to our species. If music performance promotes social bonding, and social bonding promotes evolutionary fitness, then we may make predictions about music performance ability. By this logic, we may predict that the ability to perform music (in general) emerges spontaneously in development, that this ability is widespread in the population, and that the lack of this ability may be associated with reproductive costs.

Singing is arguably the best form of music performance on which to test these predictions. In contrast to other forms of music performance, singing does not require learning of a human-made instrument and thus may emerge without any specialized training. Furthermore, singing involves the ability to regulate actions along both dimensions cited as important for social bonding in the MSB hypothesis: synchronization and harmonization. I focus here on the latter dimension.

Singing in harmony requires accurate tuning of vocal pitch control, or pitch accuracy. If singing in accurately arose from a process of gene-culture coevolution, one would expect singing accuracy to be widespread and learned early in life without considerable intervention. Singing does emerge spontaneously in children at around the same time as language acquisition appears (Stadler Elmer, 2020). By adulthood, most individuals appear able to match pitch within the boundaries set by their culture's tuning system. Tests of singers matching pitch within equal tempered tuning (which dominates Western music) suggests a majority can match pitch within a musical semitone. For instance, a recent large-scale study of over 1,000 individuals found that roughly 1/3 of participants matched over 90% of pitches within these boundaries on average (Pfordresher & Demorest, 2021). By contrast, relatively few participants exhibit a dominant tendency to deviate from a target pitch by a half semitone or more (the dividing line between pitch classes in the equal tempered scale). In the aforementioned sample, only 25% of participants exhibited this tendency; a figure that accords well with other reports (Dalla Bella, Giguère, & Peretz, 2007; Hutchins & Peretz, 2012; Pfordresher & Brown, 2007; Pfordresher, Brown, Meier, Belyk, & Liotti, 2010). Existing data, therefore, suggest most humans should be able to harmonize within acceptable limits as per the claims of MSB hypothesis.

Given the fact that accurate singing is widely represented in the population, thus facilitating harmonization, one might predict that collective singing is widespread. Unfortunately, self-evaluation of singing in Western cultures tends to run against the levels of accuracy found in the data, and this may suppress participation in collective singing. Pfordresher and Brown (2007) found that in a sample of over 1,000 psychology undergraduates in Texas, 59% believed they were unable to imitate a melody via singing. In the more recent large data set discussed above (Pfordresher & Demorest, 2021), only 34% agreed with the claim "I am a good singer" whereas 62% agreed with the claim "I enjoy singing." Although self-assessments did correlate with singing accuracy in this data set, there is clearly a disconnect between overall levels of self-evaluation and levels of pitch accuracy. These dissociations can have significant consequences given that self-assessment of musical skill predicts future musical participation during late childhood (Demorest, Kelley, & Pfordresher, 2016). Reported experiences of inaccurate singers of their early choral singing experiences tell of significant embarrassment and avoidance of future singing (Welch, 2006). Thus, if collective singing is beneficial to social bonding and to our species, this strong tendency in Western culture may have considerable negative consequences.

This leads to a prediction of the MSB hypothesis that is less clearly supported: What are the consequences of music production deficits, including inaccurate singing? The fact that a trait may have had adaptive value in the past does not necessarily mean that the success of modern humans relies on this trait; nevertheless, the importance of traits in modern life is often held up as evidence for such claims (as is the case for language, for instance). Pinker's (1997) speculations about the non-adaptive nature of music, for instance, arose in part from the lack of evidence that music plays a fundamental role in modern-day survival. In the case of singing, the evidence is mixed. To date, there has been no evidence that inaccurate singers, or individuals with other deficits such as congenital amusia, fare worse in life than others. On the other hand, participation in singing may promote well-being at a physiological level, via reduction of cortisol secretion (Grebosz-Haring & Thun-Hohenstein, 2018; Kreutz, Bongard, Rohrmann, Hodapp, & Grebe, 2004), the release of oxytocin (Keeler et al., 2015), and through the promotion of social relationships.

Evidence from research on singing accuracy among modern Westerners thus accords with the MSB hypothesis in some ways but not others. More important, the MSB hypothesis helps to place research on singing accuracy in an informative broader context. Ultimately, although pitch accuracy is an important component of musical communication, it may be the case that the evolution of music was not so much based on the ability to perform accurately but may instead be rooted in the benefits of engaging in this collective activity. The fact that feedback from multiple singers may lead to a collective chorusing effect that obscures accuracy of an individual singer supports a limited role for accuracy of an individual. Moreover, fine-grained tuning of singing is negligibly related to the esthetic quality of singing (Hutchins, Roquet, & Peretz, 2012; Pfordresher & Brown, 2017). If music evolved to be participatory, those of us in Western nations may wish to reconsider the current value placed on virtuoso solo performances.

Financial support. This research was funded in part by NSF grant BCS-1848930.

Conflict of interest. I am presently collaborating with Patrick Savage as part of a larger research team.

References

Dalla Bella, S., Giguère, J. F., & Peretz, I. (2007). Singing proficiency in the general population. *Journal of the Acoustical Society of America*, *121*, 1182–1189.

Demorest, S. M., Kelley, J., & Pfordresher, P. Q. (2016). Singing ability, musical self-concept and future music participation. *Journal of Research in Music Education*, *64*, 405–420.

Grebosz-Haring, K., & Thun-Hohenstein, L. (2018). Effects of group singing versus group music listening on hospitalized children and adolescents with mental disorders: A pilot study. *Heliyon*, *4*, e01014.

Hutchins, S., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General*, *141*, 76–97.

Hutchins, S., Roquet, C., & Peretz, I. (2012). The vocal generosity effect: How bad can your singing be? *Music Perception*, *30*, 147–159.

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*, 518.

Kreutz, G., Bongard, S., Rohrmann, S., Hodapp, V., & Grebe, D. (2004). Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state. *Journal of Behavioral Medicine*, *27*, 623–635.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness." *Music Perception*, *25*, 95–115.

Pfordresher, P. Q., & Brown, S. (2017). Vocal mistuning reveals the origin of musical scales. *Journal of Cognitive Psychology*, *29*, 35–52.

Pfordresher, P. Q., Brown, S., Meier, K., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *Journal of the Acoustical Society of America*, *128*, 2182–2190.

Pfordresher, P. Q., & Demorest, S. M. (2021). The prevalence and correlates of accurate singing. *Journal of Research in Music Education*, *69*, 5–23.

Pinker, S. (1997). *How the mind works*. Norton.

Stadler Elmer, S. (2020). From canonical babbling to early singing and its relation to the beginnings of speech. In F. Russo, B. Ilari, & A. Cohen (Eds.), *Routledge companion to interdisciplinary studies in singing: Vol. 1 development* (pp. 17–24). Routledge.

Welch, G. F. (2006). Singing and vocal development. In G. McPherson (Ed.), *The child as musician: A handbook of musical development* (pp. 311–329). Oxford University Press.

Musical features emerging from a biocultural musicality

Tudor Popescu[a,b], Nathan Oesch[c] and Bryony Buck[d]

[a]Department of Behavioural and Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria; [b]Department of Neurology, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria; [c]Department of Psychology, Western University, London, ON N6A 3K7, Canada and [d]Institute of Musicians' Medicine, Carl Maria von Weber University of Music, Wettiner Platz 13, 01067 Dresden, Germany.
tudor.popescu@univie.ac.at, tudorpopescu.com
nathanoesch@yahoo.com
bryony.buck@gmail.com

doi:10.1017/S0140525X20001491, e87

Abstract

Savage et al. make a compelling case, Mehr et al. less so, for social bonding and credible signalling, respectively, as the main adaptive function of human musicality. We express general advocacy for the former thesis, highlighting: (1) overlap between the two; (2) direct versus derived biological functions, and (3) aspects of music embedded in cultural evolution, for example, departures from tonality.

Although not a novel hypothesis, Savage et al.'s music and social bonding (MSB) is an impressive synthesis of evidence suggesting social bonding is the primary adaptive function for the origin and evolution of musicality in humans. This intriguing idea has seen rapid development in recent years, both in terms of theoretical refinement and increasing evidence across human and nonhuman animals (Oesch, 2019). Putting aside relative strengths, three aspects, relevant to both articles, seem worthy of elaboration.

First, the analogy with vision borders on tautology. In opening, Savage et al. curiously state: "Turning to music, 'social bonding' provides an umbrella explanation analogous to 'vision is for seeing'." Indeed, vision is for seeing; but testing an adaptive explanation for a trait is to determine what *kind* of vision is best adapted for a particular organism within a specific ecological niche. For example, colour vision seems to have specifically evolved in howler monkeys in foraging for ripe fruit and leaves, whereas convergently evolved in bower birds for conspecific identification and assessing mate quality during courtship (for a review, see Gerl & Morris, 2008). Similarly, invoking social bonding as an "umbrella explanation" for the function of musicality blurs the contexts and functions in which musicality exists in human and nonhuman animals. For example, recent evidence suggests song can facilitate both social bonding (Oesch, 2019; Whittingham, Kirkconnell, & Ratcliffe, 1997) and courtship in various bird species (Catchpole & Slater, 2008), whereas similar effects are observed in humans (Oesch, 2020). Although Savage et al. do acknowledge it as "unlikely that a single 'main' evolutionary function for complex, multi-component abilities like language or music exists," we believe an ideal analysis would, when space allows, simultaneously: (1) avoid circular reasoning, (2) invoke substantive data from both human and nonhuman animals, and (3) be based on a distinction between the ancestral, primary, or direct function of a trait or adaptation like musicality and its derived evolutionary function(s) (Oesch, 2019, 2020). Fortunately, evolutionary theory offers clear criteria for distinguishing direct from derived functions: shared ancestry and phylogenetic history in the former case, and unshared with a common ancestor in the latter (Ridley, 2003). Moreover, in addition to social bonding, it seems likely that there were several other selective factors and derived functions at play, which Savage et al. at least implicitly acknowledge.

Second, is credible signalling truly disjunct from the social-bonding proposal? Ultimately, cooperation and even signalling are arguably part of the family of hypotheses epitomized by MSB. As such, the target articles appear closer to one another than to hypotheses following Pinker's "cheesecake" analogy. If Mehr et al.'s main thesis is that music-making is a proxy measure for signalled group coherence, then we fail to see how this is not a special case of MSB, albeit that signalling focuses on between-groups and social-bonding on within-group relations. In their group of friends example, surely *something* is required to explain why they are already closely-bonded enough to walk together in a dangerous neighbourhood at night: manifestations of musicality could themselves plausibly have contributed to the bonding being signalled. In fact, much of the evidence presented by Mehr et al. support a social-bonding interpretation at least as well as a signalling one – if, in fact, the dichotomy is warranted.

Third, although social bonding hypotheses provide a plausible seed for musicality, accounting for specific features of *music* necessitates deeper examination of the by-products of the gene-culture feedback loop, as played out many iterations later. The application of consonance versus dissonance in music is an interesting case in point. Musicality may well be what allows for the affective value of dissonance, just like – as Savage et al. note – metre aids synchronized performance. Indeed, dissonance's role in music, as in animal vocalizations, is arguably to calibrate the arousal "dial" (Fitch, Neubauer, & Herzel, 2002). Because of its deep biological substrate in the harmonic series of vertebrates' affective vocalizations (Bowling, Hoeschele, Kamraan, & Fitch, 2017; Bowling & Purves, 2015), sensory dissonance imparts strong affect that music aptly builds upon (Bravo et al., 2019) – once again, within a biology-culture loop. We feel the logic of such feedback loops – not lost on Savage, Mehr and colleagues – is emblematic for the apparent (mechanistic, rather than logical) circularity intrinsic to theories of cultural or biological evolution, and indeed ultimately intrinsic to any systematic account of connected events whereby past effects become future causes, as seen in history.

In Western tonal music, for example, although by no means rare, dissonance tends to be used as a "spice"; one that, in the language of predictive coding, "provides its own reward prediction error," as Savage et al. remark. Sensory and tonal dissonance (Johnson-Laird, Kang, & Leong, 2012) abounds in rock's distorted guitars, jazz's flatted-5th chords, European art music back to at least Monteverdi – and, plentifully, in the Second Viennese School's atonality. That ideas revolutionary at their time – such as Schoenberg's wish to "emancipate dissonance" and strive for total chromaticism (Harrison, 1996; Schoenberg, 2010) – only survived as an elitist niche, whereas music, including most art music, remained firm in its tonal structure, is evidence of the biologically-based invariance of universal features of music such as the cognitive framework of tonality (Mehr et al., 2019). It seems unlikely Schoenberg's idea would have arisen from social-related pressures, but rather from pressures of *cultural* evolution (arguably, the demand for novelty).

The Western musical canon is anchored in salient features of cognition (e.g., Huron, 2016). Although current evidence is mixed, humans and other species (Izumi, 2000; Sugimoto et al., 2009; Watanabe, Uozumi, & Tanaka, 2005) appear to prefer sounds with minimal roughness and/or abundant fusion (Masataka, 2006; Popescu et al., 2019; Trainor, Tsang, & Cheung, 2002). Thus, consonance preference appears not merely an arbitrary cultural invention – despite recent suggestions to the contrary (Bowling et al., 2017; McDermott, Schultz, Undurraga, & Godoy, 2016; McPherson et al., 2019) – but plausibly emerges from iterated cycles of gene-culture coevolution. In effect, Savage et al.'s MSB explains a biologically-evolved set of propensities upon which music builds, using universal principles of cognition such as prediction (Koelsch, Vuust, & Friston, 2019) and dendrophilia (Fitch, 2014). Once established, this basis can give rise to culturally-evolved features such as scales, some more pervasive across cultures than others (Fitch & Popescu, 2019; Mehr et al., 2019).

Acknowledgment. We gratefully thank W. Tecumseh Fitch for helpful discussions.

Financial support. TP was supported by the "Music and Language in the Human Brain" research-cluster grant from the Medical University of Vienna and the University of Vienna.

Conflict of interest. None.

References

Bowling, D. L., Hoeschele, M., Kamraan, Z. G., & 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*, *112*(36), 11155–11160. https://doi.org/10.1073/pnas.1505768112.

Bravo, F., Cross, I., Hopkins, C., Gonzalez, N., Docampo, J., Bruno, C., & Stamatakis, E. A. (2019). Anterior cingulate and medial prefrontal cortex response to systematically controlled tonal dissonance during passive music listening. *Human Brain Mapping*, *41*(1), 46–66. https://doi.org/10.1002/hbm.24786.

Catchpole, C. K., & Slater, P. J. (2008). *Bird song: Biological themes and variations*. Cambridge University Press.

Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews*, *11*(3), 329–364. https://doi.org/10.1016/j.plrev.2014.04.005.

Fitch, W. T., Neubauer, J., & Herzel, H. (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, *63*(3), 407–418. https://doi.org/10.1006/anbe.2001.1912.

Fitch, W. T., & Popescu, T. (2019). The world in a song. *Science*, *366*(6468), 944–945. https://doi.org/10.1126/science.aay2214.

Gerl, E. J., & Morris, M. R. (2008). The causes and consequences of color vision. *Evolution: Education and Outreach*, *1*(4), 476–486.

Harrison, T. (1996). *The emancipation of dissonance*. University of California Press.

Huron, D. (2016). *Voice leading: The science behind a musical art*. MIT Press.

Izumi, A. (2000). Japanese Monkeys perceive sensory consonance of chords. *The Journal of the Acoustical Society of America*, *108*(6), 3073–3078.

Johnson-Laird, P. N., Kang, O. E., & Leong, Y. C. (2012). On musical dissonance. *Music Perception: An Interdisciplinary Journal*, *30*(1), 19–35.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, *23*(1), 63–77. https://doi.org/10.1016/j.tics.2018.10.006.

Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental Science*, *9*(1), 46–50. https://doi.org/10.1111/j.1467-7687.2005.00462.x.

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*(7613), 547–550. https://doi.org/10.1038/nature18635.

McPherson, M. J., Dolan, S. E., Ossandon, T., Valdes, J., Undurraga, E. A., Jacoby, N., Godoy, R., & McDermott, J. (2019). Perceptual fusion of musical notes suggests universal representations of dissonance despite culture-dependent aesthetic associations. *The Journal of the Acoustical Society of America*, *145*(3), 1784–1784. https://doi.org/10.1121/1.5101526.

Mehr, S. A., Singh, M., Knox, D., Ketter, D., Pickens-Jones, D., Atwood, S., …, & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(970), 957–970. https://doi.org/10.31234/osf.io/emq8r.

Oesch, N. (2019). Music and language in social interaction: Synchrony, antiphony, and functional origins. *Frontiers in Psychology*, *10*, article 1514, 1–13. https://doi.org/10.3389/fpsyg.2019.01514.

Oesch, N. (2020). Evolutionary musicology. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of evolutionary psychological science* (pp. 1–6). Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6_2845-1.

Popescu, T., Neuser, M. P., Neuwirth, M., Bravo, F., Mende, W., Boneh, O., Moss, F. C., & Rohrmeier, M. (2019). The pleasantness of sensory dissonance is mediated by musical style and expertise. *Scientific Reports*, *9*(1), 1070–1080. https://doi.org/10.1038/s41598-018-35873-8.

Ridley, M. (2003). *Evolution* (3rd ed.). Wiley.

Schoenberg, A. (2010). *Style and idea: Selected writings*. University of California Press.

Sugimoto, T., Kobayashi, H., Nobuyoshi, N., Kiriyama, Y., Takeshita, H., Nakamura, T., & Hashiya, K. (2009). Preference for consonant music over dissonant music by an infant chimpanzee. *Primates*, *51*(1), 7. https://doi.org/10.1007/s10329-009-0160-3.

Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception: An Interdisciplinary Journal*, *20*(2), 187–194. https://doi.org/10.1525/mp.2002.20.2.187.

Watanabe, S., Uozumi, M., & Tanaka, N. (2005). Discrimination of consonance and dissonance in Java sparrows. *Behavioural Processes*, *70*(2), 203–208. https://doi.org/10.1016/j.beproc.2005.06.001.

Whittingham, L. A., Kirkconnell, A., & Ratcliffe, L. M. (1997). The context and function of duet and solo songs in the red-shouldered blackbird. *The Wilson Bulletin*, *109*, 279–289.

Isochrony, vocal learning, and the acquisition of rhythm and melody

Andrea Ravignani

Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics,
Nijmegen, 6525 XD, The Netherlands.
andrea.ravignani@mpi.nl; https://www.mpi.nl/people/ravignani-andrea

doi:10.1017/S0140525X20001478, e88

Abstract

A cross-species perspective can extend and provide testable predictions for Savage et al.'s framework. Rhythm and melody, I argue, could bootstrap each other in the evolution of musicality. Isochrony may function as a temporal grid to support rehearsing and learning modulated, pitched vocalizations. Once this melodic plasticity is acquired, focus can shift back to refining rhythm processing and beat induction.

Musicality consists of the (neuro)biological underpinnings to perceive and produce music. Research in the evolution of musicality needs cross-species evidence. As a parallel, to understand the evolution of bat wings, one asks why all other mammals lack wings and why other flying animals have evolved them. Similarly, our species only constitutes one datapoint to construct evolutionary hypotheses on musicality. Comparisons with other species are necessary to avoid post-hoc explanations of evolutionary traits.

Four concepts discussed in Savage et al. are key for understanding musicality, both in humans and other animals (Fig. 1). *Isochrony* describes metronomic temporal regularity, similar to the ticking of a clock (Merker, Madison, & Eckerdal, 2009; Ravignani & Madison, 2017). *Synchrony* is the perfect co-occurrence in time of two series of events, with no strong teleological or mechanistic focus (Kotz, Ravignani, & Fitch, 2018; Ravignani, 2017). *Vocal learning* is the ability to learn and modify non-innate vocalizations, including melodies (Lattenkamp & Vernes, 2018). *Beat induction* denotes a top-down capacity to induce a regular pulse from music and move in synchrony to it (Grahn & Brett, 2007; Honing, 2012).

Do other animals have these capacities supporting musicality? Isochrony appears in many species' communication (e.g., from lobster rattles to sea lion barks: Patek & Caldwell, 2006; Schusterman, 1977), autonomously-regulated behavior or (neuro)physiology. Synchrony is widespread but scattered across taxonomic groups (Ravignani, Bowling, & Fitch, 2014; Wilson & Cook, 2016). Vocal learning is rare but potentially arose multiple times in evolution because of different pressures across species (Garcia & Ravignani, 2020; Martins & Boeckx, 2020; Nowicki & Searcy, 2014). Beat induction has only been found in a few animals, as acknowledged by Savage and colleagues (Kotz et al., 2018; cf. Mehr et al., claiming its presence in many species).

Savage and colleagues briefly characterize these four abilities; this invites discussion of cross-species implications and predictions as to how they evolved to support musicality. I add a fifth, still largely unexplored capacity: vocal rhythms, which consist of producing, perceiving, learning, or imitating signals with accuracy in the temporal – as opposed to the spectral – domain. Although this capacity to precisely time one's vocalizations is related to its spectral counterpart, vocal rhythms also have their own mechanistic and communicative value (Wirthlin et al., 2019). I argue that, across species, these five capacities are linked, mapping them to Savage et al.'s framework.

The core of Savage et al.'s idea of melodic and rhythmic musicality features vocal learning and beat induction. These are also at the core of an influential hypothesis in evolutionary neuroscience (Patel, 2006), predicting in some cases their joint co-occurrence across species. However, a few outlier species point to a mismatch between the current data and the hypothesis' predictions (Cook, Rouse, Wilson, & Reichmuth, 2013), requiring an updated theoretical framework.

Within Savage et al.'s framework, I argue that rhythm and melody may have bootstrapped each other in humans and other species gradually, especially in social interactions, such as chorusing, turn-taking, and so forth (Christophe, Millotte, Bernal, & Lidz, 2008; Hannon & Johnson, 2005; Höhle, 2009; Ravignani et al., 2014). An isochronous sequence, such as the repetitive bark of a sea lion, provides a temporal grid of predictable sound events. Both the producer of an isochronous rhythm and its conspecifics can rely on this periodicity to learn and experiment in the spectral, hence melodic, domain during vocal learning: vocal emissions could be anchored to the onsets of the isochronous sequence (Merker et al., 2009). Hence, rhythmic isochrony may function as temporal grid to rehearse learnt vocalizations (and possibly orient attention; Bolger, Coull, & Schön, 2014; Cason, Astésano, & Schön, 2015; Jones, 2010; Norton, 2019). In turn, learnt, consolidated vocalizations may serve as a "spectral anchor" to segment conspecifics' temporal sequences (Hyland Bruno, 2017; Lipkind et al., 2013), also generating vocal rhythms. Therefore, melodic templates acquired via vocal learning can afford increased attentional or cognitive resources spent on the rhythmic domain, including temporal segmentation and regularization. This provides a bootstrapping mechanism for Savage et al.'s co-evolutionary dynamics to work, and a testbench for some signaling hypotheses in Mehr and colleagues.

This hypothesis generates several testable predictions. First, by testing species along the vocal learning continuum (Martins & Boeckx, 2020), and extending this continuum to beat induction,

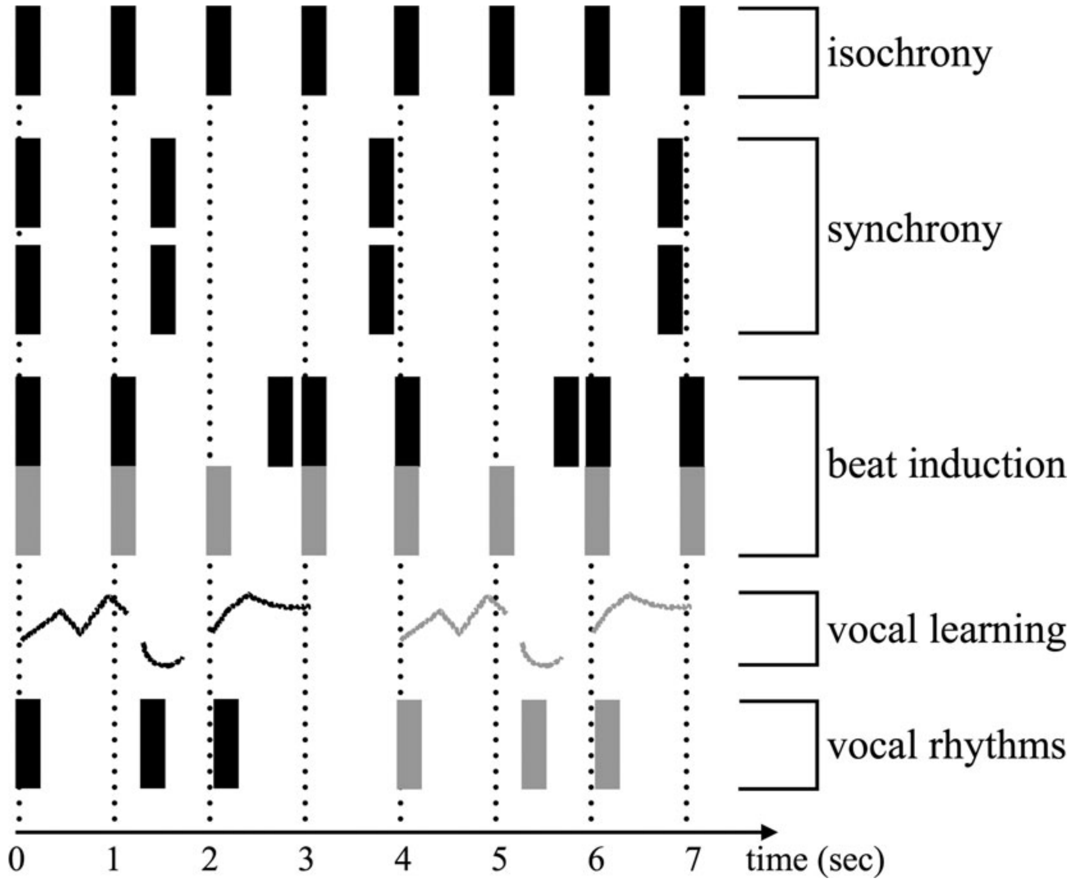


Figure 1. (Ravignani) Conceptualization of the four abilities partly explored in the target articles plus a fifth one, vocal rhythms, which deserves entering the discussion. Isochrony, when present in acoustic or motoric behaviors, may provide a clear, extremely predictable temporal grid, similar to squared notebooks guiding children who learn how to write. An isochronous pattern is, per se, neither musical nor demanding to produce or perceive. Isochrony has low entropy, definitely lower than expected for "musical" patterns (Milne & Herff, 2020; Ravignani & Madison, 2017). Production of isochrony can result from a motoric behavior entraining to a neural oscillator. Perception of isochrony requires, at least, comparing pairs of temporal intervals, an ability found in several species (e.g., Church & Lacourse, 1998; Heinrich, Ravignani, & Hanke, 2020; Ng, Garcia, Dyer, & Stuart-Fox, 2020). Although isochrony is characterized by equal timing in a series of events, synchrony requires pairwise coincidence of events from two series, neither of which needs to be isochronous (Ravignani, 2017). Given an acoustic sequence (black), beat induction consists of inferring an isochronous pulse (gray), which need not physically exist in the sequence (Honing, 2012; Kotz et al., 2018). Synchronization differs from beat induction in being independent from isochrony, relatively inflexible, achievable for a narrow range of tempi and unimodal (Patel, Iversen, Bregman, & Schulz, 2009). Vocal learning – here with emphasis in its spectral domain – includes, among other things, the capacity to copy (gray) a vocal signal (black) (Lattenkamp & Vernes, 2018; Wirthlin et al., 2019). A vocal rhythm (black) is a temporal pattern of events, which conveys most information in the temporal domain (Ravignani et al., 2019) and could also be learnt or imitated (gray).

species with a stronger sense of beat should be found among those with more developed vocal learning capacities. Chickens, great apes, parrots, and humans are examples of species predicted to show, in this order, increasing abilities in both domains. Second, isochrony should go hand in hand with synchrony but not with beat induction, so that species with developed isochrony should also synchronize. Third, empirical evidence for the rhythm–melody scaffolding process (Cason et al., 2012; Emmendorfer, Correia, Jansma, Kotz, & Bonte, 2020) could be obtained from large-scale developmental datasets, which should feature both humans and nonhuman animals, and contain data from as many capacities as possible from Figure 1. As ontogeny sometimes recapitulates phylogeny (e.g., Heldstab, Isler, Schuppli, & van Schaik, 2020), one would test whether the same stepwise processes hypothesized above appear in the first years of human life (Höhle, 2009). Fourth, a partial neural dissociation between rhythm and melody may occur early in life and become less severe over development; the dynamics of this dissociation could be tested via longitudinal neuroimaging studies (Bengtsson & Ullén, 2006; Salami, Wåhlin, Kaboodvand, Lundquist, & Nyberg, 2016). Fifth, within Savage et al.'s framework, physiological evidence for the rhythm–melody gradual interplay could come from measurements or manipulations of the dopaminergic reward system and the endogenous opioid system, testing whether they provide complementary, alternating effects. Finally, most of these putative links can be, following Savage et al., modulated by species-specific social factors, such us group density and social networks. Similarly, their value as honest signals can be tested to provide empirical support for Mehr et al. using, among others, methods from cultural evolution research (e.g., Lumaca et al., commentary on the target article by Mehr et al.; Miton, Vesper, Wolf, Knoblich, & Sperber, 2020).

To conclude, the frameworks proposed in both target articles can benefit from a finer dissection of core abilities for musicality (Fig. 1 and Honing, commentary on the target article by Savage et al.). These must then be tested across species to infer plausible evolutionary scenarios.

Acknowledgments. I am grateful to Henkjan Honing, Koen de Reus, Laura Verga, Massimo Lumaca, and Sonja Kotz for helpful discussion and feedback.

Financial support. Andrea Ravignani is supported by the Max Planck Society via an Independent Research Group Leader position.

Conflict of interest. None.

References

Bengtsson, S. L., & Ullén, F. (2006). Dissociation between melodic and rhythmic processing during piano performance from musical scores. *NeuroImage*, *30*(1), 272–284.

Bolger, D., Coull, J. T., & Schön, D. (2014). Metrical rhythm implicitly orients attention in time as indexed by improved target detection and left inferior parietal activation. *Journal of Cognitive Neuroscience*, *26*(3), 593–605.

Cason, N., Astésano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta Psychologica*, *155*, 43–50.

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652–2658.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, *51*(1–2), 61–75.

Church, R. M., & Lacourse, D. M. (1998). Serial pattern learning of temporal intervals. *Animal Learning & Behavior*, *26*(3), 272–289.

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, *127*(4), 412.

Emmendorfer, A. K., Correia, J. M., Jansma, B. M., Kotz, S. A., & Bonte, M. (2020). ERP mismatch response to phonological and temporal regularities in speech. *Scientific Reports*, *10*(1), 1–12.

Garcia, M., & Ravignani, A. (2020). Acoustic allometry and vocal learning in mammals. *Biology Letters*, *16*(7), 20200081.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, *50*(4), 354–377.

Heinrich, T., Ravignani, A., & Hanke, F. H. (2020). Visual timing abilities of a harbour seal (*Phoca vitulina*) and a South African fur seal (*Arctocephalus pusillus pusillus*) for sub- and supra-second time intervals. *Animal Cognition*, *23*(5): 851–859.

Heldstab, S. A., Isler, K., Schuppli, C., & van Schaik, C. P. (2020). When ontogeny recapitulates phylogeny: Fixed neurodevelopmental sequence of manipulative skills among primates. *Science Advances*, *6*(30), eabb4685.

Höhle, B. (2009). Bootstrapping mechanisms in first language acquisition. *Linguistics*, *47*(2), 359–382.

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, *1252*(1), 85–91.

Hyland Bruno, J. (2017). *Song rhythm development in zebra finches*. City University of New York.

Jones, M. R. (2010). Attending to sound patterns and the role of entrainment. In Nobre, A. C. & Coull, J. T. (eds.) *Attention and time* (pp. 317–330). Oxford University Press.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, *22*(10), 896–910.

Lattenkamp, E. Z., & Vernes, S. C. (2018). Vocal learning: A language-relevant trait in need of a broad cross-species approach. *Current Opinion in Behavioral Sciences*, *21*, 209–215.

Lipkind, D., Marcus, G. F., Bemis, D. K., Sasahara, K., Jacoby, N., Takahasi, M, … Tchernichovski, O. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*, *498*(7452), 104–108.

Martins, P. T., & Boeckx, C. (2020). Vocal learning: Beyond the continuum. *PLoS Biology*, *18*(3), e3000672.

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.

Milne, A. J., & Herff, S. A. (2020). The perceptual relevance of balance, evenness, and entropy in musical rhythms. *Cognition*, *203*, 104233.

Miton, H., Wolf, T., Vesper, C., Knoblich, G., & Sperber, D. (2020). Motor constraints influence cultural evolution of rhythm. *Proceedings of the Royal Society B*, *287* (1937), 20202001.

Ng, L., Garcia, J. E., Dyer, A. G., & Stuart-Fox, D. (2020). The ecological significance of time sense in animals. *Biological Reviews*, *96*(2), 526–540.

Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current opinion in Neurobiology*, *28*, 48–53.

Norton, P. (2019). *Isochronous rhythmic organization of learned animal vocalizations*. Doctoral dissertation.

Patek, S. N., & Caldwell, R. L. (2006). The stomatopod rumble: Low frequency sound production in *Hemisquilla californiensis*. *Marine and Freshwater Behaviour and Physiology*, *39*(2), 99–111.

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*(1), 99–104.

Patel, A., Iversen, J., Bregman, M., & Schulz, I. (2009). Studying synchronization to a musical beat in nonhuman animals. *Annals of the New York Academy of Sciences*, *1169*(1), 459–469.

Ravignani, A. (2017). Interdisciplinary debate: Agree on definitions of synchrony. *Nature*, *545*, 158.

Ravignani, A., Bowling, D. L., & Fitch, W. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology*, *5*, 1118.

Ravignani, A., Kello, C. T., De Reus, K., Kotz, S. A., Dalla Bella, S., Méndez-Aróstegui, M., …de Boer, B. (2019). Ontogeny of vocal rhythms in harbor seal pups: An exploratory study. *Current Zoology*, *65*(1), 107–120.

Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, *8*, 1820.

Salami, A., Wåhlin, A., Kaboodvand, N., Lundquist, A., & Nyberg, L. (2016). Longitudinal evidence for dissociation of anterior and posterior MTL resting-state connectivity in aging: Links to perfusion and memory. *Cerebral Cortex*, *26*(10), 3953–3963.

Schusterman, R. J. (1977). Temporal patterning in sea lion barking (*Zalophus californianus*). *Behavioral Biology*, *20*(3), 404–408.

Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin & Review*, *23*(6), 1647–1659.

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., … Yartsev, M. M. (2019). A modular approach to vocal learning: Disentangling the diversity of a complex behavioral trait. *Neuron*, *104*(1), 87–99.

A boldly comparative approach will strengthen co-evolutionary accounts of musicality's origins

Luke Rendell[a], Emily L. Doolittle[b], Ellen C. Garland[c] and Alex South[d]

[a]Centre for Social Learning and Cognitive Evolution & Sea Mammal Research Unit, University of St Andrews, St Andrews, Fife KY16 9TH, UK; [b]Research and Knowledge Exchange, Royal Conservatoire of Scotland, Glasgow G2 3DB, UK; [c]Centre for Social Learning and Cognitive Evolution & Sea Mammal Research Unit, University of St Andrews, St Andrews, Fife KY16 8LB, UK and [d]Centre for Social Learning and Cognitive Evolution & Sea Mammal Research Unit, University of St Andrews, and Royal Conservatoire of Scotland, Glasgow G2 3DB, UK.

ler4@st-andrews.ac.uk; www.st-andrews.ac.uk/biology/people/ler4
e.doolittle@rcs.ac.uk; https://www.rcs.ac.uk/staff/emily-doolittle
ecg5@st-andrews.ac.uk; www.st-andrews.ac.uk/biology/people/ecg5;
a.south@rcs.ac.uk; www.alexsouth.org

doi:10.1017/S0140525X20001715, e89

Abstract

Focus on the evolutionary origins of musicality has been neglected relative to attention on language, so these new proposals are welcome stimulants. We argue for a broad comparative approach to understanding how the elements of musicality evolved, and against the use of overly simplistic evolutionary accounts.

> "there is no reason to imagine that it emerged one day wholly made by evolution … recognize that there is no 'music in and of itself,' no musical essence, but only some distinct capacities that one day converged toward what we today call music." (Molino, 2000, p. 169)

It is exciting to see the evolution of music, or rather, musicality (Honing, 2018), neglected in the evolutionary sciences relative to the question of language, being highlighted by Savage et al. and Mehr et al. As with language, we are stymied by a sample size of one, but we can approach the evolutionary question by recognizing musicality is not one thing (Molino, 2000), but a collection of cognitive abilities and physiological responses, and studying its constituent elements. The origins of synchronicity, rhythmic entrainment, vocal learning, sequence learning, convergence on a series of pitches, or creative innovation can all be informed by studies on nonhumans, to explore the evolutionary forces behind human musicality. Here, we expand upon the cross-species considerations briefly mentioned by Savage et al. to argue that research into the evolutionary origins of music can be enriched by a broad comparative perspective.

Both Savage et al. and Mehr et al. emphasize social bonds in their accounts. The former emphasize social bonding and the latter the credible signalling of that bonding. The contrast is primarily about who the audience is – social partners, or potential competitors. From a signal evolution perspective, these accounts can be unified – any acoustic signal produced to support or strengthen a social bond is simultaneously a cue for any third party, containing pertinent information about that bond. We see no reason why selection for bonding and credible signalling should not work simultaneously and in the same direction – consider, for example, the apparent intermixing of these functions in duetting neotropical wrens, which are both thought to promote vocal learning (Rivera-Cáceres, Quirós-Guerrero, Araya-Salas, & Searcy, 2016; Templeton, Ríos-Chelén, Quirós-Guerrero, Mann, & Slater, 2013).

Nonetheless, one surely has to have a social bond before one can credibly signal about it. Several of the social bonding functions proposed by Savage et al. can be supported by examples from animal studies. Humpback whales provide prime examples of "learned patterns of features" along with a tendency to "deviate from predicted combinations of features" (Allen, Garland, Dunlop, & Noad, 2018) whereas the vocal clans of sperm whales reflect large scale social structures akin to Savage et al.'s "imagined communities" (Rendell & Whitehead, 2003). The synchronized and overlapping production of the latter is also analogous to Savage et al.'s "prediction" and overall supports a social bonding hypothesis (Whitehead & Rendell, 2015).

To illustrate how comparative approaches can both support and challenge Savage et al.'s account, consider how the absence and presence of a particular form of temporal synchronization in these two species relates to their differing social structures. Humpback song contains some music-like features: hierarchical structure, repetitive themes, and consistent phrase lengths (Martinelli, 2008; Payne, 1995). Savage et al.'s account suggests this would arise within a complex social system in which other social bonding mechanisms had proven to be insufficient for the expected group size. However, when compared with odontocetes, humpback social networks are relatively unstructured, with little evidence of long-term bonds between adults (Allen, Weinrich, Hoppitt, & Rendell, 2013). Why do humpbacks have such strikingly music-like features in structurally complex song whereas sperm whales, with highly structured societies featuring enduring social bonds (Whitehead et al., 2012) have only patterned series of clicks?

One answer relates to the kind of musical features absent in humpback song, specifically the lack of temporal synchronization between individuals, resulting in an "asynchronous chorus" (Herman, 2017). Savage et al. present convincing evidence that synchronization can give rise to prosocial behaviour, we shouldn't be too surprised to find it absent in communities lacking long-term bonds. Sperm whales, in contrast, do coordinate their codas (Schulz, Whitehead, Gero, & Rendell, 2008), so we have an illustration of how music-like features of vocalization systems do not come in fixed packages (Nettl, 1999; Nketia, 1984), and how particular components of musicality may correlate with social structures. The larger point is that both communication systems and social structures exist in a plurality of different forms, and comparative study can reveal contingency in relationships that can superficially appear evolutionarily necessary when only examined in one species.

Mehr et al.'s critiques of the social bonding hypothesis, although explicitly rooted in an evolutionary psychology perspective, rely on over-simplistic evolutionary arguments and do not reflect the nuance of current evolutionary thinking. First, their "superfluous" claim assumes that the mere presence of fitness benefits negates the need for stress reduction, but this ignores at a fundamental level the real tensions between cooperation and conflict that need to be resolved with sometimes costly suppressor mechanisms in any cooperative system, even down to the genome level (Scott & West, 2019). Mehr et al.'s credible signal account turns the explanation of developing evolutionarily stable social bonds in the first place into somebody else's problem. This seems to us inherently weaker than an account where the elements of musicality coevolved with social bonds, simultaneously providing useful cues that could themselves become targets of selection, as in the duetting wrens. Second, the claimed confusion between proximate and ultimate causation doesn't reflect the difficulty inherent in such distinctions engendered by gene-culture co-evolutionary dynamics (Laland, Sterelny, Odling-Smee, Hoppitt, & Uller, 2011). Such arguments against gene-culture co-evolutionary accounts can only endure by ignoring the accumulating evidence for gene-culture co-evolutionary dynamics in nonhumans (Whitehead, Laland, Rendell, Thorogood, & Whiten, 2019), evidence that lends credence to Savage et al.'s co-evolutionary account.

Finally, the different accounts offered here map onto debates about information and influence in animal communication (see Stegmann, 2013). Although Mehr et al. give an explicitly information-based account of music, musicologists more typically characterize it in terms of influence, or "affect" (e.g., Juslin, 2010), and this may provide both a stronger link to potential evolutionary precursors as revealed by a comparative perspective and support for an account under which music coevolved with human hypersociality.

Financial support. ECG is supported by a Royal Society University Research Fellowship (UF160081), AS is supported by a St Leonard's Interdisciplinary Scholarship from the University of St Andrews and Royal Conservatoire of Scotland.

Conflict of interest. None.

References

Allen, J., Garland, E. C., Dunlop, R. A., & Noad, M. J. (2018). Cultural revolutions reduce complexity in the songs of humpback whales. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1891), 20182088. https://doi.org/10.1098/rspb.2018.2088.

Allen, J., Weinrich, M., Hoppitt, W., & Rendell, L. (2013). Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science (New York, N.Y.)*, *340*(6131), 485–488. https://doi.org/10.1126/science.1231976.

Herman, L. M. (2017). The multiple functions of male song within the humpback whale (*Megaptera novaeangliae*) mating system: Review, evaluation, and synthesis. *Biological Reviews*, *92*, 1795–1818. https://doi.org/https://doi.org/10.1111/brv.12309.

Honing, H. (Ed.). (2018). *The origins of musicality*. MIT Press.

Juslin, P. N. (2010). *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230143.001.0001.

Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and effect in biology revisited: Is Mayr's proximate-ultimate dichotomy still useful? *Science (New York, N.Y.)*, *334*(6062), 1512–1516. https://doi.org/10.1126/science.1210879.

Martinelli, D. (2008). *Of birds, whales and other musicians: An introduction to zoomusicology*. University of Scranton Press.

Molino, J. (2000). Toward an evolutionary theory of music and language. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 165–176). The MIT Press.

Nettl, B. (1999). An ethnomusicologist contemplates universals in musical sound and musical culture. In C. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 473–472). MIT Press. https://doi.org/10.7551/mitpress/5190.003.0032.

Nketia, J. H. K. (1984). Universal perspectives in ethnomusicology. *The World of Music*, *26*(2), 3–24. http://www.jstor.org/stable/43560981.

Payne, R. (1995). *Among whales*. Simon and Schuster.

Rendell, L. E., & Whitehead, H. (2003). Vocal clans in sperm whales (*Physeter macrocephalus*). *Proceedings of the Royal Society B: Biological Sciences*, *270*, 225–231. https://doi.org/10.1098/rspb.2002.2239.

Rivera-Cáceres, K. D., Quirós-Guerrero, E., Araya-Salas, M., & Searcy, W. A. (2016). Neotropical wrens learn new duet rules as adults. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1843), 20161819. https://doi.org/10.1098/rspb.2016.1819.

Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: Insights into communication function. *Animal Behaviour*, *76*, 1977–1988. https://doi.org/10.1016/j.anbehav.2008.07.032.

Scott, T. W., & West, S. A. (2019). Adaptation is maintained by the parliament of genes. *Nature Communications*, *10*(1), 5163. https://doi.org/10.1038/s41467-019-13169-3.

Stegmann, U. (Ed.). (2013). *Animal communication theory: Information and influence*. Cambridge University Press.

Templeton, C. N., Ríos-Chelén, A. A., Quirós-Guerrero, E., Mann, N. I., & Slater, P. J. B. (2013). Female happy wrens select songs to cooperate with their mates rather than confront intruders. *Biology Letters*, *9*(1), 20120863. https://doi.org/10.1098/rsbl.2012.0863.

Whitehead, H., Antunes, R., Gero, S., Wong, S. N. P., Engelhaupt, D., & Rendell, L. (2012). Multilevel societies of female sperm whales (*Physeter macrocephalus*) in the Atlantic and Pacific: Why are they so different? *International Journal of Primatology*, *33*(5), 1142–1164. https://doi.org/10.1007/s10764-012-9598-z.

Whitehead, H., Laland, K. N., Rendell, L., Thorogood, R., & Whiten, A. (2019). The reach of gene–culture coevolution in animals. *Nature Communications*, *10*, 2405. https://doi.org/10.1038/s41467-019-10293-y.

Whitehead, H., & Rendell, L. (2015). *The cultural lives of whales and dolphins*. University of Chicago Press.

Clarifying the link between music and social bonding by measuring prosociality in context

Matthew E. Sachs[a], Oriel FeldmanHall[b] and Diana I. Tamir[c]

[a]Department of Psychology and Center for Science and Society, Columbia University, New York, NY 10027, USA; [b]Department of Cognitive, Linguistic, Psychological Sciences, Brown University, Providence, RI 02912, USA and [c]Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA.
ms5924@columbia.edu
Oriel_FeldmanHall@brown.edu
dtamir@princeton.edu

doi:10.1017/S0140525X2000117X, e90

Abstract

To corroborate the music and social bonding hypothesis, we propose that future investigations isolate specific components of social bonding and consider the influence of context. We deconstruct and operationalize social bonding through the lens of social psychology and provide examples of specific measures that can be used to assess how the link between music and sociality varies by context.

Savage et al. (2020) present cross-disciplinary evidence that the evolutionary origins of musicality stem primarily from its ability to foster social bonding. Although the authors provide predictions for potential investigations that could corroborate this, we suggest future human research would benefit from combining two complementary approaches: (1) specifying which features of musicality (synchronization, learning, and listening) relate to which building blocks of social bonding (identity fusion and coalition formation) and (2) characterizing how context modulates the relationship.

Music has the ability to influence the *process* of forming affiliative connections and the downstream *effects* of forming those bonds. These separable components of social bonding are often conflated in Savage et al. (2020), which can muddle the predicted associations with different musical experiences. In an attempt to clarify, one avenue in which music can initiate the process of bonding is through *identity fusion* – the feeling of oneness with others through alignment of actions, affect, and/or preferences (Swann, Jetten, Gómez, Whitehouse, & Bastian, 2012). Playing music together can synchronize movement (drumming and dancing), voice (singing; Mogan, Fischer, and Bulbulia, 2017), and/or emotional states and such experiences can lead to increases in perceived identity fusion (Lawendowski & Besta, 2020; Páez, Rimé, Basabe, Wlodarczyk, & Zumeta, 2015; Swann, Gómez, Seyle, Morales, & Huici, 2009). Non-synchronized musical experiences, such as simply liking the same music, may also lead to a shared sense of identity (Boer et al., 2012), although the extent of the feelings of fusion have yet to be fully explored. *Coalition formation* – which involves combining efforts to achieve a common goal – is another way in which social bonds can be formed through music. Playing music together, whether exactly synchronized or not, is a collective experience that involves shared intentionality (Tomasello, Carpenter, Call, Behne, & Moll, 2005). People perceive that musicians playing in time together are more likely to help each other than musicians playing out of time (Hagen & Bryant, 2003).

Playing music together can also influence the downstream effects of forming bonds, such as increasing trusting and cooperative behaviors. In lab settings, this can be measured with economic games, where people decide how to allocate money. After a joint singing task, for example, students were more willing to contribute money to the group during a public-goods game (Wiltermuth & Heath, 2009), than those who sang asynchronously or did not sing at all. The effects of forming social bonds can also be assessed with out-of-lab measures of *situational prosociality*, in which, after an experimental manipulation, participants are asked by a confederate for help (Lefevor, Fowers, Ahn, Lang, & Cohen, 2017). Joint music-making has also been shown to increase subsequent helping behaviors, relative to non-musical, non-synchronized tasks (Kirschner & Tomasello, 2010; Reddish, Tong, Jong, Lanman, & Whitehouse, 2016).

The influence of musical experiences that do not involve movement synchronization, such as listening to music together or learning musical sequences from others, on the process of forming social bonding is less clear. Although it may be that listening to music together leads to forming bonds because of the aural or emotional synchronization, the framework provided in Savage et al. (2020) does not explicitly state the process by which different features of non-synchronized musical experiences play a role in forming bonds and their subsequent effects.

To systematically characterize the links between musicality and various social bonding processes, future research should pair an isolated feature of musicality (e.g., playing, learning, dancing, or listening together) with measures of identity fusion or coalition formation. By fully mapping this space, research can more precisely assess the subsequent effects of each pairing. Imagine a study in which participants collectively listen to music (compared to a non-musical stimulus) and are subsequently asked to report feelings of perceived identity fusion. This can be compared to the effects of social bonding processes, as measured by the types of decisions made in the public-goods game and helping behaviors outside of the lab. Providing evidence of this link between self-report and actual behavior would establish the effect of musicality on components of social bonding.

It is also important to consider *how* social context modulates the relationship between features of musicality and social bonding (Tamir & Hughes, 2018). Situational constraints can impact the degree to which a person desires or seeks out social (FeldmanHall, Raio, Kubota, Seiler, & Phelps, 2015) or musical rewards (Sachs, Damasio, & Habibi, 2020; Thielmann, Spadaro, & Balliet, 2020). The number of people around, the degree of closeness one feels toward them, and the degree of certainty that they will behave prosocially within the group can swiftly shift the tensions embedded in the social dynamic (FeldmanHall & Shenhav, 2019). Research reveals that these factors influence emotional engagement when listening to music and willingness to synchronize with another (Miles, Griffiths, Richardson, & Macrae, 2009). In short, music's ability to foster social bonds is likely determined by the social context.

We take the Registered Report (Savage et al., 2020) as a practical example of how context can be incorporated into empirical investigations; it is predicted that cooperation would increase in groups performing a joint vocalization task with an accompanying beat. Should a null finding be observed between groups who experience a beat compared to those who do not, the authors suggest this can be taken as evidence that musical synchronization does not facilitate cooperation any more so than basic language synchronization. However, a null finding may alternatively reflect the modulatory effect of context. Synchronizing to a beat with strangers might prove to be more uncomfortable compared to synchronizing to a beat with loved ones. Including additional conditions in which participants perform the task in dyads versus groups or with close others versus strangers would help drill down on how social context modulates the relationship between musicality and social bonding.

Here, we offer two points to sharpen the hypotheses laid out in Savage et al. (2020). Considering both the types of processes that lead to social bonding and the context in which they arise can clarify the evolutionary significance of musicality as a unique source of social bonding. We hope this framework will inform and inspire future empirical research aiming to test the theory that music is unique in its capacity to foster affiliative connections at a larger scale.

Financial support. No specific grant from any funding agency in the public, commercial, or not-for-profit sectors was received for this manuscript.

Conflict of interest. None.

References

Boer, D., Fischer, R., Tekman, H. G., Abubakar, A., Njenga, J., & Zenger, M. (2012). Young people's topography of musical functions: Personal, social and cultural experiences with music across genders and six societies. *International Journal of Psychology*, *47*(5), 355–369. https://doi.org/10.1080/00207594.2012.656128.

FeldmanHall, O., Raio, C. M., Kubota, J. T., Seiler, M. G., & Phelps, E. A. (2015). The effects of social context and acute stress on decision making under uncertainty. *Psychological Science*, *26*(12), 1918–1926.

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*(5), 426–435.

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*(1), 21–51.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364.

Lawendowski, R., & Besta, T. (2020). Is participation in music festivals a self-expansion opportunity? Identity, self-perception, and the importance of music's functions. *Musicae Scientiae: The Journal of the European Society for the Cognitive Sciences of Music*, *24*(2), 206–226.

Lefevor, G. T., Fowers, B. J., Ahn, S., Lang, S. F., & Cohen, L. M. (2017). To what degree do situational influences explain spontaneous helping behaviour? A meta-analysis. *European Review of Social Psychology*, *28*(1), 227–256.

Miles, L. K., Griffiths, J. L., Richardson, M. J., & Macrae, C. N. (2009). Too late to coordinate: Contextual influences on behavioral synchrony. *European Journal of Social Psychology*, *85*, 52–60. https://doi.org/10.1002/ejsp.721.

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.

Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*(5), 711–729.

Reddish, P., Tong, E. M. W., Jong, J., Lanman, J. A., & Whitehouse, H. (2016). Collective synchrony increases prosociality towards non-performers and outgroup members. *British Journal of Social Psychology*, *55*(4), 722–738.

Sachs, M. E., Damasio, A., & Habibi, A. (2020). Unique personality profiles predict when and why sad music is enjoyed. *Psychology of Music* (p. 030573562093266), 1–20. https://doi.org/10.1177/0305735620932660.

Savage, P. E., Yamauchi, M., Hamaguchi, M., Tarr, B., Kitayama, Y., & Fujii, S. (2020). Rhythm, synchrony, and cooperation. https://psyarxiv.com/46bd9/download?format=pdf.

Swann, W. B., Jr., Gómez, A., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, *96*(5), 995–1011.

Swann, W. B., Jetten, J., Gómez, A., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review*, *119*(3), 441–456.

Tamir, D. I., & Hughes, B. L. (2018). Social rewards: From basic social building blocks to Complex social behavior. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *13*(6), 700–717.

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin*, *146*(1), 30–90.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences*, *28*(5), 675–691, discussion 691–735.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5.

The evolution of music as artistic cultural innovation expressing intuitive thought symbolically

Valerie van Mulukom

Centre for Trust, Peace and Social Relations, Coventry University, Innovation Village IV5, Coventry CV1 2TL, UK.
valerie.vanmulukom@coventry.ac.uk; https://valerievanmulukom.com/

doi:10.1017/S0140525X20001260, e91

Abstract

Music is an artistic cultural innovation, and therefore it may be considered as intuitive thought expressed in symbols, which can efficiently convey multiple meanings in learning, thinking, and transmission, selected for and passed on through cultural evolution. The symbolic system has personal adaptive benefits besides social ones, which should not be overlooked even if music may tend more to the latter.

Savage et al. and Mehr et al. (S&M) have put forward an impressive synthesis of evolutionary theories of music and musicality, but overlooked the fact that music is an art. I suggest that the core biological components of human musicality evolved as they did for the other arts: to support the symbolic system, which is important at both personal and social levels. I argue that: (1) the features of music which make them memorable and rewarding are a result of cultural evolution for symbols rather than gene-culture coevolution; and (2) music, like other arts, does not just have social benefits but also personal ones.

The arts encompass a wide range of apparently disjointed media: music, dance, painting, poetry, theatre, sculpture, fiction, and so forth. One would be hard-pressed to imagine an underlying, heritable cognitive mechanism to produce all, and a "cognitive toolkit" (Savage et al., target article) may indeed be more likely. Nevertheless, we understand "the arts" as a valid category. I argue that what unifies these activities is that they (1) express intuitive thoughts symbolically, rendering them meaningful, and (2) can be accompanied by extraordinary experiences, rendering them significant.

Intuitive thought is considered here as the main cognitive mechanism with which we process the world around us, but which does not reach conscious awareness. Specific gut feelings in this framework are what happens when such subconscious "thoughts" reach a sufficiently high level of salience and just breach conscious awareness. The arts are intuitive thoughts expressed in a symbolic way (to ourselves, or an audience), thus not fully rationalized and explained in language, but made tangible nonetheless. These symbolic artefacts may be material, but need not be, for example, as with music or plays.

The way in which these intuitive thoughts or feelings are expressed artistically is through simplifications, formalizations, repetitions, exaggerations and elaborations of ordinary materials, thus elevating their status to extraordinary (Dissanayake, 2009), and imbuing the symbols with personal and social significance (Alcorta, 2013). The symbols are created in a way that attracts attention, sustains interest, and creates, shapes, and activates emotion and intuition in their creator and/or audience (Dissanayake, 2009). This facilitates their mnemonic retention, social and cultural transmission, and therefore cultural selection (Atran & Norenzayan, 2004). The rewarding feelings of fulfilment of expectation within prediction (common for predictive processing across the board, and not specific to music) further promotes their endurance. Savage et al. appear to mix up the evolution of music (cultural) and musicality (biological) in this regard.

Symbols, by capturing many meanings – previously only existent in intuitive "thought" – at once in an efficiently packaged, tangible, and memorable form (Alcorta, 2013), allow for better learning and transmission of, and thinking about, ideas and feelings (Deacon, 1998). Therefore, symbolization has important personal *and* social benefits, and evolved through both, in a way similar to language, which also may be considered necessary for both thinking and communicating.

Cognitive play theories explain the personal benefits by the ability to mentally "try out" scenarios and ideas (Boyd, 2009), including creatively recombined imaginations (van Mulukom, 2020). S&M pay too little attention to spontaneous creativity of cultural innovations, particularly in solo contexts or in contexts where no group identification is present or required, such as singing or playing an instrument at home. Such creative solo acts are widespread and important, and contribute to psychological well-being (MacDonald, 2013).

The symbols' facilitation of social transmission on the other hand allows for the regulation of social interactions (Carroll, 2012) and promotion of social cohesion. Social bonding is thus a highly important aspect of the symbolic system (including music), but not the only major one. In the case of music, it may be that the social benefit is emphasized over the personal one, as music is more abstract and intuitive, and less language-reliant, than the other arts (cf. Fitch, 2013), and importantly includes the facilitation of synchrony (e.g., through an external beat), which is well-known to increase endorphins and support social bonding (Lang, Bahna, Shaver, Reddish, & Xygalatas, 2017; Launay, Tarr, & Dunbar, 2016).

Another aspect overseen by Savage et al. is the (evolutionary) initial use of music, despite comparing it to fire-making or dairy farming. I suggest that the initial utilitarian trigger of the arts may have been the symbolic representation of experiences (Zaidel, 2018), which could have aided in learning from the past and planning for the future (van Mulukom, 2020).

Moreover, the events of making and experiencing music can bring about extraordinary experiences, filled with awe, feelings of connectedness to something bigger, and other significant feelings. Such experiences are common in rituals (Charles et al., 2020a, 2020b), and can be induced by music, psychedelics (van Mulukom, Patterson, & van Elk, 2020), and other ritual behaviours. Their mechanisms and underlying effects are the same: The reduction of our "rational voice" (i.e., executive or cognitive control, supported by the prefrontal cortex) and "rational self" ("me" rather than "I" in William James' terminology; James [1950], supported by the default mode network). These reductions allow our intuitive thinking to "take over" and let "I" thrive (van Elk, van der Zwaag, Arciniegas, van Schie, & Sauter, 2019), in a process also called "flow" (Csikszentmihalyi, 1990) and "absorption" (Luhrmann, Nusbaum, & Thisted, 2010). As a result of being in this state, declarative memories are often reduced (van Mulukom, 2017). The events also induce surges of dopamine and endorphins (cf. Savage et al.), further boosting the significance, motivational force, and memorability of the resulting symbols.

The experience of rational ego dissolution in these events can lead to connectedness with others, but also with God, nature, or the universe, as well as values which the symbols represent. Such experiences can have transformative effects, contributing to one's identity and well-being (van Mulukom, 2017), as well as to social bondedness, especially when these experiences are shared with a small group (Tasuji, Reese, van Mulukom, & Whitehouse, 2020). Therefore, such events may be used as a costly signal (cf. Mehr, Krasnow, Bryant, & Hagen, target article) or as a sign of fitness in sexual selection, but it was not evolved for it specifically – the symbols and their functions are themselves inherently valuable.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Alcorta, C. S. (2013). Religious ritual and modes of knowing: Commentary on the cognitive resource depletion model of ritual. Religion, *Brain & Behavior*, *3*(1), 55–58.

Atran, S., & Norenzayan, A. (2004). Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences*, *27*(6), 713–730.

Boyd, B. (2009). *On the origin of stories: Evolution, cognition, and fiction*. Harvard University Press.

Carroll, J. (2012). The truth about fiction: Biological reality and imaginary lives. *Style*, *46* (2), 129–160.

Charles, S., van Mulukom, V., Brown, J. E., Watts, F., Dunbar, R., & Farias, M. (2020a). United on Sunday: The effects of secular rituals on social bonding and affect. *PLoS ONE*, *16*(1), e0242546. doi: https://doi.org/10.1371/journal.pone.0242546.

Charles, S., van Mulukom, V., Farias, M., Brown, J. E., Delmonte, R., de Maraldi, E. O., … Dunbar, R. (2020b). Religious rituals increase social bonding and pain threshold. *PsyArXiv*, 1–27. https://doi.org/10.31234/osf.io/my4hs.

Csikszentmihalyi, M. (1990). *Flow*. Harper & Row.

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. WW Norton & Company.

Dissanayake, E. (2009). The artification hypothesis and its relevance to cognitive science, evolutionary aesthetics, and neuroaesthetics. *Cognitive Semiotics*, *5*, 136–191.

Fitch, W. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, *7*, 68.

James, W. (1950). *The principles of psychology, vols. I and II*. Dover Publications.

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.

Launay, J., Tarr, B., & Dunbar, R. I. (2016). Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology*, *122*(10), 779–789.

Luhrmann, T. M., Nusbaum, H., & Thisted, R. (2010). The absorption hypothesis: Learning to hear God in evangelical Christianity. *American Anthropologist*, *112*(1), 66–78.

MacDonald, R. A. (2013). Music, health, and well-being: A review. *International Journal of Qualitative Studies on Health and Well-Being*, *8*(1), 20635.

Tasuji, T., Reese, E., van Mulukom, V., & Whitehouse, H. (2020). Band of mothers: Childbirth as a female bonding experience. *PLoS ONE*, *15*(10), e0240175.

van Elk, M., van der Zwaag, W., Arciniegas, A., van Schie, H. T., & Sauter, D. (2019). Reduced default mode network activity during the experience of awe. *Human Brain Mapping*, *40*(12), 3561–3574.

van Mulukom, V. (2017). Remembering religious rituals: Autobiographical memories of high-arousal religious rituals considered from a narrative processing perspective. *Religion, Brain & Behavior*, *7*(3), 191–205.

van Mulukom, V. (2020). The evolution of imagination and fiction through generativity and narrative. In J. Carroll, M. Clasen & E. Jonsson (Eds.), *Evolutionary perspectives on imaginative culture* (pp. 53–70). Springer.

van Mulukom, V., Patterson, R., & van Elk, M. (2020). Broadening your mind to include others: The relationship between serotonergic psychedelic experiences and maladaptive narcissism. *Psychopharmacology*, *237*, 2725–2737.

Zaidel, D. W. (2018). Culture and art: Importance of art practice, not aesthetics, to early human culture. In J. F. Christensen & A. Gomila (Eds.), *Progress in brain research* (Vol. *237*, pp. 25–40). Elsevier.

Singing is not associated with social complexity across species

Jan Verpooten[a] and Marcel Eens[b]

[a]Behavioral Engineering Research Group, KU Leuven, Naamsestraat 69 – bus 3545, 3000 Leuven, Belgium and [b]Behavioural Ecology & Ecophysiology, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium.
jan.verpooten@kuleuven.be; https://www.janverpooten.com/
marcel.eens@uantwerpen.be; https://www.uantwerpen.be/en/staff/marcel-eens/

doi:10.1017/S0140525X20001120, e92

Abstract

Based on their social bonding hypothesis, Savage et al. predict a relation between "musical" behaviors and social complexity across species. However, our qualitative comparative review suggests that, although learned contact calls are positively associated with complex social dynamics across species, songs are not. Yet, in contrast to songs, and arguably consistent with their functions, contact calls are not particularly music-like.

Savage et al. posit that if their main hypothesis is correct, namely that music serves to enhance social bonds, "across species, production or proficiency in 'musical' behaviors should predict both the number and complexity of social bonds."

However, this is not borne out by cross-species evidence of natural vocal production behavior. Across species, social complexity and the need for enhanced social bonding is positively associated with the production of learned contact calls, but not with songs (Nowicki & Searcy, 2014; Sewall, 2015). Production of song appears to be independent from social complexity. Yet, considering their design features, songs are more music-like than learned contact calls. From a signaling theory perspective, these differences in form between songs and contact calls make functional sense.

Both Savage et al. and Mehr et al. highlight vocal learning as an important design feature of music. Hence, we restrict our discussion to learned vocalizations. Fitch (2006) defines songs as "complex learned vocalizations" and notes that this definition "almost coincidentally" also applies to human song. However, not all complex vocal learners produce song, rather several use their vocal learning abilities to produce vocalizations commonly called contact calls. Marler's (2004) distinction highlights their design differences: "On a structural level, songs are usually longer and more complex acoustically, involving a variety of different notes and syllables, ordered in statistically reliable sequences; calls are often short, monosyllabic with simple frequency patterning, delivered in what often appears to be a disorderly fashion." This contrast between songs and calls suggests that songs are music-like in design (i.e., complex, with ordered notes and syllables in reliable sequences), which is why they may actually be called "songs," whereas calls (i.e., simple, short, and disordered) are not. Savage et al. and Fitch (2006) highlight repetition as a key design feature of music, distinguishing it from language. However, although animal songs may involve repetition on several hierarchical levels, contact calls do not (Catchpole & Slater, 2008).

The distinction between songs and contact calls allows us to explore whether they differ regarding their relationship to social complexity across vocal learners. Savage et al. state that "melodic, learned song among songbirds, whales, or other vocal learners are predicted to enhance social bonding in these species." However, among cetaceans, baleen whales sing (e.g., humpback whales produce long bouts of complex and hierarchically structured song) to attract mates rather than to enhance social bonds, although spending most of their time solitary or in small groups (Janik, 2014; Whitehead & Rendell, 2012). In contrast, toothed whales have not been reported to sing but instead produce short individually distinctive learned contact calls to maintain social cohesion in their complex social systems (e.g., signature whistles in bottlenose dolphins) (King, Sayigh, Wells, Fellner, & Janik, 2013). Similarly, elephants live in complex fission–fusion societies, produce learned contact calls to maintain individual-specific bonds within changing social groupings, but are not known to sing

(Poole, Tyack, Stoeger-Horwath, & Watwood, 2005). Many parrot species live in large fission–fusion social groups and produce short and simple, individually distinctive contact calls to mediate social and foraging dynamics. And although several parrot species also produce complex duets (e.g., yellow-naped amazon parrots), their primary function is to defend nest sites rather than to support social cohesion within these groups (Bradbury & Balsby, 2016). Finally, hummingbirds live quite solitary lives, but sing sometimes phonologically and syntactically quite complex learned songs to defend territories and attract mates (Araya-Salas et al., 2019; Del Hoyo, Elliot, & Sargatal, 1999).

Hence, our qualitative comparative review of complex vocal learners suggests that learned contact calls enhance social cohesion and, correspondingly, are positively associated with social complexity (Nowicki & Searcy, 2014; Sewall, 2015). In contrast, songs function to attract mates and defend territories and do not appear to be associated with social complexity. This, therefore, counts as evidence against Savage et al.'s social bonding hypothesis as songs share important design features with music whereas learned contact calls do not.

From a signaling perspective, the observed form-function associations of these vocalizations make a lot of sense. The shortness, simplicity, and resulting behaviorally low-costliness of contact calls are consistent with the fact that they are cooperative signals, which do not need to persuade, but merely transparently and efficiently communicate relevant information, such as individual identity and group membership. In contrast, the elaborateness and associated costliness of songs (in terms of production and opportunity costs as well as predation risk, cf. Mehr et al.) correspond to the costly signaling prediction that songs should be costly to advertise and credibly indicate covert qualities (e.g., genetic quality). Hence, because music shares costly design features with animal songs rather than with contact calls, our comparative review lends support to Mehr et al.'s argument against Savage et al.'s social bonding hypothesis that, because a low-cost signaling system such as language can efficiently facilitate social coordination within groups, music would be needlessly costly to fulfill that role.

Instead of enhancing social bonding, Mehr et al. argue that music evolved as a costly, credible signal of covert qualities such as coalition strength and parental attention. Although we concur that music has the design features of a credible signal, plausibly for these functions, we are less convinced by both Mehr et al.'s and Savage et al.'s refutation of the sexual selection hypothesis, namely that music also evolved as a credible signal of mate quality. Their main argument against this hypothesis is the lack of musical sex differences. Yet, this counterargument is weakened by the fact that musical ability, production, and perception may be selected in both sexes under mutual mate choice (because of male parental investment), which is well-established in humans (Stewart-Williams & Thomas, 2013). Indeed, also in many bird species both sexes sing and, for instance, the greater vasa parrot demonstrates that female song can function to attract males that take up a provisioning role (Ekstrom, Burke, Randrianaina, & Birkhead, 2007; Riebel, Odom, Langmore, & Hall, 2019). Hence, although we do not wish to deny potential weaknesses of the sexual selection hypothesis, the lack of sex differences might not necessarily be one of them.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Araya-Salas, M., Smith-Vidaurre, G., Mennill, D. J., González-Gómez, P. L., Cahill, J., & Wright, T. F. (2019). Social group signatures in hummingbird displays provide evidence of co-occurrence of vocal and visual learning. *Proceedings for the Royal Society B*, *286*: 20190666. http://dx.doi.org/10.1098/rspb.2019.0666.

Bradbury, J. W., & Balsby, T. J. S. (2016). The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology*, *70*(3), 293–312. https://doi.org/10.1007/s00265-016-2068-4.

Catchpole, C. K., & Slater, P. J. B. (2008). *Bird song: Biological themes and variations* (2nd ed.). Cambridge University Press.

Del Hoyo, J., Elliot, S. A., & Sargatal, J. (1999). *Handbook of the birds of the world* (5th ed.). Lynx.

Ekstrom, J. M. M., Burke, T., Randrianaina, L., & Birkhead, T. R. (2007). Unusual sex roles in a highly promiscuous parrot: The greater vasa parrot *Caracopsis vasa*. *Ibis*, *149*(2), 313–320. https://doi.org/10.1111/j.1474-919X.2006.00632.x.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*(1), 173–215. https://doi.org/10.1016/j.cognition.2005.11.009.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28*, 60–65. https://doi.org/10.1016/j.conb.2014.06.010.

King, S. L., Sayigh, L. S., Wells, R. S., Fellner, W., & Janik, V. M. (2013). Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1757), 20130053–20130053. https://doi.org/10.1098/rspb.2013.0053.

Marler, P. (2004). Bird calls: A cornucopia for communication. In M. Peter & H. Slabbekoorn (Eds.), *Nature's music: The science of birdsong* (pp. 132–177). Academic Press. https://doi.org/10.1016/B978-012473070-0/50008-6.

Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, *28*, 48–53. https://doi.org/10.1016/j.conb.2014.06.007.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Elephants are capable of vocal learning. *Nature*, *434*(7032), 455–456. https://doi.org/10.1038/434455a.

Riebel, K., Odom, K. J., Langmore, N. E., & Hall, M. L. (2019). New insights from female bird song: Towards an integrated approach to studying male and female communication roles. *Biology Letters*, *15*(4). https://doi.org/10.1098/rsbl.2019.0059.

Sewall, K. B. (2015). Social complexity as a driver of communication and cognition. *Integrative and Comparative Biology*, *55*(3), 384–395. https://doi.org/10.1093/icb/icv064.

Stewart-Williams, S., & Thomas, A. G. (2013). The ape that thought it was a peacock: Does evolutionary psychology exaggerate human sex differences? *Psychological Inquiry*, *24*, 137–168. https://doi.org/10.1080/1047840X.2013.804899.

Whitehead, H., & Rendell, L. (2012). *The cultural lives of whales and dolphins*. University of Chicago Press.

Music as a trait in evolutionary theory: A musicological perspective

Melanie Wald-Fuhrmann, Lara Pearson, Tina Roeske, Christian Grüny and Rainer Polak

Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany.
melanie.wald-fuhrmann@ae.mpg.de
lara.pearson@ae.mpg.de
tina.roeske@ae.mpg.de
christian.grueny@ae.mpg.de
rainer.polak@ae.mpg.de
https://www.aesthetics.mpg.de/en/research/department-of-music.html

doi:10.1017/S0140525X20001193, e93

Abstract

Although it can be straightforward to define the features of physical traits, complex cultural categories tend to elude widely accepted definitions that transcend cultural and historical context. Addressing papers by Mehr et al. and Savage et al., which both aim to explain music as an evolved trait, we discuss fundamental problems that arise from their conceptualizations of music.

When evolutionary theory seeks to describe traits as function-driven adaptations, it is fundamental to correctly identify a trait and define its main features, in order to specify its ultimate functions. For complex and variable cultural phenomena such as "music," this is far from being obvious and strongly dependent on the cultural concepts researchers bring to their study. Across cultures, no common concept of "music/musicality" exists – there is no directly observable, uncontroversial set of identifying features allowing demarcation between practices variously identified as heightened speech, chant and song, and also between dance, gesture, and music. Our comment, therefore, targets the conceptualization of the trait in question from an (ethno)musicological perspective.

The two target articles define the trait in question differently, reflecting contrasting Western concepts of music, conventionally addressed as presentational (in Mehr et al.) versus participatory (in Savage et al.) (Besseler, 1926/2011, 1959; Nettl, 1921; Small, 1998; Turino, 2008). Mehr et al. define music primarily as auditory communication, thus excluding cross-modal aspects of music perception involving vision and proprioception (Phillips-Silver & Trainor, 2005; Vuoskoski, Thompson, Clarke, & Spence, 2014, 2016). This focus weakens their account of group coalition signaling, because the closest modern equivalent – ceremonial cultural performance – fully integrates auditory aspects (music, speech, and sound) with gesture (procession, dance, and theater), and material culture (ornamentation and objects) into a unified mode of display (Brown & Dissanayake, 2018; Schechner, 2013). At a basic level, rhythmic entrainment as an indicator of coalition quality already goes beyond auditory communication as it involves multimodally complex sensorimotor-synchronization (Phillips-Silver & Keller, 2012), the intra-group effects of which are as prominent as its signaling function. In comparison, Savage et al. apply a more multimodal and action/experience-based understanding, which prioritizes intra-group effects over representational communication across groups.

Neither of the author groups substantiates their reasons for preferring one concept of music over the other, and alternative definitions are not considered. This raises the fundamental question of how one can arrive at a description of the relevant trait as the basis of an evolutionary argument. A primordial form of music is clearly no longer observable. Consequently, one has to rely on existing or documented historical forms of music to draw reasoned inferences from, and test evolutionary hypotheses on, with the evident risk of circularity. A method that suggests itself would involve identifying the most prominent features and functions of music-related behaviors around the globe, relying on musicological research. However, efforts toward this have not led to a clear-cut definition (Nettl, 2001; Simon, Riethmüller, & Hüschen, 2016). Furthermore, the results of such a method still depend on prior criteria for inclusion and exclusion, as well as weighting. Authors from both groups have made noteworthy earlier contributions to such comparative endeavors, and both groups use references to existing music to support one argument or reject another. Yet, Savage et al. do this in a more consistent and explicit manner, listing a number of concrete musical design features regarding rhythm, dance, melody, harmony, and structure, which they associate with a number of universal functions in an attempt to find a common denominator. Mehr et al., in contrast, grant the status of ultimate function to only one out of many equally common functions – credible signaling. The selective evidence referred to by Mehr et al. works rather as an illustration than as a demonstration. For example, when they reject the mating quality hypothesis, their concept of music as auditory communication leads them to focus exclusively on songs and the gender distribution of performers (of Western popular music in the last 100 years), but prevents them from considering musical mating-related practices that involve dance (Garfinkel, 2018; Hanna, 2010).

We suggest that any evolutionary scenario should target the set of most commonly observable forms and functions of music-related behaviors. For instance, solitary musicking and musicking for mood regulation and/or pleasure are extremely common and historically observable forms of music-related behavior that are not predicted by any of the proposed evolutionary explanations. Examples of such behavior include playing or singing for oneself (Killick, 2006), and listening to music alone (Herbert, 2011). Mehr et al. do not acknowledge the existence of these forms of musicking, whereas Savage et al. view them as byproducts of social bonding. This may indeed be the case. But equally, considering the widespread existence of solo musicking behaviors, a counter-proposal would be that music-like practices evolved for achieving homeostasis of emotions, feelings, and associated body states (Habibi & Damasio, 2014), and that other functions, such as social bonding, are supported by that.

Finally, seeking to account for the evolution of a human trait places a heavier burden on the concept of "music" than is usually the case, when less is at stake with the use of the term. We propose that the search for the evolution of "music" is flawed because music is a contemporary concept of European heritage without direct equivalent in many other cultures and eras. Although musicologists can certainly point to practices around the world they believe to be similar to what they understand as music, this does not make such practices "music" – at most, it makes them "music-like." It is, therefore, unclear whether "music" is really the evolutionary trait needing to be explained. Two alternatives are possible; the first would be to identify a more inclusive and neutral trait such as "multimodal performance," involving sound and movement, communicative signaling, and participatory experience; the second would be to focus explicitly on narrower traits addressed by the target papers, such as rhythm/entrainment and tonality/melody, without claiming they represent the core of an erroneously universalized notion of music. This involves accepting that what is now commonly considered music may be understood as a composition of diverse behaviors, only recently subsumed under a unified concept in modern discourse. With this in mind, a more plausible approach would be to follow those who argue that what is termed music today is one result of various human socio-cultural inventions (Patel, 2018) that form a multi-stranded, non-linear history (Tomlinson, 2018), which cannot be accounted for by any single (however broadly conceived) adaptivity-based evolutionary explanation.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Besseler, H. (1926). Grundfragen des musikalischen Hörens. *Jahrbuch der Musikbibliothek Peters*, *32*, 35–52. English edition: Besseler, H. (2011) Fundamental issues of musical listening (M. Pritchard, Transl.). *Twentieth-Century Music 8*(1), 49–70. https://doi.org/10.1017/S1478572212000011.

Besseler, H. (1959). *Das musikalische Hören der Neuzeit*. Akademie-Verlag. Republished as: Besseler, H. (1978) Das musikalische Hören der Neuzeit. In P. Gülke (Ed.), *Aufsätze zur Musikästhetik und Musikgeschichte* (pp. 104–173). Philipp Reclam.

Brown, S., & Dissanayake, E. (2018). The synthesis of the arts: From ceremonial ritual to "total work of art." *Frontiers in Sociology*, *3*, 9. https://doi.org/10.3389/fsoc.2018.00009.

Garfinkel, Y. (2018). The evolution of human dance: Courtship, rites of passage, trance, calendrical ceremonies and the professional dancer. *Cambridge Archaeological Journal*, *28*(2), 283–298. https://doi.org/10.1017/S0959774317000865.

Habibi, A., & Damasio, A. (2014). Music, feelings, and the human brain. *Psychomusicology: Music, Mind, and Brain*, *24*(1), 92. http://dx.doi.org/10.1037/pmu0000033.

Hanna, J. L. (2010). Dance and sexuality: Many moves. *Journal of Sex Research*, *47*(2), 212–241. https://doi.org/10.1080/00224491003599744.

Herbert, R. (2011). Reconsidering music and trance: Cross-cultural differences and cross-disciplinary perspectives. *Ethnomusicology Forum*, *20*(2), 201–227. http://dx.doi.org/10.1080/17411912.2011.592402.

Killick, A. (2006). Holicipation: Prolegomenon to an ethnography of solitary music-making. *Ethnomusicology Forum*, *15*(2), 273–299. http://dx.doi.org/10.1080/17411910600915414.

Nettl, B. (2001). Music. In Grove music online. Retrieved 17 Nov., 2020, from https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000040476. https://doi.org/10.1093/gmo/9781561592630.article.40476.

Nettl, P. (1921). Beiträge zur Geschichte der Tanzmusik im 17. Jahrhundert. *Zeitschrift für Musikwissenschaft*, *4*, 257–265.

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. https://doi.org/10.7551/mitpress/10636.003.0009.

Phillips-Silver, J., & Keller, P. E. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Frontiers in Human Neuroscience*, 6, article 26.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430. https://doi.org/10.1126/science.1110922.

Schechner, R. (2013). *Performance studies: An introduction* (3rd ed.). Routledge.

Simon, A., Riethmüller, A., & Hüschen, H. (2016). Musiké – musica – Musik. In L. Lütteken (Ed.), MGG Online. Retrieved 17 Nov., 2020, from https://www.mgg-online.com/mgg/stable/12696.

Small, C. (1998). *Musicking: The meanings of performing and listening*. University Press of New England.

Tomlinson, G. (2018). *A million years of music. The emergence of human modernity*. Zone Books.

Turino, T. (2008). *Music as social life: The politics of participation*. University of Chicago Press.

Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception & Psychophysics*, *76*(2), 591–604. https://doi.org/10.3758/s13414-013-0582-2.

Vuoskoski, J. K., Thompson, M. R., Spence, C., & Clarke, E. F. (2016). Interaction of sight and sound in the perception and experience of musical performance. *Music Perception*, *33*(4), 457–471. https://doi.org/10.1525/mp.2016.33.4.457.

Why language survives as the dominant communication tool: A neurocognitive perspective

Qing Zhang[a] and Edward Ruoyang Shi[b]

[a]Department of Psychology, Sun Yat-Sen University, Guangzhou 510006, China and [b]Department of Catalan Philology and General Linguistics, University of Barcelona, Gran Via de Les Corts Catalanes, 585, 08007 Barcelona, Spain.
zhangqing3@mail.sysu.edu.cn
edwardshiruoyangend@gmail.com;
http://departament-filcat-linguistica.ub.edu/directori-organitzatiu/edward-ruoyang-shi

doi:10.1017/S0140525X20000692, e94

Abstract

By focusing on the contributions of subcortical structures, our commentary suggests that the functions of the hippocampus underlying "displacement," a feature enabling humans to communicate things and situations that are remote in space and time, make language more effective at social bonding. Based on the functions of the basal ganglia and hippocampus, evolutionary trajectory of the subcomponents of music and language in different species will also be discussed.

Savage et al.'s article presents multidisciplinary evidence supporting their music and social bonding hypothesis. The authors emphasize that in the situations where language is less effective, music enhances the social bonding functions. Then, why language outstrips music in many situations in evolution is still a question. We propose that it is because of "displacement," one of the design features of language allowing humans to communicate events beyond here and now. From the perspective of cognitive neuroscience, we focus on the contributions of the subcortical structures of the brain in both music and language. In addition to the domain-general function of the basal ganglia, we propose that the functions of the hippocampus could underlie "displacement" which makes language more effective in general. Furthermore, comparative studies reveal that various subcomponents of music and language have been identified in nonhuman animals, and thus music and social bonding hypothesis cannot explain why only humans have music/language for social bonding or communication.

First, it has been clear that some aspects of music and language have the common neural basis (Brown, 2000). From the clinical perspective, Shi and Zhang (2020) highlight the function of rhythm processing of the cortical-basal ganglia loop for both cognitive domains. To be more specific, we propose that the basal ganglia loop is responsible for transferring hierarchy to linearization in music and language, which is supported by the mechanism of temporal prediction, motor programing, and execution. However, this domain-general function of the basal ganglia loop cannot explain why language succeeded in outstripping music as the main means for communication.

Second, what makes language more effective than music in some situations most likely depends on its feature of displacement, and the hippocampus is proposed to be the neural basis for this property. Displacement is one of the design features of language enabling humans to "talk about things that are remote in space or time (or both) from where the talking goes on" (Hockett, 1960, p. 6). It was assumed the most salient property of human language (Bickerton, 2009). Displacement requires mental time/space travel in mind, which was proposed to depend on episodic memory (Tulving, 1983) and the ability to put oneself in different timescales (Tulving, 2001). Neuroimaging studies have shown that the hippocampus is responsible for episodic memory (e.g., Dickerson & Eichenbaum, 2010; Ergorul & Eichenbaum, 2004). The hippocampus not only binds disparate elements across both space and time, but it can also compare already formed representations with current perceptual input (Olsen, Moses, Riggs, & Ryan, 2012). Covington and Duff (2016) proposed that the shared predictive processing of memory and language is supported by the hippocampus. In the case of language, this predictive processing associates the incoming words and semantic knowledge and builds the interface between episodic memory and communication, thinking of the case of megafauna scavenging of ancient humans. If one member of a group detected a dead deinotherium, he must exchange information, such as where and when he found it, because only by himself he cannot exploit it, he must persuade other members in the group to cooperate. It is this kind of high-end scavenging that distinguishes human ancestors with bone-crunching *garhi* and *habilis*. In this sense, the feature of displacement subserved by the function of the hippocampus enhances the power of language in social bonding.

Third, from the bottom-up perspective of evolutionary biology (De Waal & Ferrari, 2010), analogous or homologous mechanisms implicated in language and music have been found in other animals (Fitch, 2015; Hauser, Chomsky, & Fitch, 2002). Comparative studies have shown that the subcomponents of rhythm processing and episodic-like memory are present in diverse species, which are supported by the basal ganglia and hippocampus. Rhythm processing was proposed to be subdivided into four subcomponents, among which beat perception and synchronization has been detected in vocal learning birds and mammals, and entrainment of conspecific signaling can be found in both vertebrates and invertebrates (Kotz, Ravignani, & Fitch, 2018). The involvement of the basal ganglia circuit in vocal learning in birds (Jarvis, 2007) and rhythm processing in humans (Grahn, 2009) encouraged Patel (2008) to come up with the "vocal learning and rhythm synchronization hypothesis." Damaging the basal ganglia in zebra finches produces stuttering-like songs, a behavior with the disrupted rhythm, resembling stuttering in humans with impaired function of the basal ganglia (Ravignani et al., 2019). With respect to the episodic-like memory, with behavioral criteria "where-what-when," it has been identified in scrub jays (Clayton & Dickinson, 1998), rodents (Crystal & Smith, 2014), and nonhuman primates (Martin-Ordas, Haun, Colmenares, & Call, 2010). Evidence has shown that the hippocampus is involved in episodic memory in mice (Ergorul & Eichenbaum, 2004) and monkeys (Buckley & Gaffan, 2000). Although no direct connection between the hippocampus and episodic memory in birds has been reported, Gould et al. (2013) have found that the avian relative hippocampal size is closely related to food caching, a behavior related to episodic-like memory. Interestingly, the hippocampal size is potentially linked to song plasticity in open-ended vocal learning birds and language learning in human adults (Zhang & Alamri, 2016). It is also worth noting that the basal ganglia and hippocampus are conserved brain structures, and as different species evolve, they may be involved in more advanced cognitive abilities with conserved functions. The identified subcomponents of rhythm processing and episodic-like memory subserved by the basal ganglia and hippocampus above seem to all contribute to social bonding in different species. However, the social bonding theory proposed in the target article cannot explain the evolutionary trajectory of these subcomponents.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Bickerton, D. (2009). *Adam's tongue*. Hill and Wang.

Brown, S. (2000). The "musilanguage" model of musical evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). MIT Press.

Buckley, M. J., & Gaffan, D. (2000). The hippocampus, perirhinal cortex, and memory in the monkey. In J. J. Bolhuis (Ed.), *Brain, perception, and memory: Advances in cognitive neuroscience* (pp. 279–298). Oxford University Press.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*(6699), 272–274. https://doi.org/10.1038/26216.

Covington, N. V., & Duff, M. C. (2016). Expanding the language network: Direct contributions from the hippocampus. *Trends in Cognitive Sciences*, *20*(12), 869–870.

Crystal, J. D., & Smith, A. E. (2014). Binding of episodic memories in the rat. *Current Biology*, *24*(24), 2957–2961. https://doi.org/10.1016/j.cub.2014.10.074.

De Waal F., & Ferrari P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends Cognitive Science*, *14*, 201–207.

Dickerson, B. C., & Eichenbaum, H. (2010). The episodic memory system: Neurocircuitry and disorders. *Neuropsychopharmacology*, *35*(1), 86–104. https://doi.org/10.1038/npp.2009.126.

Ergorul, C., & Eichenbaum, H. (2004). The hippocampus and memory for "what," "where," and "when." *Learning & Memory*, *11*(4), 397–405. https://doi.org/10.1101/lm.73304.

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140091. https://doi.org/10.1098/rstb.2014.0091.

Gould, K. L., Gilbertson, K. E., Hrvol, A. J., Nelson, J. C., Seyfer, A. L., Brantner, R. M., & Kamil, A. C. (2013). Differences in relative hippocampus volume and number of hippocampus neurons among five corvid species. *Brain Behavior and Evolution*, *2013*(81), 56–70.

Grahn, J. A. (2009). The role of the basal ganglia in beat perception: Neuroimaging and neuropsychological investigations. *Annals of the New York Academy of Sciences*, *1169*, 35–45. http://doi.org/10.1111/j.1749-6632.2009.04553.x.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science (New York, N.Y.)*, *298*(5598), 1569–1579. https://doi.org/10.1126/science.298.5598.1569.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*(3), 89–96. https://doi.org/10.1038/scientificamerican0960-88.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology*, *148*(Suppl. 1), S35–44. https://doi.org/10.1007/s10336-007-0243-0.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002.

Martin-Ordas, G., Haun, D., Colmenares, F., & Call, J. (2010). Keeping track of time: Evidence for episodic-like memory in great apes. *Animal Cognition*, *13*(2), 331–340. https://doi.org/10.1007/s10071-009-0282-4.

Olsen, R. K., Moses, S. N., Riggs, L., & Ryan, J. D. (2012). The hippocampus supports multiple cognitive processes through relational binding and comparison. *Frontiers in Human Neuroscience*, 6, 146.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.

Ravignani, A., Dalla Bella, S., Falk, S., Kello, C. T., Noriega, F., & Kotz, S. A. (2019). Rhythm in speech and animal vocalizations: A cross-species perspective. *Annals of the New York Academy of Sciences*, *1453*, 79–98. https://doi.org/10.1111/nyas.14166.

Shi, E. R., & Zhang, Q. (2020). A domain-general perspective on the role of the basal ganglia in language and music: Benefits of music therapy for the treatment of aphasia. *Brain and Language*, *206*, 104811.

Tulving, E. (1983) *Elements of episodic memory*. Oxford University Press.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*(1413), 1505–1515.

Zhang, E. Q., & Alamri, S. (2016). Revisiting the role of the hippocampus in vocal learning. Ducog 2016, Dubrovnik, Croatia.

Music as social bonding: A cross-cultural perspective

Ivan Yifan Zou[a] and William S.-Y. Wang[b]

[a]Department of Music, The University of Hong Kong, 11.03 11/F Run Run Shaw Tower, Pok Fu Lam, Hong Kong and [b]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, GH144, Hung Hom, Kowloon, Hong Kong.
ivan_yfzou@connect.hku.hk;
wsywang@polyu.edu.hk;
https://www.polyu.edu.hk/cbs/web/en/people/staff_detail/4

doi:10.1017/S0140525X20001326, e95

Abstract

We extend Savage et al.'s music and social bonding hypothesis by examining it in the context of Chinese music. First, top-down functions such as music as political instrument should receive more attention. Second, solo performance can serve as important cues for social identity. Third, a right match between the tones in lyrics and music contributes also to social bonding.

Savage et al.'s article makes an impressive attempt to integrate a broad array of theories on music evolution into their overarching music and social bonding (MSB) hypothesis. Here, we propose three extensions which would enrich the MSB hypothesis from a cross-cultural perspective.

First, Savage et al. eloquently demonstrate how the design features of music are well-suited for social bonding, but they do so predominantly from a bottom-up perspective: The adaptive function of music is to meet the basic human needs such as mating, territory advertisement, infant care, and social cohesion. Although such needs are spontaneous and the satisfaction of them allows humans to connect at an unprecedented level, social bonding can manifest itself on an even larger scale when music functions in a top-down manner to address the needs that are planned and intended. One example is the idea of music as a political instrument to maintain the social order.

At about the same time in history, Plato and Confucius began to examine the moral implications of music and to depict their own utopia of music. They both regarded music as intimately linked to virtue: "good music" cultivates self-constrained emotions and noble feelings, whereas "bad music" makes people overindulge in sensual pleasure and therefore leads to moral decadence (Carr, 2006; Liu, 2014). The enormous ethical and educational strength of music makes a perfect tool for the governors to regulate citizens' behaviors, promote the morality, and eventually attain the goal of maintaining the social order.

Perhaps no other country had brought the idea of music as social bonding to the national level and practiced it more rigorously than Imperial China (221 BC–1912 AD) (Brindley, 2012). Throughout the history, *liyue zhidu* 禮樂制度 (the system of ritual and music) lies at the very core of Confucianism and had been heavily relied upon by the imperial courts to reinforce social bonding (Liu, 2014). According to Confucius, music and ritual (or propriety) are not only mutually dependent, but they also complement each other: whereas the former is for the stratification of the society, the latter is for the harmonization and unification. Each social class has its own form of music to appreciate just as each planet revolves around the sun in its own orbit. Social bonding can thus be maximized in such a well-ordered hierarchical society where everyone finds their own position and be content with what they have.

Second, as noted by Savage et al., the design features of music (e.g., music as cue to social identity) are most typically exhibited in group performance and are much less evident in solo performance. We would like to call attention to *guqin* 古琴 music, which serves as a strong cue to group identity but is most typically played and appreciated in solo (Van Gulik, 2011).

Guqin music has long enjoyed the highest prestige in Chinese musical culture owning much to its symbolization of what is considered to be the utmost virtue in Chinese philosophy: *tian ren he yi* 天人合一 (the unity between nature and human) (Lindqvist, 2006; Tien, 2015). A complete mastery of guqin is widely regarded as the foremost accomplishment one can expect from a scholar-official (or literati).

Traditionally, guqin music was circulated exclusively among small groups of intellectuals. Shared expertise on guqin music, therefore, has gradually become a strong cue to the membership of the elite social class (Lai & Mok, 1981). Such social hierarchy built upon the guqin culture is highly rigid. Sometimes, guqin schools function more like a sect, because they not only indoctrinate ways of playing, but also advocate a way of life. Moreover, most guqin masters are extremely cautious when recruiting their disciples. What they value most is not the technical virtuosity but whether the disciples can truly inherit the essential ideology and virtue embraced by their own guqin school. Every guqin school cherishes its own unique style of playing, making the same piece of music varies considerably from one school to the other. Although one can recognize the same piece of music interpreted in different national piano schools (e.g., Austro-German, Russian, and French) with relative ease, it would be difficult to perform the same task across different guqin schools even for the trained ear.

Third, we also note the design features of music more from the perspective of music-language coevolution. The success of singing in tone languages depends crucially on both the tones in lyrics and the melodic contours in music. Tones are mainly used in languages such as Mandarin and Cantonese to assign distinctive lexical meanings to words. They can be characterized by their distinctive pitch contours depending on whether the fundamental frequency sustains, ascends, or descends (Wang, 1973). Because pitch variation is also an essential feature in music, it is, therefore, important to examine the alignment between the tones in lyrics and their associated melodic contours (or the "tone-tune correspondence") (Wee, 2007). Songs with a better tone-tune alignment are perceptually more salient, thus attracting more people to sing and eventually promoting social bonding (Chao, 1956; Pian, 2000). A higher-than-chance tone-tune alignment has been found in a cross-comparison among nine tone languages although some cases of mismatch were also reported (Schellenberg, 2012).

Moreover, music and language could also interact and coevolve in a more implicit way. It has been proposed that the intervals of the chromatic scale by which many cultures create and represent music have evolved to reflect the spectral characteristics inherent in speech. Based on speech data from American English and Mandarin, Ross, Choi, and Purves (2007) showed that justly tuned chromatic intervals are embedded in the vowel formant ratios of speech, although this hypothesis has yet to be tested against many other languages.

Here, we explore the idea of music as social bonding mainly from the perspective of Chinese music. However, given that music is such a human universal, it would be reasonable to examine the issue by taking into account other musical cultures before we can fully accept the validity of the MSB hypothesis.

Financial support. This study was supported in part by grants GRF14611615 and GRF15601718 awarded by the Research Grant Council of Hong Kong.

Conflict of interest. None.

References

Brindley, E. (2012). *Music, cosmology, and the politics of harmony in early China*. State University of New York Press.

Carr, D. (2006). The significance of music for the moral and spiritual cultivation of virtue. *Philosophy of Music Education Review*, *14*(2), 103–117.

Chao, Y. R. (1956). Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese. In M. Halle, H. G. Lunt, H. McLean & C. H. Van Schooneveld (Eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday* (pp. 52–59). Morton.

Lai, T. C., & Mok, R. (1981). *Jade flute: The story of Chinese music*. Swindon Book.

Lindqvist, C. (2006). *Qin*. Albert Bonniers.

Liu, J. (2014). Art and aesthetics of music in classical Confucianism. In V. Shen (Ed.), *Dao companion to classical Confucian philosophy* (pp. 227–244). Springer. https://doi.org/10.1007/978-90-481-2936-2_10.

Pian, R. C. (2000). Tone and tone: Applying musical elements to Chinese words. *Journal of Chinese Linguistics*, *28*(2), 181–200.

Ross, D., Choi, J., & Purves, D. (2007). Musical intervals in speech. *Proceedings of the National Academy of Sciences*, *104*(23), 9852–9857. https://doi.org/10.1073/pnas.0703140104.

Schellenberg, M. (2012). Does language determine music in tone languages? *Ethnomusicology*, *56*(2), 266–278.

Tien, A. (2015). *The semantics of Chinese music: Analysing selected Chinese musical concepts*. John Benjamins.

Van Gulik, R. H. (2011). *Lore of the Chinese lute* (3rd ed.). Orchid Press.

Wang, W. S.-Y. (1973). The Chinese language. *Scientific American*, *228*(2), 50–63.

Wee, L.-H. (2007). Unraveling the relation between Mandarin tones and musical melody. *Journal of Chinese Linguistics*, *35*(1), 128–144.

Sound sleep: Lullabies as a test case for the neurobiological effects of music

Miriam Akkermann[a], Ugur Can Akkaya[a], Cagatay Demirel[b], Dirk Pflüger[c] and Martin Dresler[b]

[a]Institute of Art and Music, TU Dresden, D-01062 Dresden, Germany; [b]Donders Institute, Radboudumc, Nijmegen, Netherlands and [c]Systeme (IPVS), Universität Stuttgart, 70569, Germany.
miriam.akkermann@tu-dresden.de; https://tu-dresden.de/gsw/phil/ikm/das-institut/muwi; u.c.akkaya@gmail.com; caghangir@gmail.com; Dirk.Pflueger@ipvs.uni-stuttgart.de; https://www.ipvs.uni-stuttgart.de/departments/sc/; martin.dresler@donders.ru.nl; https://dreslerlab.org

doi:10.1017/S0140525X20001259, e96

Abstract

Music is part of the cultural practice and, at the same time, is interwoven with biology through its effects on the brain and its likely evolutionary origin. Studies on music, however, are traditionally based on the humanities and often carried out in a purely historical context, without much input from neuroscience and biology. Here, we argue that lullabies are a particularly suited test case to study the biological versus cultural aspects of music.

Music is traditionally studied in historical contexts as part of the cultural practice, but also regarding its social and cultural contexts (see among others, Kramer, 1990, 2001). A fundamental question tapping into diverse fields of research is how music affects the listener. This has been examined from a historic and socio-cultural perspective (e.g., Kassabian, 2013), but gained particular interest in psychology and biology (Prince, 1972). Three distinct underlying dimensions for listening to music have been proposed: to achieve self-awareness, to express social relatedness, and to regulate mood and arousal (Schäfer, Sedlmeier, Städtler, & Huron, 2013). In particular, the latter requires biological investigations, which, however, are still sparse.

A well-known example of the effects music can have on the human mind, behavior, and brain are the soothing effects of lullabies, which are known and used across human cultures and historical ages. In their target articles, Savage et al. and Mehr et al. discuss lullabies for infants as examples for music that especially fosters bonding, showing evidence for both a one-by-one level bonding promoting group cohesion, and evidence that this effect can be found at a cross cultural level. Musical features in lullabies throughout centuries show cross-culturally consistent aspects fostering the feeling of being socially secure, which goes together with soothing effects for the babies, and they even work in the case of unfamiliar foreign lullabies (see e.g., Bainbridge & Bertolo et al. 2021; Sands & Sekaquaptewa, 1978; Spitz, 1979; Trehub, Unyk, & Trainor, 1993). This suggests that musical effects likely involve a strong biological component. In fact, lullabies can be seen as particularly suited to be studied as examples for musical effects on the biological level: their intentional aim is a change in the physiological state of the recipient, from wakefulness to sleep.

Importantly, lullabies work for infants as well as for adults – the market for relaxation/relax music/sounds, sleep (aid) music, or deep/easy/healing sleep music has again come into focus and is commercially growing with the COVID-19 crisis in 2020, being reviewed regularly also in major newspapers (see e.g., Chow, 2020; Times Staff, 2020). From a practical research perspective, the effects of lullabies can thus conveniently be studied in the most easily available research population, namely young adults.

That music can indeed exert "somnogenic" effects is backed up by a body of empirical research. A number of studies have demonstrated therapeutic effects of music in cases of sleep disorders (for a review and meta-analysis see Feng et al., 2018; Jespersen, Otto, Kringelbach, Someren, & Vuust, 2019; Wang, Sun, & Zang, 2014). Also in healthy volunteers, music has been shown to affect sleep quality as assessed by questionnaires (e.g., Chang, Lai, Chen, Hsieh, & Lee, 2012; Field, 1999; Harmat, Takács, & Bódizs, 2007; Johnson, 2003; Lai & Good, 2005; Lamboley, 1998; Trahan, Durrant, Müllensiefen, & Williamson, 2018).

As a more objective measure, several studies in recent years have used polysomnography to test the effects of music on the brain (e.g., Chen et al., 2014; Cordi, Ackermann, & Rasch, 2019; DuRousseau, Mindlin, Insler, & Levin, 2011; Loewy, Hallan, Friedman, & Martinez, 2005). Sleep is a physiologically exceptionally well characterized state in which even subtle changes can be robustly detected, in the absence of motor artifacts that would be difficult to avoid in awake music listeners. Oscillatory patterns in the sleep electroencephalogram (EEG) such as an increase in theta activity and a disappearance of alpha activity can be used as objective markers of the process of falling asleep, whereas slow wave activity can be used as an objectively quantifiable indicator of sleep depth (Chen et al., 2014; Cordi et al., 2019; Lazic & Ogilvie, 2007). Beyond the effects of music before sleep and during transition phases, different kinds of acoustic stimulation are increasingly used during sleep to enhance sleep depth (Ngo, Martinez, Born, & Mölle, 2013) or cognitive functions of sleep (Hu, Cheng, Chiu, & Paller, 2020).

Beyond these advantages of lullabies as test cases for the neurobiological effects of music, two obstacles have to be noted. First, classical music repertoire includes an abundance of lullabies, and the contemporary music market provides an even broader plethora of soothing music. The choice for specific lullaby candidates is thus often based on heuristics rather than systematic evaluation (Cordi et al., 2019; Loewy, 2020; but see Trahan et al., 2018). Second, full polysomnography in the sleep laboratory as the gold standard of sleep research is effortful and costly, thus restricting the scope of a more systematic investigation. However, recent advances in sleep technology and computer science provide promising options to overcome these limitations: A growing market of sleep wearables including convenient sleep EEG headbands allow for the large-scale, low-cost acquisition of longitudinal sleep data with considerable samples sizes (Depner et al., 2020; Scott, Lack, & Lovato, 2020). Resulting big sleep data sets can, in turn, be processed with (semi-)automatic machine-learning-based analysis pipelines (see e.g., sleeptrip.org). These possibilities enable novel interdisciplinary approaches to the study of music: Historically, sociologically, and psychologically informed sets of candidate lullabies can be tested in large populations using wearable sleep recording technology. The resulting large-scale data sets can be analyzed with machine-learning approaches to extract particularly effective somnogenic musical features, which, in turn, can be fed back into traditional musicological analysis. Overall, lullabies can thus be considered an ideal musical genre to elucidate the associations and differences between neurobiological and cultural aspects of music.

Note

Cagatay Demirel's name was misspelled in the original online version of this commentary. This has been corrected here and a corrigendum has been published.

Financial support. This study was supported by *Die Junge Akademie* of the German National Academy of Sciences.

Conflict of interest. None.

References

Bainbridge, C. M., Bertolo, M., Youngers, J., Atwood, S., Yurdum, L., Simson, J., … Mehr, S. A. (2021). Infants relax in response to unfamiliar foreign lullabies. *Nature Human Behaviour*, *5*(2), 256–64. doi: 10.1038/s41562-020-00963-z.

Chang, E. T., Lai, H. L., Chen, P. W., Hsieh, Y. M., & Lee, L. H. (2012). The effects of music on the sleep quality of adults with chronic insomnia using evidence from polysomnographic and self-reported analysis: A randomized control trial. *International Journal of Nursing Studies*, *49*(8), 921–930.

Chen, C. K., Pei, Y. C., Chen, N. H., Huang, L. T., Chou, S. W., Wu, K. P., … Wu, C. K. (2014). Sedative music facilitates deep sleep in young adults. *The Journal of Alternative and Complementary Medicine*, *20*(4), 312–317.

Chow, A. R. (2020). How music could become a crucial part of your sleep hygiene. Time. https://time.com/5844980/sleep-music/.

Cordi, M. J., Ackermann, S., & Rasch, B. (2019). Effects of relaxing music on healthy sleep. *Scientific Reports*, *9*(9079). doi: 10.1038/s41598-019-45608-y.

Depner, C. M., Cheng, P. C., Devine, J. K., Khosla, S., de Zambotti, M., Robillard, R., … Drummond, S. P. A. (2020), Wearable technologies for developing sleep and circadian biomarkers: A summary of workshop discussions. *Sleep*, *43*(2), zsz254. doi: 10.1093/sleep/zsz254.

DuRousseau, D. R., Mindlin, G., Insler, J., & Levin, I. (2011). Operational study to evaluate music-based neurotraining at improving sleep quality, mood and daytime function in a first responder population. *Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience*, *15*(4), 389–398.

Feng, F., Zhang, Y., Hou, J., Jiyai, C., Jiang, Q., Li, X., Zhao, Q., & Li, B.-A. (2018). Can music improve sleep quality in adults with primary insomnia? A systematic review and network meta-analysis. *International Journal of Nursing Studies*, *77*, 189–196. https://doi.org/10.1016/j.ijnurstu.2017.10.011.

Field, T. (1999). Music enhances sleep in preschool children. *Early Child Development and Care*, *150*(1), 65–68.

Harmat, L., Takács, J., & Bódizs, R. (2007). Music improves sleep quality in students. *Journal of Advanced Nursing*, *62*(3), 327–335.

Hu, X., Cheng, L. Y., Chiu, M. H., & Paller, K. A. (2020). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulletin*, *146*(3), 218–244. https://doi.org/10.1037/bul0000223.

Jespersen, K. V., Otto, M., Kringelbach, M., Someren, E. V., & Vuust, P. (2019). A randomized controlled trial of bedtime music for insomnia disorder. *Journal of Sleep Research*, *28*(4), e12817.

Johnson, J. E. (2003). The use of music to promote sleep in older women. *Journal of Community Health Nursing*, *20*(1), 27–35.

Kassabian, A. (2013). *Ubiquitous listening. Affect, attention, and distributed subjectivity*. University of California Press.

Kramer, L. (1990). *Music as cultural practice,* 1800–1900. University of California Press.

Kramer, L. (2001). *Musical meaning: Toward a critical history*. University of California Press.

Lai, H.-L., & Good, M. (2005). Music improves sleep quality in older adults. *Journal of Advanced Nursing*, *49*(3), 234–244.

Lamboley, D. (1998). *Einschlafen und durchschlafen – ohne Medikamente. Wirksame natürliche Methoden, gesunden Schlaf zu finden*. Herder.

Lazic, S. E., & Ogilvie, R. D. (2007). Lack of efficacy of music to improve sleep: A polysomnographic and quantitative EEG analysis. *International Journal of Psychophysiology*, *63*(3), 232–239.

Loewy, J. (2020). Music therapy as a potential intervention for sleep improvement. *Nature and Science of Sleep*, *12*, 1–3. https://doi.org/10.2147/NSS.S194938.

Loewy, J., Hallan, C., Friedman, E., & Martinez, C. (2005). Sleep/sedation in children undergoing EEG testing: A comparison of chloral hydrate and music therapy. *Journal of PeriAnesthesia Nursing*, *20*(5), 323–332.

Ngo, H.-V., Martinez, T., Born, J., & Mölle, M. (2013). Auditory closed-loop stimulation of the sleep slow oscillation enhances memory, *Neuron*, *87*(3), 545–553. doi: 10.1016/j.neuron.2013.03.006.

Prince, W. F. (1972). A paradigm for research on music listening. *Journal of Research in Music Education*, *20*(4), 445–455. https://doi.org/10.2307/3343802.

Sands, K. M., & Sekaquaptewa, E. (1978). Four Hopi lullabies: A study in method and meaning. *American Indian Quarterly*, *4*(3), 195–210. https://doi.org/10.2307/1184620.

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Front Psychology*, *4*(511), 1–33. https://doi.org/10.3389/fpsyg.2013.00511.

Scott, H., Lack, L., & Lovato, N. (2020). A systematic review of the accuracy of sleep wearable devices for estimating sleep onset, *Sleep Medicine Review*, *49*, 101227. doi: 10.1016/j.smrv.2019.101227.

Spitz, S. A. (1979). Social and psychological themes in east Slavic folk lullabies. Slavic and East. *The Slavic and East European Journal*, *23*(1), 14–24. https://doi.org/10.2307/307796.

Times Staff. (2020). The best music to help you fall asleep and combat coronavirus anxiety. Los Angeles Times. https://www.latimes.com/entertainment-arts/music/story/2019-07-25/la-et-ms-sleep-insomnia-best-music.

Trahan, T., Durrant, S. J., Müllensiefen, D., & Williamson, V. J. (2018). The music that helps people sleep and the reasons they believe it works: A mixed methods analysis of online survey reports. *PLoS ONE*, *13*(11), e0206531. https://doi.org/10.1371/journal.pone.0206531.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, *16*(2), 193–211. https://doi.org/10.1016/0163-6383(93)80017-3.

Wang, C. F., Sun, Y. L., & Zang, H. X. (2014). Music therapy improves sleep quality in acute and chronic sleep disorders: A meta-analysis of 10 randomized studies. *International Journal of Nursing Studies*, *51*(1), 51–62.

Progress without exclusion in the search for an evolutionary basis of music

Daniel L. Bowling[a], Marisa Hoeschele[b,c] and Jacob C. Dunn[c,d,e]

[a]Stanford University School of Medicine, Stanford, CA 94306, USA; [b]Acoustic Research Institute, Austrian Academy of Sciences, 1040 Vienna, Austria; [c]Department of Cognitive Biology, University of Vienna, 1090 Vienna, Austria; [d]Behavioral Ecology Research Group, Anglia Ruskin University, Cambridge CB1 1PT, UK and [e]Biological Anthropology, University of Cambridge, Cambridge CB2 1QH, UK.
dbowling@stanford.edu; https://profiles.stanford.edu/daniel-bowling
marisa.hoeschele@oeaw.ac.at; https://tinyurl.com/marisahoeschele
jacob.dunn@aru.ac.uk; https://aru.ac.uk/people/jacob-c-dunn

doi:10.1017/S0140525X20001466, e97

Abstract

Mehr et al.'s hypothesis that the origins of music lie in credible signaling emerges here as a strong contender to explain early adaptive functions of music. Its integration with evolutionary biology and its specificity mark important contributions. However, much of the paper is dedicated to the exclusion of popular alternative hypotheses, which we argue is unjustified and premature.

Human musicality poses a longstanding evolutionary puzzle (Darwin, 1871), and Savage et al. and Mehr et al. provide much needed updates. Their perspectives consolidate and refine ideas from the past two decades of research, marking an important milestone (cf. Brown, Merker, & Wallin, 2000). We focus on Mehr et al., which argues that music's origins lie in credibly signaling coalition quality and parental attention. These adaptive hypotheses are formulated within the well-established framework of signaling theory in evolutionary biology, and build upon comparative evidence for musical behavior in nonhuman animals. We find tremendous value in the breadth and specificity of this work, but weaknesses in its dismissal of alternative hypotheses show that the historical genesis of music remains unclear, if indeed there is a principal one.

Mehr et al. dismiss the music and social bonding (MSB) hypothesis on three counts. The first derives from the premise that primate sociality evolved under predation pressure associated with diurnal foraging. Mehr et al. imply that this ultimate-level pressure renders superfluous any fitness benefits that accrue from variation in group social dynamics. This conflates the selection pressures that drive the evolution of social versus solitary living, with those that drive the evolution of social behavior within a group. We are not aware of any evidence that ties variation in the social group dynamics to differences in fitness, but differences in fitness between groups are self-evident, and we see no reason for assuming that environmental and/or genetic shifts that facilitate social bonding cannot have profound consequences in this context.

Mehr et al.'s second argument against MSB is that it conflates ultimate and proximate levels of explanation by connecting music's function to the neurobiology of social reward (Machin & Dunbar, 2011; Savage et al., target article; Tarr, Launay, & Dunbar, 2014). They correctly point out that music causing social bonding today (assayed behaviorally or neurobiologically) is not evidence that it evolved to do so. This recalls Gould and Lewontin's (1979) critique of adaptationism. Current function tells us little about evolutionary process, particularly in complex aspects of human behavior/cognition (like musicality), where exaptations are expected to comprise "a mountain to the adaptive molehill" (Gould, 1991). However, we object to the implication that this fundamental issue uniquely undercuts MSB. Mehr et al.'s own adaptive hypotheses derive the majority of their empirical support from current functions of musical behavior (in war, intimidation, territoriality, alliance-forging, and infant-directed song). We are all trapped by the present, and inappropriate evidentiary standards for identifying adaptations are not specific to any particular theory (Andrews, Gangestad, & Mathew, 2002; Williams, 1966).

Mehr et al.'s third argument against MSB is that music is poorly designed to coordinate groups. They derive this counterfactual from the notion that language is a superior facilitator of coordinated collective action, offering the example of a coxswain's use of language (rather than music) to coordinate rowing as support. This reflects a dubious imposition of the modern distinction between music and language onto their evolutionary foundations. Language (or its primary behavioral manifestation speech) exists on a continuum with music and many intermediates (public oratory, poetry, rap, chant, etc.). Features that are held in common across this continuum (e.g., auditory-vocal channel is default, highly ordered, infinitely generative, fundamentally social) exceed those which may be considered unique to either pole (e.g., music's spectrotemporal regularity, speech's explicit referentiality). From this perspective, the coxswain's rhythmic calls to "row!" appear more musical than linguistic. Their support to coordination, in particular, seems musical, as temporal regularity characterizes music more than speech (Brown & Jordania, 2013; Dauer, 1983). By contrast, more linguistic features (like the meaning of the word "row") are inessential; a nonsense word or a drum beat (the norm in Chinese dragon boat racing) works just fine. Undoubtedly, speech is superior for coordinating rational thought and planning, but music, and the more musical aspects of speech, clearly support temporal and emotional coordination (Filippi, Hoeschele, Spierings, & Bowling, 2019). The MSB hypothesis is not undone by language.

Finally, Mehr et al. dismiss the mate quality hypothesis (Darwin, 1871). The crux of their argument is that if music evolved via sexual selection in a substantive way, human musicality would be sexually dimorphic, which they argue it is not. There are a number of problems here. One is that it contradicts the author's earlier acknowledgement that current function does not imply original function. Another is that sexual selection does not always produce sexual dimorphism (Darwin, 1871; Hooper & Miller, 2008; Jones & Ratterman, 2009). Another is that sexual selection has almost certainly shaped the evolution of primate loud calls, which Mehr et al. identify as musical precursors (Delgado, 2006; Dunn et al., 2015). But, a more pressing problem is the claim that there are no sex differences in human musicality relevant to this argument. This seems premature given how few studies have addressed the issue directly, particularly when considering the difficulty of separating predisposition from experience at this level (a point which Mehr et al. also acknowledge). The authors' assertion that musical behavior is invariant across the human lifespan is also suspect. Musical preferences emerge as a critical part of self-identity during adolescence, musical performances peaks in young adulthood when courtship is most intense, and musical tastes support strong assortative mating (Miller, 2000; North & Hargreaves, 1999). Finally, it should be noted that humans are more sexually dimorphic in voice frequency than any other ape (Puts et al., 2016). Male and female singing voices fall roughly an octave apart (Titze, 2000), which has potential implications for the esthetics of chorusing (Bowling & Purves, 2015; Hoeschele, 2017).

In sum, we find Mehr et al.'s proposed hypothesis of music evolution to be extremely valuable for its integration with evolutionary biology, breadth, and specificity, but we see no present reason to rule out any of the other hypotheses discussed above as (co-)functional drivers of human musicality.

Financial support. DLB is supported by NIMH grant K01-MH122730-01; JCD is supported by Royal Society grant RSG/R1/180340.

Conflict of interest. None.

References

Andrews, P. W., Gangestad, S. W., & Mathew, D. (2002). Adaptationism – How to carry out an exaptationist program. *Behavioral and Brain Sciences*, *25*(4), 489–504. https://doi.org/10.1017/S0140525X02000092.

Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences*, *112*(36), 11155–11160. https://doi.org/10.1073/pnas.1505768112.

Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, *41*(2), 229–248. https://doi.org/10.1177/0305735611425896.

Brown, S., Merker, B., & Wallin, N. L. (Eds.). (2000). *The origins of music*. MIT Press.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. John Murray.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*(1), 51–62.

Delgado, R. A. (2006). Sexual selection in the loud calls of male primates: Signal content and function. *International Journal of Primatology*, *27*(1), 5–25. https://doi.org/10.1007/s10764-005-9001-4.

Dunn, J. C., Halenar, L. B., Davies, T. G., Cristobal-Azkarate, J., Reby, D., Sykes, D., … Knapp, L. A. (2015). Evolutionary trade-off between vocal tract and testes dimensions in howler monkeys. *Current Biology*, *25*(21), 2839–2844. https://doi.org/10.1016/j.cub.2015.09.029.

Filippi, P., Hoeschele, M., Spierings, M., & Bowling, D. L. (2019). Temporal modulation in speech, music, and animal vocal communication: Evidence of conserved function. *Annals of the New York Academy of Sciences*, *1453*, 99–113. https://doi.org/10.1111/nyas.14228.

Gould, S. J. (1991). Exaptation: A crucial tool for an evolutionary psychology. *Journal of Social Issues*, *47*(3), 43–65. https://doi.org/10.1111/j.1540-4560.1991.tb01822.x.

Gould, S. J., & Lewontin, R. C., (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society*, *205* (1161), 581–598. https://doi.org/10.1098/rspb.1979.0086.

Hoeschele, M. (2017). Animal pitch perception: Melodies and harmonies. *Comparative Cognition & Behavior Reviews*, *12*, 5–18. https://doi.org/10.3819/CCBR.2017.120002.

Hooper, P. L., & Miller, G. F. (2008). Mutual mate choice can drive costly signaling even under perfect monogamy. *Adaptive Behavior*, *16*(1), 53–70. https://doi.org/10.1177/1059712307087283.

Jones, A. G., & Ratterman, N. L. (2009). Mate choice and sexual selection: What have we learned since Darwin? *In the Light of Evolution*, *3*, 169–190. https://doi.org/10.17226/12692.

Machin, A., & Dunbar, R. (2011). The brain opioid theory of social attachment: A review of the evidence. *Behaviour*, *148*(9), 985–1025. https://doi.org/10.1163/000579511X596624.

Miller, G. (2000). Evolution of human music through sexual selection. In S. Brown, B. Merker & C. Wallin (Eds.), *The origins of music* (pp. 329–360). The MIT Press.

North, A. C., & Hargreaves, D. J. (1999). Music and adolescent identity. *Music Education Research*, *1*(1), 75–92. https://doi.org/10.1080/1461380990010107.

Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., … Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1829), 1–7. https://doi.org/10.1098/rspb.2015.2830.

Tarr, B., Launay, J., & Dunbar, R. (2014). Music and social bonding: "self-other" merging and neurohormonal mechanisms. *Frontiers in Psychology*, *5*, 1096.

Titze, I. R. (2000). *Principles of voice production*. National Center for Voice and Speech.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton University Press.

The evolution of music: One trait, many ultimate-level explanations

Edgar Dubourg[a], Jean-Baptiste André[a,b,c] and Nicolas Baumard[a,b,c]

[a]ENS-PSL, 75005 Paris, France; [b]CNRS, 75005 Paris, France and [c]EHESS, 75005 Paris, France.
edgar.dubourg@gmail.com,
jeanbaptisteandre@gmail.com nbaumard@gmail.com

doi:10.1017/S0140525X20001156, e98

Abstract

We propose an approach reconciling the ultimate-level explanations proposed by Savage et al. and Mehr et al. as to why music evolved. We also question the current adaptationist view of culture, which too often fails to disentangle distinct fitness benefits.

Savage et al. focus on the social functions of music-related behaviors, but they don't explain why musicality, and not music, is often preferred to other things such as food, perfume, and painting, to increase social bonding in humans. Similarly, Mehr et al.'s article focuses on the adaptive nature of musicality, which might well be rooted in credible signaling for coalition strength, but they don't expand on how and why music develops new social functions beyond coalition signaling. We argue that both hypotheses need each other to provide a comprehensive and consistent evolutionary understanding of music. We aim at showing why this is the case by disentangling three evolutionary steps.

The first step should account for the fitness benefits of producing music-like sounds in the first place. Mehr et al. provide evidence that contact calls and territorial advertisements may have been the evolutionary precursors of music-related cognitive processes and behavior in humans. Coordinated rhythm, according to Mehr et al., enhanced fitness in local environments when humans evolved because it credibly signaled high level of interindividual coordination. Their hypothesis is seducing because it explains both how music could emerge by natural selection (because it provides a unique way to signal coalition that food or smell cannot provide) and the cognitive constraints such an evolution would have put on all future music-related inventions (they need a least some rhythm). This hypothesis explains why humans evolved cognitive mechanisms to detect and enjoy music-like inputs, at the proximate level. However, it doesn't explain why music evolved culturally with much variability and extend beyond situations of coalitional signaling.

This is the second step. In many species, once a behavioral or phenotypic trait has emerged, it can be co-opted for new adaptive functions. This is very often the case, in particular for traits involved in signaling, as signaling evolves by recycling traits that have first evolved for other functions (Krebs & Dawkins, 1978; Lorenz, 1966). For instance, the female frog *Physalaemus pustulosus* had pre-existing preferences for lower-frequency chuck sounds, and then males evolved the ability to produce such sounds to exploit this sensory preference (Ryan, Fox, Wilczynski, & Rand, 1990). In nonhuman animals, this recycling usually emerges by natural selection.

But, it can also emerge by cultural evolution. Humans are very plastic. Thanks to their cognitive flexibility, they can recycle existing behaviors and preferences and use evolved preferences (e.g., for sugar, sex, social information, and musicality) to shape sophisticated cultural things (e.g., cheesecakes, pornography, stories, and music) that other people enjoy consuming. As many have noted, musicians, singers, and dancers honestly signal skills and qualities through their performances and this leads to sexual, reputational, or material benefits (André, Baumard, & Boyer, 2020; Miller, 2001). Even in modern industrialized societies, musicians and singers take advantage of their productions with economic benefits. We suggest this underlies the producers' motivation to craft such cultural items in the first place.

This second step crucially explains why music appeared in human culture: because (1) humans had evolved with a preference for music-like sounds and (2) people adaptively used this preference to do other things. However, it does not explain why,

Table 1. (Dubourg et al.) Fitness costs and benefits of music-related behavior on both producers and consumers of music, with the framework from social evolution theory (André et al., 2020; Hamilton, 1964)

		Effect on recipients	
		Positive	Negative
Effect on actor	Positive	*Mutualism* (1) Producing artificial signal of coordination to send and receive signals of personal quality (2) Producing artificial signals of coordination to create bonding for mutual benefit (e.g., sport)	*Selfishness* (1) Producing artificial signal of coordination to send and receive false signals of personal quality (2) Producing artificial signals of coordination to create bonding for manipulation (e.g., supermarket)
	Negative	*Altruism* (1) Producing artificial signals of coordination in the form of lullabies to provide social support to children	*Spite*

according to many empirical studies reported by Savage et al., music promotes social bonding.

The third step implies that signaling is not necessarily selfish. At least in small-scale societies, consumers should have fitness benefits too. The most obvious one is the acquisition of the information about the musicians, inferred from their music. A similar phenomenon is well described in the nonhuman animal literature: Peacocks impress peahens with their large and beautiful tails, but the peahens are adaptively drawn to them because the size of the tails honestly signal the genetic quality of the peacocks, leading to a more informative sexual partner choice (Petrie, 1994; Petrie, Tim, & Carolyn, 1991; Zahavi, 1975). Similarly, in small-scale societies, consumers know how skilled producers of music are merely by listening to their music, and they can arguably better choose skilled cooperative agents or mating partners. In large-scale societies, consumers can take advantage of the fact that other people are also attracted by music to signal preferences, skills, and qualities of their own to other people (Bourdieu, 1979; Veblen, 1899).

But. there is another social use of music that relates to the first adaptive function of musicality. Coordinated rhythm evolved to be perceived as a credible signal of coalitional bond. Thus, when we listen to coordinated rhythm outside an agonistic context, we cannot help but analyze this signal as a cue that, somehow, we are part of a well-coordinated coalition. Hence, we feel the pleasure of having social support. This would explain why, by default, music makes people cheerful: it mimics the signal that we have coalitional allies (exactly like pornography is arousing because it mimics the signal that we have an opportunity to reproduce). Humans will thus use music in all cases when they need to artificially create social bonding, with actual fitness consequences (Table 1). We propose that most of the findings reported by Savage et al. are best understood in light of this adaptive recycling.

In sum, we believe evolutionary approaches of cultural items such as music should carefully distinguish fitness costs and benefits from separate evolutionary steps. In that sense, we believe that the publication of these two articles as a pair is an exciting event in the field, if we prevent ourselves from setting them against each other.

Financial support. This study is supported by FrontCog ANR-17-EURE-0017.

Conflict of interest. None.

References

André, J.-B., Baumard, N., & Boyer, P. (2020). *The Mystery of Symbolic Culture: What fitness costs? What fitness benefits?* 18. https://osf.io/kdh7t/download?format=pdf.

Bourdieu, P. (1979). *La distinction: Critique sociale du jugement*. Éditions de Minuit.

Hamilton, W. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, *7*(1), 1–16.

Krebs, J. R., & Dawkins, R. (1978). Animal signals: Mind reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioral ecology: An evolutionary approach* (pp. 282–309), Blackwell Scientific Publications.

Lorenz, K. Z. (1966). Evolution of ritualization in the biological and cultural spheres. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *251*(772), 273–284.

Miller, G. (2011). *The mating mind: How sexual choice shaped the evolution of human nature*. Anchor.

Petrie, M. (1994). Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature*, *371*(6498), 598–599. https://doi.org/10.1038/371598a0.

Petrie, M., Tim, H., & Carolyn, S. (1991). Peahens prefer peacocks with elaborate trains. *Animal Behaviour*, *41*(2), 323–331. https://doi.org/10.1016/S0003-3472(05)80484-1.

Ryan, M. J., Fox, J. H., Wilczynski, W., & Rand, A. S. (1990). Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. *Nature*, *343*(6253), 66–67. https://doi.org/10.1038/343066a0.

Veblen, T. (1899). *Theory of the leisure class: An Economic Study of Institutions*. Macmillan.

Zahavi, A. (1975). Mate selection – A selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214. https://doi.org/10.1016/0022-5193(75)90111-3.

Human evolution of gestural messaging and its critical role in the human development of music

Martin F. Gardiner

Department of Epidemiology, Brown University, Providence, RI 02912, USA.
martin_gardiner@brown.edu

doi:10.1017/S0140525X20001697, e99

Abstract

By fostering bonding (Mehr et al.; Savage et al.), music illustrates marvelously its ability to induce emotional experience. But, music can induce emotion more generally as well. To help explain how music fosters bonding and induces other emotions,

> I propose that music derives this power from the evolution of what I term "gestural messaging."

The subjective experience music provides (Dewey, 1934) is what fosters bonding (Mehr et al.; Savage et al.). Music builds bonding by inducing joint emotional experience, and also, as I now discuss, induces emotional experience more broadly.

Often neglected by psychology and brain research, brain development of subjective awareness and its functional significance is now receiving fresh attention (Damasio, 2010; Dehaene & Changeux, 2011; Edelman & Seth, 2009; Fuster, 2015; Griffin, 2000; Jaynes, 1990; Lamme, 2006; Mashour, Roelfsema, Changeux, & Dehaene, 2020; Nieder, Wagener, & Rinnert, 2020; Sperry, 1983, 1984). Jerison (1973, 1989) has proposed that the nature of human awareness has evolved in ways that as now discussed may have been critical to the evolution of music.

Building from extensive study of fossil evidence, Jerison (1973, 1989) proposes that the current composition of human subjective awareness was created from two periods of dramatic brain evolution. Four hundred million years ago, our vertebrate ancestors escaped heavily visual dinosaur enemies by becoming able to live in the dark. To do so, they evolved new brain design which constructed representation of external world by coordinating and integrating information from all the senses. This evolution developed brain function underlying what we now term sensory awareness. Then, 40–50 million years ago, social hominids through substantial further cortical development evolved verbal communication that separated them from primate ancestors. This communication allowed them to move into safer niches and migrate when necessary in especially challenging environments (Jerison, 1973, 1989; see also Jaynes, 1990; Lieberman, 2006, 2017). I propose that evolution developing verbal communication catalyzed more general development of brain function as well that achieved more general capability for what I term *gestural messaging* which augmented human subjective awareness.

By a gestural message, I refer to a mentally composed act of communication felt to be initiated intentionally even if the mental actions starting the initiation appear to take place outside of conscious awareness (Libet, 1985).

Verbal messaging probably began with individual spoken gestures (Jerison, 1989), but as verbal function developed messages became increasingly composed through sequential combinations of spoken verbal gestures. Then much later (Jaynes, 1990), written verbal gestures were developed. This allowed verbally constructed gestures to pass information widely in historical space and time, and also to be developed more richly.

As we all witness, an essential feature of human evolution of verbal gesturing is that it brings much information received from a verbal message to conscious awareness thus augmenting and interacting usefully with sensory awareness. Jerison (1989) proposes that the need to keep straight these different components of awareness required development of "self" which I interpret as a metaphor for awareness related to executive function (Damasio, 2010; Freeman, 1995; Fuster, 2015; Luria, 1980). Sperry (1984) in his analysis of non-verbal function implies need for such awareness, as well. Notably, our awareness concerning development of verbal messaging (Libet, 1985) allows awareness not of formation, but rather of consequences of intentional actions such as messaging. This, for example, allows a writer to consider, correct, and compose a message even if not yet or even never sent.

With musical messaging, it is emotional reaction both by the sender and receiver to the created message that is of greatest interest. Similar to verbal messaging, musical messaging may have started with single gestures, such as sound made by hitting a rock or tree or vocally that seemed interesting or enjoyable or scary to producer, listener, or both. Once such acts began to be used to generate emotional reactions, reactions to composed single and then multiple simultaneous sequences began to be explored, and finally, with the development of written notation musical messaging could be sent broadly across historical space and time.

Emotional reactions concern reorganizations of physiology of body and brain according to need that are essential to survival (Damasio, 1999). They are largely controlled outside of conscious awareness (Damasio, 1999; Darwin, 1872, 2006) but music takes advantage of evolution of gestural messaging to initiate, develop, and communicate emotional experiences under voluntary control. We become aware of consequences of emotional change by perceiving changes in bodily activity and behavioral capability (Damasio, 1999; James, 1890). The perceived changes in ways being explored are integrated into awareness we term feelings (Carvalho & Damasio, 2021; Damasio & Corvalho, 2013; Habibi & Damasio, 2014).

Darwin (1872, 2006) already discussed in detail the importance of evolved expression of emotion at signaling emotional state of one animal to another. Witnessed effects of emotion on perceived action can accomplish such signaling as well (Gardiner, 2016). A mother calming a crying child by physical actions, speaking soothingly or singing is not signaling, but rather inducing emotion (Juslin, 2016; Juslin & Zentner, 2002). Music's emotional induction appears to depend on the choice and organization of music creating gestures in time (Epstein, 1988; Gardiner, 2012, 2015; Turner & Pöppel, 1988) and the manner of music creating performance (Gardiner, 2012, 2016; Juslin & Lindstrom, 2016). As seen with verbal messaging, the acts of generating the musical messaging bring the nature and consequence of what is produced into awareness. Indeed a musician may respond still more deeply than a listener to a musical message.

General characteristics appear to allow listeners to distinguish with some reliability basic emotions indicated by music (Juslin & Lindstrom, 2016). I have witnessed consistent basic induced effects when playing music to newborns (unpublished data). Music that arouses warriors, builds bonding, or advances love illustrates quite specific emotional induction. As discussed previously (Gardiner, 2015), research by Clynes implies that not only general characteristics, but also very specific details of music messaging gestures as created individually and together are affected by emotion and thus may well influence emotion when heard. Western Classical music, as an example, shows what rich complexity of emotional induction can be created by artful composition of extended musical messaging (Aldrich, 1966; Meyer, 1956; Rattner, 1962, 1980). Our evolved capacity for esthetic judgment, and evaluation of beauty is a type of emotional reaction with great evolutional significance (Darwin, 1871) that can be explored through specific properties of music (Eibl-Eibesfeldt, 1988; Epstein, 1988; Hodges, 2016; Huron, 2016; Levy, 1988; Turner & Pöppel, 1988). Our development of musical messaging opens important windows on ourselves (Gardiner 2020a, 2020b; Gardiner et al., 1996; Winner et al., 2013).

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Aldrich, P. (1966). *Rhythm in seventeenth-century Italian monody*. Norton.

Carvalho, G. B., & Damasio, A. (2021). Interoception and the origin of feelings: A new synthesis. *BioEssays*. 10.1002/bies.202000261.

Damasio, A. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Harcourt.

Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. Pantheon.

Damasio, A., & Corvalho, G. B. (2013). The nature of feelings: Evolutionary and neurobiological origins. *Nature Reviews: Neuroscience*, *14*(2), 143–152.

Darwin, C. (1871). The descent of man and selection in relation to sex. In E. O. Wilson (Ed.), *From so simple a beginning: The four great books of Charles Darwin* (pp. 767–1252). Norton.

Darwin, C. (1872, 2006). The expression of emotion in man and animals. In E. O. Wilson (Ed.), *From so simple a beginning: The four great books of Charles Darwin* (pp. 1255–1478). Norton.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*, 200–227.

Dewey, J. (1934). *Art as experience*. (1989) Southern Illinois Press.

Edelman, B., & Seth, A. K. (2009). Animal consciousness: A synthetic approach. *TINS*, *32* (9), 476–484.

Eibl-Eibesfeldt, I. (1988). The biological foundation of aesthetics. In Rentschler, I., Herzberger, B. and Epstein, D. (Eds.). Beauty and the Brain: Biological aspects of aesthetics (pp 29–69). Springer.

Epstein, D. (1988). Tempo relations in music. In I. Rentschler, B. Herzberger, & D. Epstein (Eds.), *Beauty and the brain: Biological aspects of aesthetics* (pp. 91–116). Springer.

Freeman, W. G. (1995). *Societies of brains*. Erlbaum, N.Y. (2009) Psychology Press.

Fuster, J. M. (2015). *The prefrontal cortex* (5th ed.). Elsevier.

Gardiner, M. F. (2012). Emotional participation in musical and non-musical behaviors. *Behavioral and Brain Sciences*, *35*(3), 149–150.

Gardiner, M. F. (2015). Integration of cognition and emotion in physical and mental actions in musical and other behaviors. *Behavioral and Brain Sciences*, *38*, 24–25.

Gardiner, M. F. (2016). Modulation of behavior in communicating emotion. *Animal Sentience*, *4*(4), 1–5.

Gardiner, M. F. (2020a). Chapter 32: Relationships between intrinsic and broader educational benefits of singing training. In *Routledge companion to interdisciplinary studies in singing* (Vol. *2*, pp. 529–539). Education.

Gardiner, M. F. (2020b). Academic skill learning and the problem of complexity: II: Underdeveloped mental action, its role in students who struggle as their formal education begins, and evidence from Kodàly music training for opportunity to address this problem. *International Journal for Complexity in Education*, *1*(2), 185–207.

Gardiner, M. F., Fox, A., Knowles, F., & Jeffrey, D. (1996). Learning improved by arts training. *Nature*, *381*, 284.

Griffin, D. R. (2000). Scientific approaches to animal consciousness.*American Anthropologist*, *4*, 889–892.

Habibi, A., & Damasio, A. (2014). Music, feelings, and the human brain. *Psychomusicology: Music, Mind and Brain*, *24*(1), 92–102.

Hodges, D. A. (2016). The neuroaesthetics of music. In S. Hallam, I. Cross and M. Thaut (Eds.), *The Oxford handbook of music Psychology* (2nd ed., pp. 247–262). Oxford University Press.

Huron, D. (2016). Aesthetics. In S. Hallam, I. Cross and M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 233–246). Oxford University Press.

James, W. (1890). *The principles of psychology*. Dover, 1950.

Jaynes, J. (1990). *The origin of consciousness in the breakdown of the bicameral mind*. Houghton Mifflin.

Jerison, H. J. (1973). *Evolution of the brain and intelligence*. Academic Press.

Jerison, H. J. (1989). Brain size and the evolution of mind. *59th James Arthur Lecture on the evolution of the human brain. American museum of Natural History*. Retrieved from H. J. Jerison web site, UCLA, Los Angeles, CA.

Juslin, P. (2016). Emotional reactions to music. In S. Hallam, I. Cross & M. Thaut (Eds.), *The Oxford handbook of music Psychology* (2nd ed., pp. 197–214). Oxford University Press.

Juslin, P., & E. Lindstrom. (2016). Emotion in music performance. In S. Hallam, I. Cross & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 597–614). Oxford University Press.

Juslin, P., & Zentner, M. R. (2002). Current trends in the study of music and emotion: Overture. *Musicae scientiae*. Special issue, 2001–2002, 3–21.

Lamme, V. A. F. (2006). Towards a neural stance on consciousness. *TIGS*, *10*(11), 494–501.

Levy, J. (1988). Cerebral asymmetry and aesthetic experience. In I. Rentschler, B. Herzberger & D. Epstein (Eds.), *Beauty and the brain: Biological aspects of aesthetics* (pp. 219–242). Springer.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*(4), 529–539.

Lieberman, P. (2006). *Toward an evolutionary biology of language*. Belknap Press of Harvard University Press.

Lieberman, P. (2017). *The theory that changed everything: "On the origin of species" as a work in progress*. Columbia University Press.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). Consultants Bureau.

Mashour, G. A., Roelsfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious processing and the neuronal workspace hypothesis. *Neuron*, *105*, 776–798.

Meyer, L. B. (1956). *Emotion and the meaning of music*. Univ. Chicago Press.

Nieder, A., Wagener, L., & Rinnert, P. (2020). A neural correlate of sensory consciousness in a corvid bird. *Science (New York, N.Y.)*, *369*(6511), 1626–1629.

Rattner, L. G. (1962). *Harmony, structure and style*. McGraw-Hill.

Rattner, L. G. (1980). *Classic music: Expression, form and style*. Schirmer.

Sperry, R. (1983). *Science and moral priority: Merging mind, brain and human values*. Columbia University Press.

Sperry, R. (1984). Consciousness, personal identity and the divided brain. *Neuropsychologia*, *22*(6), 661–673.

Turner, F., & Pöppel, E. (1988). Metered poetry, the brain and time. In I. Rentschler, B. Herzberger & D. Epstein (Eds.), *Beauty and the brain: Biological aspects of aesthetics* (pp. 71–90). Springer.

Winner, E., Goldstein, T. R., & Vincent-Lancrin, S. (2013). *Art for art's sake? The impact of arts education*. OECD Publishing.

Why musical hierarchies?

Courtney B. Hilton[a,b], Rie Asano[c] and Cedric Boeckx[d]

[a]Department of Psychology, Harvard University, Cambridge, MA 02138, USA; [b]School of Psychology, The University of Sydney, Camperdown, NSW 2006, Australia; [c]Institute of Musicology, Systematic Musicology, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany and [d]ICREA (Catalan Institute for Advanced Studies and Research)/Universitat de Barcelona, Section of General Linguistics/Universitat de Barcelona Institute for Complex Systems (UBICS), Gran Via de les Corts Catalanes, 585, 08007 Barcelona, Spain.
courtneyhilton@g.harvard.edu
rie.asano@uni-koeln.de
cedric.boeckx@ub.edu

doi:10.1017/S0140525X20001338, e100

Abstract

Credible signaling may have provided a selection pressure for producing and discriminating increasingly elaborate proto-musical signals. But, why evolve them to have hierarchical structure? We argue that the hierarchality of tonality and meter is a byproduct of domain-general mechanisms evolved for reasons other than credible signaling.

The target article by Mehr and colleagues provides a welcome critique of prevailing evolutionary theories of music while also advancing their own credible signaling proposal. We find many aspects of this promising. However, although adaptations for rhythm and melody seem plausible, we take issue with the claim that credible signaling resulted in a "grammar-like, combinatorially generative interface" based on the "hierarchical organization of meter and tonality."

Coalition signaling provides plausible reasons to evolve the capacity to produce and discriminate rhythmically coordinated displays. This is supported in the cited data on birds (Hall & Magrath, 2007; Tobias et al., 2016) and primates (Geissmann, 2000). But, it does not, as far as we can see, provide reasons to evolve hierarchical means of doing so. And indeed, these data only show evidence of rhythmic coordination in terms of temporal precision or synchronization and provide no evidence for or against hierarchy.

Similarly, parent–infant signaling provides evolutionary reasons for melodic signals and infant sensitivity to them. A comparison of animal contact calls (Bouchet, Blois-Heulin, & Lemasson, 2013; Leighton, 2017) and data on genomic imprinting disorders in humans (Mehr, Kotler, Howard, Haig, & Krasnow, 2017) supports these claims. But, here too: why hierarchies? For the purpose of signaling attention to an infant, or for contact calls more generally, hierarchical organization poses no obvious advantage. There is also limited evidence for contact calls being hierarchically organized. Moreover, although some brain areas show differential responses to tonal structure from birth (Perani et al., 2010), behavioral sensitivity only begins to manifest at around 4 years of age before continuing to develop into the teenage years (Brandt, Gebrian, & Slevc, 2012; Corrigall & Trainor, 2014).

Taken together, although the hierarchical properties of meter and tonality are a design feature of the musical capacity, their presence is not so clearly motivated by credible signaling.

Hierarchies, however, are not unique to music. They are found in other cognitive domains such as language (Chomsky, 1957), vision (Bill, Pailian, Gershman, & Drugowitsch, 2020), metacognition (Frith, 2012), and action planning (Miller, Galanter, & Pibram, 1960). In nonhuman primates, they are found in social learning (Byrne & Russon, 1998) and tool use (Byrne, Sanz, & Morgan, 2013; Greenfield, 1991). Musical hierarchality may, therefore, be better conceived as using generic mechanisms evolved for reasons other than as a specific adaptation for credible signaling.

We have previously argued that the hierarchality of both musical and linguistic structures derives from mechanisms originally evolved for action planning (Asano & Boeckx, 2015; see also: Fitch & Martins, 2014; Jackendoff, 2009). The inspiration for much of this thinking was Karl Lashley's (1951) prescient insight that complex actions generally, and those for music and language specifically, control their sequential manifestation through hierarchical plans. Doing so, he argued, was necessary for flexibility and robustness, especially for more complex and abstractly motivated actions in which the limitations of control by linear associative chaining are laid bare.

The primary neurocognitive mechanism underlying this capacity is hierarchical cognitive control and comprises a combination of executive functions (maintenance, selection, and inhibition). Maintenance is subserved by prefrontal areas (together with their parietal connections) and selection and inhibition by the basal ganglia. The orchestration of these functional areas through a number of distinct cortico-basal ganglia-thalamocortical circuits enables complex and flexible behavior (Badre & Nee, 2018). Consistent with Lashley's insight, these neural circuits are not only implicated in action planning, but also for processing musical and linguistic hierarchies (Asano et al., 2021; Fitch & Martins, 2014; Jeon, Anwander, & Friederici, 2014; Slevc & Okada, 2015).

Functional explanations of behavior are essential for understanding biological evolution. But, based on these alone, the determination of how they are translated into mechanisms is too underconstrained. Are new mechanisms evolved *de novo*? Or are existing ones tweaked and put to new use? And then how may these and other proximate mechanisms in turn constrain the space of ultimate reasons that guides selection in a reciprocal cycle? (Laland, Sterelny, Odling-Smee, Hoppitt, & Uller, 2011). As Tinbergen (1963) suggested, the biological study of behavior (and cognitive systems, in the current paper) should give equal attention to each of four questions: mechanism, ontogeny, phylogeny, and function. Each provides unique constraints whose combined consilience is the basis for robust theory.

One notable "so what?" of all this for the target article is that adaptations for credible signaling may also have implications for language. According to our proposal, the structural complexity of both music and language partly derives from generic hierarchical cognitive control mechanisms that interface with auditory and motor systems. Compared to nonhuman primates, humans have substantially greater white-matter connectivity both within the hierarchical control circuits and through the dorsal auditory pathway that links motor, auditory, and parietal areas with the prefrontal cortex (Barrett et al., 2020; Rilling et al., 2008). Adaptations for producing and perceiving rhythmically coordinated audio-motor displays and for fine-scale vocal control of pitch conceivably include an expansion of this shared connectome (Merchant & Honing, 2014; Patel & Iversen, 2014), thus entangling the evolution of both domains. This would also be consistent with claims about music-to-language transfer effects more generally in ontogeny (Patel, 2011; Zatorre, 2013).

To conclude, the credible signaling proposal of Mehr and colleagues is commendable. But, we suggest that it can be further improved by considering interactions of proximate and ultimate causes, and specifically how this may clarify the origins of musical hierarchies.

Note. The authors contributed equally to this commentary.

Financial support. This study was supported by MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas #4903 (Evolinguistics) [grant number JP17H06379] and the Spanish Ministry of Science and Innovation [grant number PID2019-107042GB-I00].

Conflict of interest. None.

References

Asano, R., & Boeckx, C. (2015). Syntax in language and music: What is the right level of comparison? *Frontiers in Psychology*, *6*, 942. https://doi.org/10.3389/fpsyg.2015.00942.

Asano, R., Boeckx, C., & Seifert, U. (2021). Hierarchical control as a shared neurocognitive mechanism for language and music. *PsyArXiv*. https://doi.org/10.31234/osf.io/25qha.

Badre, D., & Nee, D. E. (2018). Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences*, *22*(2), 170–188.

Barrett, R. L. C., Dawson, M., Dyrby, T. B., Krug, K., Ptito, M., D'Arceuil, H., … Catani, M. (2020). Differences in frontal network anatomy across primate species. *The Journal of Neuroscience*, *40*(10), 2094–2107.

Bill, J., Pailian, H., Gershman, S. J., & Drugowitsch, J. (2020). Hierarchical structure is employed by humans during visual motion perception. *Proceedings of the National Academy of Sciences*, *117*(39), 24581–24589. https://doi.org/10.1073/pnas.2008961117.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, *4*, 390. https://doi.org/10.3389/fpsyg.2013.00390.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*, *3*, 327. https://doi.org/10.3389/fpsyg.2012.00327.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, *21*(5), 667–684.

Byrne, R. W., Sanz, C. M., & Morgan, D. B. (2013). Chimpanzees plan their tool use. In C. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals* (pp. 48–64). Cambridge University Press.

Chomsky, N. (1957). *Syntactic structures*. Walter de Gruyter.

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children: Evidence from behavioral and electrophysiological methods. *Developmental Science*, *17*(1), 142–158.

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316*(1), 87–104.

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2213–2223.

Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 103–123). MIT Press.

Greenfield, P. (1991). Language, tools and the brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, *14*(4), 570–571.

Hall, M. L., & Magrath, R. D. (2007). Temporal coordination signals coalition quality. *Current Biology*, *17*(11), R406–R407.

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception*, *26*(3), 195–204.

Jeon, H.-A., Anwander, A., & Friederici, A. D. (2014). Functional network mirrored in the prefrontal cortex, caudate nucleus, and thalamus: High-resolution functional imaging and structural connectivity. *Journal of Neuroscience*, *34*(28), 9202–9212.

Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and effect in biology revisited: Is Mayr's proximate-ultimate dichotomy still useful? *Science (New York, N.Y.)*, *334*(6062), 1512–1516.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon Symposium* (pp. 112–147). Wiley.

Leighton, G. M. (2017). Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proceedings of the Royal Society B: Biological Sciences*, *284* (1868), 20171508.

Mehr, S. A., Kotler, J., Howard, R. M., Haig, D., & Krasnow, M. M. (2017). Genomic imprinting is implicated in the psychology of music. *Psychological Science*, *28*(10), 1455–1467. https://doi.org/10.1177/0956797617711456.

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*, 274. https://doi.org/10.3389/fnins.2013.00274.

Miller, G., Galanter, E., & Pibram, K. (1960). *Plans and structure of behavior*. Henry Holt and Company.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, 57.

Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., … Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, *107*(10), 4758–4763.

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, *11*(4), 426–428.

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, *22*(3), 637–652.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*, 410–433.

Tobias, J. A., Sheard, C., Seddon, N., Meade, A., Cotton, A. J., & Nakagawa, S. (2016). Territoriality, social bonds, and the evolution of communal signaling in birds. *Frontiers in Ecology and Evolution*, *4*, 74. https://doi.org/10.3389/fevo.2016.00074.

Zatorre, R. J. (2013). Predispositions and plasticity in music and speech learning: Neural correlates and implications. *Science (New York, N.Y.)*, *342*(6158), 585–589.

A neurodevelopmental disorders perspective into music, social attention, and social bonding

Anna Kasdan[a], Reyna L. Gordon[a,b] and Miriam D. Lense[a,b]

[a]Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN 37232, USA and [b]Department of Otolaryngology – Head and Neck Surgery, Vanderbilt University Medical Center, Nashville, TN 37232, USA.
anna.v.kasdan@vanderbilt.edu; reyna.gordon@vanderbilt.edu; miriam.lense@vanderbilt.edu
annakasdan.com; https://www.vumc.org/music-cognition-lab/person/faculty

doi:10.1017/S0140525X20001302, e101

Abstract

Our commentary addresses how two neurodevelopmental disorders, Williams syndrome and autism spectrum disorder, provide novel insights into the credible signaling and music and social bonding hypotheses presented in the two target articles. We suggest that these neurodevelopmental disorders, characterized by atypical social communication, allow us to test hypotheses about music, social bonding, and their underlying neurobiology.

Neurodevelopmental disorders characterized by atypical social profiles provide a unique window into the relationship between sociability, music, and the evolutionary origins of music. Here, we focus on two neurodevelopmental disorders characterized by atypical development of social communication – Williams syndrome and autism spectrum disorder (ASD). We (1) provide additional consideration about music as a credible signal for parent–infant interactions (Mehr, Krasnow, Bryant, & Hagen, target article) and (2) extend neurobiological predictions about music and social bonding (MSB; Savage et al., target article).

Individuals with Williams syndrome, a genetic neurodevelopmental disorder caused by the deletion of ~28 genes on chromosome 7, exhibit marked hypersociability accompanied by difficulties in social pragmatics (Barak & Feng, 2016). Individuals with ASD, an etiological heterogenous set of neurodevelopmental disorders, exhibit deficits in social communication and interactions (Barak & Feng, 2016). Music is a relative strength and interest in both Williams syndrome and ASD. Individuals with Williams syndrome exhibit increased emotional responsiveness to music and increased neural responses to musical stimuli compared to typical peers (Kasdan, Gordon, & Lense, 2020; Lense, Gordon, Key, & Dykens, 2014), although behavioral studies of musical skills show mixed findings and substantial individual differences (Hopyan, Dennis, Weksberg, & Cytrynbaum, 2001; Martens, Reutens, & Wilson, 2010). Individuals with ASD generally exhibit age-appropriate melody and rhythm skills (Jamey et al., 2019; Tryfon et al., 2017) and better pitch interval detection compared to controls (Heaton, 2005). Additionally, individuals with ASD show preserved emotional processing of music at both behavioral and neural levels, despite overall impairments in emotion recognition (Caria, Venuti, & De Falco, 2011; Molnar-Szakacs & Heaton, 2012).

As in typical development, music is used in these populations to leverage social attention and bonding, including during parent–child interactions (Steinberg, Shivers, Liu, Cirelli, & Lense, 2020). Because of impaired social communication of children with ASD and Williams syndrome, parents adapt and increase their social signaling in order to capture and modulate their children's attention and emotions; this frequently involves adapting the "musicality" of the social signal such as through increased rhythmic predictability and pitch contours (e.g., increased use of infant-directed speech; Cohen et al., 2013; Quigley, McNally, & Lawson, 2016). Music and song activities provide an ecologically valid means for scaffolding parental signaling during these shared social interactions (Lense & Camarata, 2020). This is consistent with and extends the idea of music as a credible signal for parent–infant interactions (Mehr et al.). The core features of music (e.g., discrete pitches and predictable and hierarchical rhythms) incorporated into both credible signaling and MSB hypotheses can predictably, reliably, and efficiently structure social rhythmic signals such as vocalizations, movement, gaze, and touch (Lense & Camarata, 2020).

The two hypotheses espoused for the evolution of musicality, which highlight the social attention components of musical activities, also inform directions regarding the therapeutic use of music in ASD and Williams syndrome. Investigations of MSB must consider the bidirectionality of the interaction and the needs and goals of both social partners. For example, musical activities between parents and their child with a neurodevelopmental disability are increasingly recognized as modulating *parent* behaviors such as parent responsiveness (e.g., physically supporting child's play through contingent imitation) (Boorom, Muñoz, Xin, Watson, & Lense, 2020; Thompson, Shanahan, & Gordon, 2019) and parent mood (Williams, Berthelsen, Nicholson, Walker, & Abad, 2012), thus enhancing the likelihood of child engagement downstream. Music and song may be a particularly meaningful platform for social communication in light of the altered language and social communication abilities in ASD and Williams syndrome (Mervis & Velleman, 2011; Tager-Flusberg, 2000). Consistent with tenets of the MSB hypothesis for the efficacy of music in large groups, music-based activities may scale to larger interaction contexts beyond the parent–child dyad, and to new interaction partners, because of the effective and rewarding aspects of musical activities for participants (Lense & Camarata, 2020).

Neurobiologically, individuals with Williams syndrome and ASD exhibit brain structural and functional differences in regions important for rhythm processing (e.g., basal ganglia and somatomotor connectivity; Campbell et al., 2009; Estes et al., 2011; Vega, Hohman, Pryweller, Dykens, & Thornton-Wells, 2015). These differences in brain structure and connectivity may be functionally linked to social and musical profiles. Children with ASD who underwent a music intervention exhibited higher communication scores and functional connectivity between auditory and motor regions (cortical and subcortical) compared to a non-music intervention control group following the intervention (Sharda et al., 2018). In addition, individuals with ASD exhibit increased functional connectivity between left frontal and temporal regions for song compared to speech (Lai, Pantazatos, Schneider, & Hirsch, 2012). Different neural mechanisms for processing music and language may provide a basis for using music-based activities in these populations (Sharda, Midha, Malik, Mukerji, & Singh, 2015). Additionally, behavioral work suggests links between beat perception and adaptive communication abilities in Williams syndrome (Lense & Dykens, 2016) and appropriate sensorimotor rhythmic synchronization under some circumstances in ASD (Tryfon et al., 2017). These behavioral studies further support investigation into if and how musical rhythm activities may modulate neural connectivity between sensorimotor and reward related regions in these populations. The neuroanatomical hubs of the MSB hypothesis (e.g., basal ganglia, implicated in both the reward system and in processing "beat-based" rhythms; Matthews, Witek, Lund, Vuust, & Penhune, 2020) provide a relevant framework for assessing how these regions may relate to music's social bonding function in individuals with atypical social communication.

Brain-to-brain synchrony (i.e., "neural resonance" component of the MSB hypothesis) is an important neurobiological mechanism for successful communication (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012; Nguyen et al., 2020). When considering music as a credible signal for parent–infant interactions, social interactions emphasizing musical elements (e.g., song) may strengthen neural synchrony across the dyad. This may be particularly effective for individuals with Williams syndrome or ASD for whom music may be a more salient cue for social interactions. Predictable musical experiences paired with salient social signals (e.g., eye gaze; Leong et al., 2017) may facilitate rewarding interactions between dyads. Populations with neurodevelopmental disorders may offer a unique lens into how brain-to-brain synchrony in parent–child dyads is modulated by musical and non-musical activities. Both the credible signaling (Mehr et al., target article) and MSB (Savage et al., target article) hypotheses provide relevant frameworks for understanding the neurobiology underlying musicality and social communication in Williams syndrome and ASD.

Financial support. This study was supported in part by the National Institute of Mental Health (NIMH), National Center for Complementary and Integrative Health (NCCIH) [R61MH123029], National Institute on Deafness and Other Communication Disorders (NIDCD) [R21DC016710], the National Institutes of Health Office of the Director (DP2HD098859), National Endowment for the Arts Research Lab (1863278-38), and the National Science Foundation (NSF) [1926794] and NSF GRFP. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

Conflict of interest. None.

References

Barak, B., & Feng, G. (2016). Neurobiology of social behavior abnormalities in autism and Williams syndrome. *Nature Neuroscience*, *19*(6), 647–655. https://doi.org/10.1038/nn.4276.Neurobiology.

Boorom, O., Muñoz, V., Xin, R., Watson, M., & Lense, M. D. (2020). Parental responsiveness during musical and non-musical engagement in preschoolers with ASD. *Research in Autism Spectrum Disorders*, *78*. https://doi.org/10.1016/j.rasd.2020.101641.

Campbell, L. E., Daly, E., Toal, F., Stevens, A., Azuma, R., Karmiloff-Smith, A., … Murphy, K. C. (2009). Brain structural differences associated with the behavioural phenotype in children with Williams syndrome. *Brain Research*, *1258*, 96–107. https://doi.org/10.1016/j.brainres.2008.11.101.

Caria, A., Venuti, P., & De Falco, S. (2011). Functional and dysfunctional brain circuits underlying emotional processing of music in autism spectrum disorders. *Cerebral Cortex*, *21*(12), 2838–2849. https://doi.org/10.1093/cercor/bhr084.

Cohen, D., Cassel, R. S., Saint-Georges, C., Mahdhaoui, A., Laznik, M. C., Apicella, F., … Chetouani, M. (2013). Do parentese prosody and fathers' involvement in interacting facilitate social interaction in infants who later develop autism? *PLoS ONE*, *8*(5), 1–10. https://doi.org/10.1371/journal.pone.0061402.

Estes, A., Shaw, D. W. W., Sparks, B. F., Friedman, S., Giedd, J. N., Dawson, G., … Dager, S. R. (2011). Basal ganglia morphometry and repetitive behavior in young children with autism spectrum disorder. *Autism Research*, *4*(3), 212–220. https://doi.org/10.1002/aur.193.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121. https://doi.org/10.1016/j.tics.2011.12.007.

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders*, *35*(6), 787–793. https://doi.org/10.1007/s10803-005-0024-7.

Hopyan, T., Dennis, M., Weksberg, R., & Cytrynbaum, C. (2001). Music skills and the expressive interpretation of music in children with Williams-Beuren syndrome: Pitch, rhythm, melodic imagery, phrasing, and musical affect. *Child Neuropsychology*, *7*(1), 42–53. https://doi.org/10.1076/chin.7.1.42.3147.

Jamey, K., Foster, N. E. V., Sharda, M., Tuerk, C., Nadig, A., & Hyde, K. L. (2019). Evidence for intact melodic and rhythmic perception in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *64*, 1–12. https://doi.org/10.1016/j.rasd.2018.11.013.

Kasdan, A., Gordon, R. L., & Lense, M. D. (2020). Neurophysiological correlates of dynamic beat tracking in individuals with Williams syndrome. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2020.10.003.

Lai, G., Pantazatos, S. P., Schneider, H., & Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain*, *135*(3), 961–975. https://doi.org/10.1093/brain/awr335.

Lense, M. D., & Camarata, S. (2020). PRESS-play: Musical engagement as a motivating platform for social interaction and social play in young children with ASD. *Music & Science*, *3*, 205920432093308. https://doi.org/10.1177/2059204320933080.

Lense, M. D., & Dykens, E. M. (2016). Beat perception and sociability: Evidence from Williams syndrome. *Frontiers in Psychology*, *7*, 1–13. https://doi.org/10.3389/fpsyg.2016.00886.

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. https://doi.org/10.1093/scan/nst017.

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(50), 13290–13295. https://doi.org/10.1073/pnas.1702493114.

Martens, M. A., Reutens, D. C., & Wilson, S. J. (2010). Auditory cortical volumes and musical ability in Williams syndrome. *Neuropsychologia*, *48*(9), 2602–2609. https://doi.org/10.1016/j.neuropsychologia.2010.05.007.

Matthews, T. E., Witek, M. A. G., Lund, T., Vuust, P., & Penhune, V. B. (2020). The sensation of groove engages motor and reward networks. *NeuroImage*, *214*, 116768. https://doi.org/10.1016/j.neuroimage.2020.116768.

Mervis, C. B., & Velleman, S. L. (2011). Children with Williams syndrome: Language, cognitive, and behavioral characteristics and their implications for intervention. *Perspectives on Language Learning and Education*, *18*(3), 98–107. https://doi.org/10.1044/lle18.3.98.

Molnar-Szakacs, I., & Heaton, P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences*, *1252*(1), 318–324. https://doi.org/10.1111/j.1749-6632.2012.06465.x.

Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2020). Neural synchrony in mother–child conversation: Exploring the role of conversation patterns. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 93–102. https://doi.org/10.1093/scan/nsaa079.

Quigley, J., McNally, S., & Lawson, S. (2016). Prosodic patterns in interaction of low-risk and at-risk-of-autism spectrum disorders infants and their mothers at 12 and 18 months. *Language Learning and Development*, *12*(3), 295–310.

Sharda, M., Midha, R., Malik, S., Mukerji, S., & Singh, N. C. (2015). Fronto-temporal connectivity is preserved during sung but not spoken word listening, across the autism spectrum. *Autism Research*, *8*(2), 174–186. https://doi.org/10.1002/aur.1437.

Sharda, M., Tuerk, C., Chowdhury, R., Jamey, K., Foster, N., Custo-Blanch, M., … Hyde, K. (2018). Music improves social communication and auditory-motor connectivity in children with autism. *Translational Psychiatry*, *8*(1), 231. https://doi.org/10.1038/s41398-018-0287-3.

Steinberg, S., Shivers, C. M., Liu, T., Cirelli, L. K., & Lense, M. D. (2020). Musical Engagement and Parent-Child Attachment in Children with and without Developmental Disabilities. *PsyArXiv*, 1–35. https://doi.org/10.31234/osf.io/xveqy.

Tager-Flusberg, H. (2000). Understanding the language and communicative impairments in autism. *International Review of Research in Mental Retardation*, *23*, 185–205. https://doi.org/10.1016/s0074-7750(00)80011-7.

Thompson, G. A., Shanahan, E. C., & Gordon, I. (2019). The role of music-based parent-child play activities in supporting social engagement with children on the autism spectrum: A content analysis of parent interviews. *Nordic Journal of Music Therapy*, *28*(2), 108–130. https://doi.org/10.1080/08098131.2018.1509107.

Tryfon, A., Foster, N. E., Ouimet, T., Doyle-Thomas, K., Anagnostou, E., Sharda, M., & Hyde, K. L. (2017). Auditory-motor rhythm synchronization in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *35*, 51–61. https://doi.org/10.1016/j.rasd.2016.12.004.

Vega, J. N., Hohman, T. J., Pryweller, J. R., Dykens, E. M., & Thornton-Wells, T. A. (2015). Resting-state functional connectivity in individuals with Down syndrome and Williams syndrome compared with typically developing controls. *Brain Connectivity*, *5*(8), 461–475. https://doi.org/10.1089/brain.2014.0266.

Williams, K. E., Berthelsen, D., Nicholson, J. M., Walker, S., & Abad, V. (2012). The effectiveness of a short-term group music therapy intervention for parents who have a child with a disability. *Journal of Music Therapy*, *49*(1), 23–44. https://doi.org/10.1093/jmt/49.1.23.

Credible signalling and social bonds: Ultimately drawing on the same idea

Patrick Kennedy and Andrew N. Radford

School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK.
patrick.kennedy@bristol.ac.uk
andy.radford@bristol.ac.uk;
http://www.bio.bris.ac.uk/research/behavior/Vocal_Communication/home.html

doi:10.1017/S0140525X20001454, e102

Abstract

The hypotheses in both target articles rely implicitly on much the same logic. For a "social-bonding" device to make sense, there must be an underlying reason why an otherwise-arbitrary behaviour sustains alliances – namely, credible signals of one's value to partners. To illustrate our points, we draw on the parallels with supposed bonding behaviours in nonhuman animals.

The target articles present themselves as diametrically opposed. Mehr and coauthors argue that human musical skill and appreciation ("musicality") has its origins in the use of rhythm and sound to communicate trustworthy information, ensuring honesty in contexts where partners are tempted to lie. Savage and coauthors, by contrast, argue that musicality evolved to establish and cement "social bonds," and to coordinate behaviours when partners have aligned interests. We believe that the hypotheses overlap much more than either side admits, given the breadth of "credible signals" in the natural world.

Savage et al.'s argument that musicality provides adaptive "social-bonding" benefits is split into three components (see their Fig. 3): (1) strengthening "bonds" with groupmates; (2) providing cues of group membership; and (3) enhancing within-group coordination. Yet, the first and second of these must ultimately be based on credible signalling, which is Mehr et al.'s proposed mechanism.

Our first point is that explaining the existence of exclusive and enduring "bonds" between individuals remains an open question in social evolution. Social bonds are "stable, equitable, and strong" dyadic relationships (Ostner & Schülke, 2014). A "social-bonding mechanism" is an adaptive behaviour that increases the strength of a specific bond. It must make it more profitable for the two individuals to invest in the dyadic relationship than to leave it, overcoming temptations to defect. Social-bonding mechanisms only make sense where conflicts of interest can arise (which do not feature explicitly in Savage et al.): if there is automatic alignment of interest, there is no need for any "bonding" to establish and stabilize a relationship.

Nonhuman animals can increase the reliability of a relationship in two main ways: (1) by basing relationships on a series of small reciprocal trades (e.g., removal of parasites by allogrooming in primates, ungulates, and birds; Akinyi et al., 2013; Mooring, Blumstein, & Stoner, 2004; Radford & Du Plessis, 2006), reducing the stakes of each step (Dixit & Nalebuff, 1991), and rewarding good behaviour; or (2) by using credible signals of ability and/or commitment to advertise a future resource (e.g., males honestly advertising parenting quality; Buchanan & Catchpole, 2000; Pettitt, Bourne, & Bee, 2020). It is difficult to see how music is a tradable resource – there is no immediate fitness benefit – and neither paper provides an account of music trading. This leaves credible signalling of trustworthiness or ability as the implicit adaptive device for strengthening "social bonds." This may occur in myriad ways; for instance, paying the cost of investing in lengthy bouts of music-making may only be in the interest of partners pursuing long-term alliance benefits (e.g., support during within-group contests). In this view, both hypotheses rely on the same logic.

Drawing on Dunbar's "grooming-at-a-distance" hypothesis (Dunbar, 2012), Savage et al. appear to argue that the benefit of music was release of dopaminergic rewards, and consequent "bonding." We agree with Mehr et al.: this argument is circular.

Selection builds proximate rewards to drive animals to pursue behaviours that are ultimately adaptive. Invoking dopaminergic rewards begs the question: the aim is to explain *why* brains are initially attracted to music (i.e., why they are rewarded with dopamine or other drivers). Savage et al. might escape the charge of circular reasoning by more explicitly distinguishing traits representing "senders" (ability to produce music) and "receivers" (attraction to hearing music). In principle at least, the second might be a non-adaptive by-product of some other feature of the brain (a "sensory bias") to which the first initially evolved as an adaptation (cf. fish swordtails; Basolo, 1995).

Confusion of proximate and ultimate explanations for affiliation extends beyond music to other aspects of animal behaviour. Affiliative acts between groupmates abound: for example, body contacts in ants (Birch, Cant, & Thompson, 2019), soft bumps in cichlid fish (Bruintjes, Lynton-Jenkins, Jones, & Radford, 2015), and preening or grooming in birds and mammals (Borgeaud & Bshary, 2015; Kern & Radford, 2018; Radford, 2011). Simply invoking the terms "social bonding," "social cohesion," and "affiliation" (or proximately "relieving stress") for these stops short of real explanations, which are explicit about ultimate costs and benefits, the origin of constraints, and – in cases where signalling plays a role – the information communicated and the guarantees of credibility.

The second component of Savage et al.'s social-bonding hypothesis is that musicality provides cues of group membership. This argument is explicitly about credible signals, and indeed also features in Mehr et al. Here, music would be analogous to the group-specific cuticular hydrocarbons of insects (van Zweden & D'Ettorre, 2010), call signatures of birds (Hopp, Jablonski, & Brown, 2001; Radford, 2005), and olfactory cues of mammals (Christensen, Kern, Bennitt, & Radford, 2016; Henkel & Setchell, 2018): difficult-to-fake indices of group identity.

Only the third component of Savage et al.'s hypothesis does not imply that music is a device allowing credible signalling. Here, musicality promotes within-group coordination by synchronizing emotions or behaviours. It is unclear how or why this coordination should occur; we read it as describing individuals that have no reason not to trust one another, sharing a common interest in coordination and looking to exchange information. This is signalling, with the credibility of the signals assumed. Here, music would be analogous to trustworthy acoustic signals that coordinate movement or recruitment within animal groups (Braune, Schmidt, & Zimmermann, 2005; Radford, 2004; Radford & Ridley, 2006).

The hypotheses proposed in each paper have much overlap: both must see music as an honest signal facilitating cooperation. The authors differ in the specific contexts of cooperation that they favour as important at the origins of musicality: Mehr et al. focus on inter-group coalitions and mother-offspring bonding, whereas Savage et al. focus on within-group coalitions. There exists sparse empirical evidence either way for these interesting, if speculative, scenarios. More generally, costs maintain the credibility of diverse signals across the natural world (Biernaskie, Perry, & Grafen, 2018). If music is a signal, perhaps the most progress will therefore be made by taking a leaf out of the animal behaviour book and asking empirically where the costs lie.

Financial support. This study was supported by a European Research Council Consolidator Grant (project no. 682253) awarded to ANR.

Conflict of interest. None.

References

Akinyi, M. Y., Tung, J., Jeneby, M., Patel, N. B., Altmann, J., & Alberts, S. C. (2013). Role of grooming in reducing tick load in wild baboons (*Papio cynocephalus*). *Animal Behaviour*, *85*(3), 559–568. doi:10.1016/j.anbehav.2012.12.012.

Basolo, A. L. (1995). Phylogenetic evidence for the role of a pre-existing bias in sexual selection. *Proceedings of the Royal Society B: Biological Sciences*, *259*(1356), 307–311. doi:10.1098/rspb.1995.0045.

Biernaskie, J. M., Perry, J. C., & Grafen, A. (2018). A general model of biological signals, from cues to handicaps. *Evolution Letters*, *2*(3), 201–209. doi:10.1002/evl3.57.

Birch, G., Cant, M. A., & Thompson, F. J. (2019). Behavioural response of workers to repeated intergroup encounters in the harvester ant *Messor barbarus*. *Insectes Sociaux*, *66*(3), 491–500. doi:10.1007/s00040-019-00710-2.

Borgeaud, C., & Bshary, R. (2015). Wild vervet monkeys trade tolerance and specific coalitionary support for grooming in experimentally induced conflicts. *Current Biology*, *25*(22), 3011–3016. doi:10.1016/j.cub.2015.10.016.

Braune, P., Schmidt, S., & Zimmermann, E. (2005). Spacing and group coordination in a nocturnal primate, the golden brown mouse lemur (*Microcebus ravelobensis*): The role of olfactory and acoustic signals. *Behavioral Ecology and Sociobiology*, *58*(6), 587–596. doi:10.1007/s00265-005-0944-4.

Bruintjes, R., Lynton-Jenkins, J., Jones, J. W., & Radford, A. N. (2015). Out-group threat promotes within-group affiliation in a cooperative fish. *The American Naturalist*, *187*(2), 274–282. doi:10.1086/684411.

Buchanan, K. L., & Catchpole, C. K. (2000). Song as an indicator of male parental effort in the sedge warbler. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1441), 321–326. doi:10.1098/rspb.2000.1003.

Christensen, C., Kern, J. M., Bennitt, E., & Radford, A. N. (2016). Rival group scent induces changes in dwarf mongoose immediate behavior and subsequent movement. *Behavioral Ecology*, *27*(6), 1627–1634. doi:10.1093/beheco/arw092.

Dixit, A., & Nalebuff, B. (1991). *Thinking strategically*. W. W. Norton & Company.

Dunbar, R. I. M. (2012). On the evolutionary function of song and dance. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 201–214). Oxford University Press. doi:10.1093/acprof.

Henkel, S., & Setchell, J. M. (2018). Group and kin recognition via olfactory cues in chimpanzees (*Pan troglodytes*). *Proceedings of the Royal Society B: Biological Sciences*, *285* (1889), 20181527. doi:10.1098/rspb.2018.1527.

Hopp, S. L., Jablonski, P., & Brown, J. L. (2001). Recognition of group membership by voice in Mexican jays, *Aphelocoma ultramarina*. *Animal Behaviour*, *62*(2), 297–303. doi:10.1006/anbe.2001.1745.

Kern, J. M., & Radford, A. N. (2018). Experimental evidence for delayed contingent cooperation among wild dwarf mongooses. *Proceedings of the National Academy of Sciences*, *115*(24), 6255–6260. doi:10.1073/pnas.1801000115.

Mooring, M. S., Blumstein, D. T., & Stoner, C. J. (2004). The evolution of parasite-defence grooming in ungulates. *Biological Journal of the Linnean Society*, *81*(1), 17–37. doi:10.1111/j.1095-8312.2004.00273.x.

Ostner, J., & Schülke, O. (2014). The evolution of social bonds in primate males. *Behaviour*, *151*(7), 871–906. doi:10.1163/1568539X-00003191.

Pettitt, B. A., Bourne, G. R., & Bee, M. A. (2020). Females prefer the calls of better fathers in a Neotropical frog with biparental care. *Behavioral Ecology*, *31*(1), 152–163. doi:10.1093/beheco/arz172.

Radford, A. N. (2004). Vocal coordination of group movement by green woodhoopoes (*Phoeniculus purpureus*). *Ethology*, *110*(1), 11–20. doi:10.1046/j.1439-0310.2003.00943.x.

Radford, A. N. (2005). Group-specific vocal signatures and neighbour-stranger discrimination in the cooperatively breeding green woodhoopoe. *Animal Behaviour*, *70*(5), 1227–1234. doi:10.1016/j.anbehav.2005.04.002.

Radford, A. N. (2011). Preparing for battle? Potential intergroup conflict promotes current intragroup affiliation. *Biology Letters*, *7*, 26–29. doi:10.1098/rsbl.2010.0507.

Radford, A. N., & Du Plessis, M. A. (2006). Dual function of allopreening in the cooperatively breeding green woodhoopoe, *Phoeniculus purpureus*. *Behavioral Ecology and Sociobiology*, *61*(2), 221–230. doi:10.1007/s00265-006-0253-6.

Radford, A. N., & Ridley, A. R. (2006). Recruitment calling: A novel form of extended parental care in an altricial species. *Current Biology*, *16*(17), 1700–1704. doi:10.1016/j.cub.2006.06.053.

van Zweden, J. S., & D'Ettorre, P. (2010). Nestmate recognition in social insects and the role of hydrocarbons. In G. J. Blomquist & A.-G. Bagnères (Eds.), *Insect hydrocarbons biology, biochemistry, and chemical ecology* (pp. 222–243). Cambridge University Press. doi:10.1017/CBO9780511711909.012.

Not by signalling alone: Music's mosaicism undermines the search for a proper function

Anton Killin[a], Carl Brusse[b,c], Adrian Currie[d] and Ronald J. Planer[e]

[a]School of Philosophy and ARC Centre of Excellence for the Dynamics of Language, Australian National University, Acton, ACT 2601, Australia; [b]Department of Philosophy and Charles Perkins Centre, The University of Sydney, Sydney, NSW 2006, Australia; [c]School of Philosophy, Australian National University, Acton, ACT 2601, Australia; [d]Department of Sociology, Anthropology and Philosophy, University of Exeter, Exeter EX4 4RJ, UK and [e]School of Language and Linguistics, University of Melbourne, Parkville, Victoria 3010, Australia.
anton.killin@anu.edu.au;
carl.brusse@sydney.edu.au
a.currie@exeter.ac.uk
ronald.j.planer@outlook.com

doi:10.1017/S0140525X20001107, e103

Abstract

Mehr et al. seek to explain music's evolution in terms of a unitary proper function – *signalling cooperative intent* – which they cash out in two guises, coalition signalling and (allo)parental attention signalling. Although we recognize the role signalling almost certainly played in the evolution of music, we reject "ultimate" causal explanations which focus on a unidirectional, narrow range of causal factors.

We're delighted to see two target articles about the evolution of music appear together in *BBS*. Each tackles music very differently: One developing an adaptationist scenario emphasizing music's role in credibly signalling cooperative intent (Mehr et al.), and the other focusing on gene-culture coevolution (Savage et al.).

Savage et al. approach music as a multifaceted, complex coevolutionary phenomena demanding a treatment which interweaves cognitive and morphological innovation with cultural and biological evolution, emphasizing feedback between these processes. Regardless of the success of Savage et al.'s specific proposal, we think something along those lines is required to explain music's evolution. In our view, music's coevolutionary origins render Mehr et al.'s approach to music via the adaptation/byproduct dichotomy misguided. Let us explain via analogy.

Distinctively hominin hand morphology, the hand's executive control, and lithic technologies all arose through processes of niche construction and coevolutionary dynamics and feedback. As Downes (2010, p. 249) has put it: "The hand did not evolve in response to a particular environmental stimulus at any particular time. Rather, various selection pressures, including bipedalism, the occupation of niches with widely varying food resources, and our own niche construction, led to the musculature and bone structure that supports the range of activities for which human hands can be used." In light of these coevolutionary dynamics, it is not productive to ask whether hominin hands and cognition were "adapted" for tool production and use, "exapted" for tool production and use, or whether tool production and use are "byproducts" of morphological and cognitive adaptations. We think the evolution of the dynamic, complex, coevolved mosaic "musicality" – the suite of morphological and cognitive capacities supporting the production and perception of musical displays – is roughly analogous. It doesn't follow from this that explanations appealing to adaptations and byproducts are never effective (say, the heart might be an adaptation for pumping blood, whereas blood's colour is a byproduct of its biochemistry). Rather, that vocabulary implies a causal simplicity which overlooks music's likely complex, niche-constructed, coevolutionary path (Killin, 2018a).

When dealing with a complex biocultural phenomenon, any account focused on a narrow range of unidirectional evolutionary pressures will be highly incomplete at best. But, that is exactly what Mehr et al. offer. They target psychological mechanisms underwriting signalling of cooperative intent, purportedly "adaptations for credible signalling, which give rise to a universal human psychology of music" (sect. 5.3, para. 2). Of course, it doesn't follow from music's multifaceted nature and complex evolution that signalling isn't an important part of its story (indeed, we suspect signalling roles were probably quite important and suggest that Savage et al. do more to acknowledge this). But, it does undermine appeals to a small number of narrow causal factors. Many other selective processes driving music's evolution are left by the wayside.

That said, we think Mehr et al.'s specific signalling hypotheses need further interrogation on their own terms. First, they see music as "a means for groups to credibly show off their qualities to other groups" (sect. 4.2.1, para. 10). As the authors point out, senders and receivers must be appropriately incentivized for signalling to evolve. The senders and receivers in this scenario are groups, with the incentive initially being territorial, "to deter intruders and avoid a fight" (sect. 4.2.1, para. 1), coopted later in human evolution for advertising/assessing alliance opportunities. But, are these sufficient, given the costs of music-acquisition? It is hard to say. This is complicated by the need for individual-level decomposition, as groups don't perceive music, individuals do, and they also pay its costs. Chimpanzees and bonobos can be vocal and rowdy but they do not sing or rhythmically entrain with conspecifics: Ancestral hominins would have needed to acquire these abilities even if they were rudimentary by modern human standards. This would have been cognitively demanding and imposed opportunity costs. We question whether the posited payoffs here could have provided a consistent, deep-time selective environment to explain coordinated music's origins, as Mehr et al. claim, or whether this is better seen as a form of stabilizing selection (which would also explain the ethnographic evidence discussed).

Consider synchrony. As the authors say, "a high level of synchronous coordination among signalers requires considerable effort to achieve" (sect. 4.2.1, para. 4). But, is it sufficiently well-correlated with a valuable trait (cooperative ability generally?) for an index signal mechanism to evolve? And, were such highly-synchronous displays really so ubiquitous as to be a valuable source of such information, incentivizing adaptive response? We find it equally plausible that (1) these hypothesized synchronous displays were actually quite loosely correlated with the actual fitness-relevant virtues supposedly being signalled for (as natural variation in musical aptitude partially suggests), and (2) that they instead worked (insofar as they did) by manipulating pre-existing responses to audible cues. Therefore, they may well be part of music's evolutionary story but the authors' case for ultimate causation here is weak. Indeed, unlike their example of bitter taste (for which their dichotomous ultimate/proximate causal framework is arguably adequate), coevolutionary feedback across evolution and development renders "ultimate" causation for music somewhat artificial.

Second, without a chronology, Mehr et al.'s hypothesis that infant-directed maternal contact calls evolved to signal (allo) parental attention is difficult to assess, partly because it requires the shaky assumption that our long-ancient ancestors' interbirth intervals resembled those of modern foragers. Whether we are talking about erectines or heidelbergensians or archaic sapiens matters here. For if interbirth intervals at the relevant time were longer – say, between those of modern foragers and chimpanzees (almost certainly true of erectines) – those ancient mothers would have had fewer fully-dependent young children at any given time, weakening selection for that signalling strategy.

Again, it doesn't follow that infant-directed vocalizations played no role in music's evolution. We think they would have. The point is that hypotheses focusing on sharply limited ranges of causal factors are bound to be inadequate for explaining complex traits (Currie, 2014, 2019) and musicality is a complex trait: We must instead point to the multifaceted and interwoven nature of the evolutionary dynamics behind music's evolution, as Savage et al. do (see also Killin, 2016, 2017, 2018b).

Financial support. C.B.'s research was supported in part by a grant from the John Templeton Foundation (Grant ID 60811).

Conflict of interest. None.

References

Currie, A. (2014). Narratives, mechanisms and progress in historical science. *Synthese*, *191*, 1163–1183.

Currie, A. (2019). Simplicity, one-shot hypotheses and paleobiological explanation. *History and Philosophy of the Life Sciences*, *41*. https://doi.org/10.1007/s40656-019-0247-0.

Downes, S. M. (2010). The basic components of the human mind were not solidified during the Pleistocene epoch. In F. J. Ayala & R. Arp (Eds.), *Contemporary debates in philosophy of biology* (pp. 243–252). Wiley-Blackwell.

Killin, A. (2016). Rethinking music's status as adaptation versus technology: A niche construction perspective. *Ethnomusicology Forum*, *25*(2), 210–233.

Killin, A. (2017). Plio-Pleistocene foundations of musicality: The coevolution of hominin cognition, sociality and music. *Biological Theory*, *12*(4), 222–235.

Killin, A. (2018a). Music and human evolution: Philosophical aspects. In R. Joyce (Ed.), *Routledge handbook of evolution and philosophy* (pp. 372–386). Routledge.

Killin, A. (2018b). The origins of music: Evidence, theory, and prospects. *Music & Science*, *1*. https://doi.org/10.1177/2059204317751971.

The origins of music in (musi) language

Evelina Leivada

Department of English and German Studies, Universitat Rovira i Virgili, 43002 Tarragona, Spain.
evelina.leivada@urv.cat, https://evelinaleivada.com/

doi:10.1017/S0140525X20000813, e104

Abstract

The view of music as a byproduct of other cognitive functions has been deemed incomplete or incorrect. Revisiting the six lines of evidence that support this conclusion, it is argued that it is unclear how the hypothesis that music has its origins in (musi)language is discarded. Two additional promising research lines that can support or discard the byproduct hypothesis are presented.

One of Mehr, Krasnow, Bryant, and Hagen's aims is to show that the hypothesis that music is a byproduct of other faculties fails in light of six lines of evidence. The evaluation of various adaptation hypotheses led to the conclusion that the byproduct view is incomplete or incorrect. These six arguments are revisited below from a biolinguistic point of view in order to evaluate their individual validity in relation to one premise: the possible origins of music in (musi)language.

The first argument of Mehr et al. is that complex, song-like vocalizations have evolved convergently across distant species, and they are socially learned, like music is. However, if these elaborate, song-like vocalizations feature semantically meaningful and functionally referential information (Arnold & Zuberbühler, 2012; Elie & Theunissen, 2016), they are analogous to human language, not human music. If so, this first argument does not show why the hypothesis that music is the byproduct of language is incomplete or incorrect.

The second argument is that music is a human universal. This claim, although correct, does not really discard the byproduct hypothesis specifically in relation to language, because language is universal too. The spontaneous development of home-signs in deaf populations (Senghas, Kita, & Özyürek, 2004) and the ability to lexicalize even in cases of severe absence of language input (Gleitman & Newport, 1995) are unparalleled indicators of the universality of language. For this second argument to successfully discard the byproduct hypothesis, cases of societies that have music but not language should be presented.

The next argument put forth by Mehr et al. is that music is complex and it features grammar-like structures analogous to those of language. The presentation of this argument makes clear that the identified feature is shared with language. This argument has not identified any flaws in the byproduct hypothesis.

The fourth argument is that music perception appears early in ontogeny. This argument has an ontogenetic basis, not a phylogenetic one, which is what would be needed for discarding the byproduct hypothesis about the origins of music. Although it is true that neonates detect the beat of music, this could well be the outcome of central auditory processing that is used when deciphering the rhythm of language *in utero* (Minai, Gustafson, Fiorentino, Jongman, & Sereno, 2017).

The fifth line of evidence that aims to show that the byproduct hypothesis is incorrect is that music perception displays evidence for neural specialization and is impaired in specific deficits (e.g., congenital amusia). The first part of this argument does not disprove that music originates in language: Identifying a neural basis does not entail a unique one-to-one mapping between a specific brain region and a specific cognitive function, because brain regions form part of networks that underlie the workings of many cognitive processes. Also, developmental disorders do not necessarily represent selective impairment to a specific locus, but may rather result from early perturbations of the interconnected neonate brain that prevent specialization of function from developing (D'Souza & Karmiloff-Smith, 2011). It is important that most studies describe language deficits in amusics (e.g., Liu, Jiang, Wang, Xu, & Patel, 2015; Sihvonen et al., 2016), leaving open the possibility of observing overlapping brain regions and shared loci of neural disruption in the two domains, language and music. The neural overlap hypothesis does not entail shared neural circuitries between language and music (Peretz, Vuvan, Lagrois, & Armony, 2015), but it does not discard the possibility of neural recycling either, in the sense that language may have been recycled during evolution for musicality (Honing, ten Cate, Peretz, & Trehub, 2015).

The last argument is that music is ancient. The claim that the byproduct hypothesis is incomplete/incorrect does not follow from this argument, especially if one entertains this hypothesis in relation to language. Mehr et al. (2020) mention that flutes are at least 40k years old, but language has been argued to play a role in our species' evolution for the last 200k years (Pagel, 2017).

None of these replies to the six arguments put forth by Mehr et al. (2020) shows that the main hypothesis they advance is incorrect. However, these replies suggest that Mehr et al. (2020) have not established why the byproduct hypothesis that links the origins of music to language is flawed or why the hypothesis of shared origin via a common ancestral precursor that could scaffold both music and language (e.g., musilanguage; Brown, 2001) is unconvincing.

Suggesting that the main claim of Mehr et al. (2020) has not been established through the six pursued lines of argument does not entail that there are no experiments that can actually support it. Arguments both in favor or against it can be adduced through two promising lines of evidence. The first one is to examine species that show signs of musical perception, although their communication system does not rely primarily on acoustic signals. One promising example is *Cyprinus carpio*, which was found to be successfully trained to discriminate music from different genres (Chase, 2001). The second approach is to determine whether music is innate as language is. The innateness of language is well-known: A typically developing infant does not have the option to suppress her innate linguistic predisposition and not pick up the language of her environment. The question is whether the same can be claimed for music. Another line of investigation of this argument boils down to examining the incidence of patients that perceive spoken language although they have never received linguistic input in the auditory modality. Congenitally deaf people may develop both auditory and tactile hallucinations such as hearing voices or being fingerspelled to (Atkinson, 2006). A PubMed search of the terms "congenital" AND "deaf" AND "music" suggests that this is not the case for musical hallucinations. If music is innate as language is, and given that musical hallucinations in people with acquired deafness have been amply reported (e.g., Fisman, 1991), the extremely low – if existent at all – occurrence of musical hallucinations in congenitally deaf people is a fact that awaits explanation. Providing one possibly goes through talking about the innateness and centrality of music in human cognition.

Financial support. This study received support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 746652 and from the Spanish Ministry of Science, Innovation and Universities under the Ramón y Cajal grant agreement no. RYC2018-025456-I.

Conflict of interest. None.

References

Arnold, K., & Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language*, *120*, 303–309.

Atkinson, J. R. (2006). The perceptual characteristics of voice-hallucinations in deaf people: Insights into the nature of subvocal thought and sensory feedback loops. *Schizophrenia Bulletin*, *32*(4), 701–708.

Brown, S. (2001). The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 271–300). MIT Press.

Chase, A. R. (2001). Music discriminations by carp (*Cyprinus carpio*). *Animal Learning & Behavior*, *29*, 336–353.

D'Souza, D., & Karmiloff-Smith, A. (2011). When modularization fails to occur: A developmental perspective. *Cognitive Neuropsychology*, *28*, 276–287.

Elie, J. E., & Theunissen, F. E. (2016). The vocal repertoire of the domesticated zebra finch: A data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition*, *19*, 285–315.

Fisman, M. (1991). Musical hallucinations: Report of two unusual cases. *The Canadian Journal of Psychiatry*, *36*(8), 609–611.

Gleitman, L. R., & Newport, E. L. (1995). The invention of language by children: Environmental and biological influences on the acquisition of language. In L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science* (pp. 1–24). The MIT Press.

Honing, H., ten 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*, 1664.

Liu, F., Jiang, C., Wang, B., Xu, Y., & Patel, A. D. (2015). A music perception disorder (congenital amusia) influences speech comprehension. *Neuropsychologia*, *66*, 111–118.

Minai, U., Gustafson, K., Fiorentino, R., Jongman, A., & Sereno, J. (2017). Fetal rhythm-based language discrimination: A biomagnetometry study. *Neuroreport*, *28*(10), 561–564.

Pagel, M. (2017). Q&A: What is human language, when did it evolve and why should we care? *BMC Biology*, *15*, 64.

Peretz, I., Vuvan, D., Lagrois, M.-E., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 1664.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science (New York, N.Y.)*, *305*, 1779–1782.

Sihvonen, A. J., Ripolles, P., Leo, V., Rodriguez-Fornells, A., Soinila, S., & Sarkamo, T. (2016). Neural basis of acquired amusia and its recovery after stroke. *Journal of Neuroscience*, *36*(34), 8872–8881.

Knowledge songs as an evolutionary adaptation to facilitate information transmission through music

Daniel J. Levitin[a,b]

[a]Minerva Schools at Keck Graduate Institute, San Francisco, CA 94103, USA and [b]Department of Psychology, McGill University, Montreal, QC H3A 1G1, Canada.
dlevitin@minerva.kgi.edu; www.daniellevitin.com

doi:10.1017/S0140525X20001090, e105

Abstract

I propose an adjunct to the two models presented in the target articles, a function of music that is ubiquitous and would have solved a clear adaptive problem, that of transmitting important survival information among pre-literate humans. This class of *knowledge songs* uniquely preserved cultural, botanical, medical, safety, and practical information that increased the adaptive fitness of societies.

Musicality comprises a set of capabilities that evolved and co-evolved to support multiple functions (Savage et al., target article). Here, I seek to fill in a gap in each target article with evidence suggesting a different putative function of music (while remaining agnostic as to which of the two models offers the more compelling account of music's origins). I present evidence for a category of ancient songs that solved an important adaptive problem: The transmission of essential knowledge that increased adaptive fitness of individuals and the community. This proposal is consistent with, and requires no modification, to either model.

Modern *Homo sapiens* have been around for 60,000–200,000 years (Wilshaw, 2018). Written language emerged autochthonously around the world only about 5,000 years ago, facilitating the preservation and transmission of information. For millennia before that, however, survival information was transmitted across generations through oral traditions – storytelling and song (Berkes, Colding, & Folke, 2000; Reyes-García & Fernández-Llamazares, 2019; Rubin, 1995); stories and songs constituted the fundamental pedagogical devices of preliterate societies (Coe, 2003; Jennings et al., 2005; Scalise-Sugiyama, 2011). The songs of preliterate societies preserved and transmitted information about fitness hazards, ethnobiological knowledge, food gathering, morality, mythology, kinship, medicine, and practical skills (Lord, 1960; Scalise-Sugiyama, 1996; Schniter et al., 2018) – what I will collectively call *knowledge songs*.

Music-making incurs opportunity costs (Mehr & Krasnow, 2017) that must be offset by the individual or collective benefit. Adaptations must show a clear fit between the unique design features they offer and the problems they putatively solved. The opportunity costs of music-making would be outweighed by songs that enable *knowledge transmission* (as well as other functions discussed in the target articles). Musicality may constitute an adaptation because it is uniquely suited to encode and preserve information, in a way that spoken language alone cannot, and that served the needs of pre-literate humans for tens of thousands of years (and still serves the needs of pre-literate cultures today).

Meaning and sound patterns are paired arbitrarily in speech (Rubin, Stoltzfus, & Wall, 1991) but music aids word recall (Rubin, 1995). In music, the mutually reinforcing cues of meter, accent structure, melodic contour, prosody, rhythm, and rhyme, create constraints on lyrics that preserve them with far greater accuracy than is found in non-musical oral memory (Hyman & Rubin, 1990; Kintsch, 1988; Lattimore, 1951; Palmer & Kelly, 1992; Rubin et al., 1991; Schwanenflugel & LaCount, 1988).

Mehr et al.'s (2019) landmark study revealed 20 widespread functional contexts for music (a single song can be assigned more than one category). Absent from their list are knowledge songs. Levitin (2008) examined corpora of recorded music from contemporary preliterate subsistence cultures from every continent, and found that knowledge songs were ubiquitous, transmitting information about *identification of plants and animals*, *kinship*, *daily routines*, *food preparation*, *healing practices*, *practical skills (e.g., fishing, hunting, trapping, and hut building)*, *tribal history*, and *prescribed behaviors*.

I reanalyzed Mehr et al.'s dataset of 4,700 ethnographies, and searched for these terms and related keywords (*transmission*, *avoidance and taboo*, *instructions & education including sex restrictions*, *geography*, *history*, *mythology* and *literary texts and literature*). Of the 4,700 songs in the corpus, 1,007 unique entries fit these criteria, or roughly 20%, converging with Levitin's (2008) analysis of a different set of corpora.

A typical song is from a 105 year-old chief of the Blackfoot tribe in Montana, described by the ethnographer as "a memory ethnograph" with songs that encapsulate oral history (Ehrafworld cultures, n.d.). Or consider the Gola of West Africa who are typical in placing a high value on the preservation and transmission of tribal and kinship history through song (D'Azevedo, 1962). The knowledge of kinship origins can help establish among contemporaries familial connections and reciprocal responsibilities; being able to claim a relative during a famine can mean the difference between life and death.

There is overlap, of course, in the assignment of songs to such categories. The plethora of songs about burial practices occurs in the context of ritual or ceremonial songs, and is also passing on procedural information. Similarly, songs about healing bridge shamanistic, spiritual, and medical practices with procedural information (do it *this* way, and in *this* order).

Today, we may think of music as primarily about emotional communication but this could be an ethnocentric bias; before written language, music shows evidence of a superior preservation system than speech alone. And there exist vestiges of knowledge music in contemporary, advanced literate society supporting an "ultimate level explanation" (Mehr et al., target article). For example, most children learn the alphabet and number line through songs, and some learn body parts and the left-right distinction (e.g., *The Hokey Pokey*), as well as social justice (*If I Had A Hammer, We Shall Overcome*).

The survival prospects of individuals and groups are enhanced by a capacity to communicate certain information about states of affairs in the physical world, and about the social world that concerns the organism (Cross, 2007). The ideal communication system would allow individuals to communicate knowledge about current conditions such as the availability and locations of resources, to make possible their sharing. Perceptions of dangers would need to be identified and appropriate actions coordinated; social relationships would need to be articulated and sustained. Why is music necessary and even better than language for such tasks? Because music, especially rhythmic, patterned music of the kind we typically associate with songs, provides a more powerful mnemonic force for encoding knowledge, vital and shared information that entire societies need to know, teachings that are handed down by parents to their children and that children can easily memorize (Levitin, 2008; Rubin, 1995).

Knowledge songs shared within a tribe and family would have constituted one of the greatest forms of social bonding, allowing for the transmission of culture and survival information before the written word existed. Our transformation from club-wielding early hominids to the architects of great metropolises, discoverers of the scientific method, and listeners to Weird Al Yankovic, owes its deep history to knowledge conveyed in music over the millennia that preceded writing systems.

Financial support. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (228175-10).

Conflict of interest. None.

References

Berkes, F., Colding, J., & Folke, C. (2000). Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications 10*, 1251–1262.

Coe, C. (2003). *The ancestress hypothesis: Visual arts as adaptation*. Rutgers University Press.

Cross, I. (2007). Symposium. In M. Klockars & M. Peltomaa (Eds.), *Music meets medicine* (pp. 5–13). Signe and Ane Gyllenberg Foundation.

D'Azevedo, W. L. (1962). Uses of the past in Gola discourse. *Journal of African History*, *3*, 11–34.

Ehrafworld cultures (n.d.) https://ehrafworldcultures.yale.edu/ehrafe/citation.do?method=citation&forward=browseAuthorsFullContext&id=nf06-054
Hyman, I. E., & Rubin, D. C. (1990). Memorabeatlia: A naturalistic study of long-term memory. *Memory & Cognition*, *18*(2), 205–214.
Jennings, J., Antrobus, K., Atencio, S., Glavich, E., Johnson, R., Loffler, G., … & Jennings, J. (2005). "Drinking beer in a blissful mood" Alcohol production, operational chains, and feasting in the ancient world. *Current Anthropology*, *46*(2), 275–303.
Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182.
Lattimore, R. (1951). *Homer: The Iliad*. University of Chicago Press.
Levitin, D. J. (2008). *The world in six songs: How the musical brain created human nature*. Penguin Random House.
Lord, A. B. (1960). *The singer of tales*. 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*(5), 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*, *366*(6468), 957–970.
Palmer, C., & Kelly, M. H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language*, *31*(4), 525–542.
Reyes-García, V., & Fernández-Llamazares, Á. (2019). Sing to learn: The role of songs in the transmission of indigenous knowledge among the Tsimane' of Bolivian Amazonia. *Journal of Ethnobiology*, *39*(3), 460–477.
Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press on Demand.
Rubin, D. C., Stoltzfus, E. R., & Wall, K. L. (1991). The abstraction of form in semantic categories. *Memory & Cognition*, *19*(1), 1–7.
Scalise-Sugiyama, M. (1996). On the origins of narrative. *Human nature*, *7*(4), 403.
Scalise-Sugiyama, M. (2011). The forager oral tradition and the evolution of prolonged juvenility. *Frontiers in Psychology 2*, 133. doi:10.3389/fpsyg.2011.00133.
Schniter, E., Wilcox, N. T., Beheim, B. A., Kaplan, H. S., & Gurven, M. (2018). Information transmission and the oral tradition: Evidence of a late-life service niche for Tsimane Amerindians. *Evolution and Human Behavior*, *39*(1), 94–105.
Schwanenflugel, P. J., & LaCount, K. L. (1988) Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning. Memory, and Cognition*, *14*, 344–354.
Wilshaw, A. (2018). Modern humans, origins of. *The International Encyclopedia of Anthropology*, 1–14.

If it quacks like a duck: The by-product account of music still stands

Debra Lieberman[a] and Joseph Billingsley[b]

[a]Department of Psychology, University of Miami, Coral Gables, FL 33124-0751, USA and [b]Department of Management, Innovation, and Entrepreneurship, Poole College of Management, North Carolina State University, Raleigh, NC 27695, USA.
debra@miami.edu
josephbillingsleypsych@gmail.com

doi:10.1017/S0140525X20000990, e106

Abstract

Discerning adaptations from by-products is a defining feature of evolutionary science. Mehr, Krasnow, Bryant, and Hagen posit that music is an adaptation that evolved to function as a credible signal. We counter this claim, as we are not convinced they have dispelled the possibility that music is an elaboration of extant features of language.

Mehr, Krasnow, Bryant, and Hagen (Mehr et al.) present a cogent argument that music evolved as a credible signal of coalitional formidability and, within parent–infant relationships, of caregiver attention. Their careful application of an adaptationist logic serves as a prime example of how to conduct work in evolutionary science, and they marshal a compelling case against both the mate quality and social bonding models of music's origins. That said, we do not believe Mehr et al. provide adequate grounds to dismiss the hypothesis that music is a by-product of adaptations for language. To illustrate, consider the six points which the authors use to dispel the by-product hypothesis.

The authors first suggest that the widespread convergent evolution of "song-like vocalizations" and the presence of "musical behaviors" across species demonstrates that "music-like adaptations" could have evolved in humans. But, as the authors acknowledge in a footnote, it's not clear what these vocalizations and other behaviors represent. Can we be sure that these phenomena are not instead "proto-language-like" evidencing that "proto-language-like" adaptations can evolve? Calling animal vocalizations "musical" or "song-like" as a new category of phenomena might be unfounded, perhaps akin to calling running a separate adaptation from walking, despite the common entrainment of psycho-motor systems.

Mehr et al. then note that music is a human universal (so is, for instance, language), that music production and perception is complex (so is language), that it has a grammar-like structure (so does language), that it isn't random (neither is language), and that artificial intelligence (AI) engineers have difficulty to replicating it (ditto for language). At this point, music is starting to look like a duck.

Mehr et al. offer that motivations and abilities to perceive music appear early, that specific neural circuitry underlies music perception, and that deficits to specific circuitry impair music perception. None of this is surprising. What would be surprising would be to find that impairments causing tone-deafness didn't also impair linguistic cadence/tonality perception. Mehr et al. cite Norman-Haignere, Kanwisher, and McDermott (2015), who report that music and speech are captured by different neural component profiles. But, the label "music" could just as easily have been "prosody," and their findings viewed as evidence that different components of language are processed by different cortical circuits, much like edges and depth in vision perception. Finally, Mehr et al.'s claim that music is culturally ancient again begs the question of whether we are merely talking about a by-product of, say, language, because adaptations for language, too, are generally regarded as ancient.

Although the authors concede that none of the six lines of evidence alone dismisses the by-product hypothesis, we suggest that, even together, all six do not adequately motivate the search for an evolved adaptation. Additional evidence and theoretical rationale are required to convincingly argue that music is a separate adaptation, either for signaling coalitional formidability or for signaling joint attention. Next, we examine issues specific to each of these two putative functions.

First, with respect to coalitions, it is unclear why signals of formidability need be credible. Predators don't signal prey from afar. As Sun Tzu in *The Art of War* states: "All warfare is based on deception." In the context of coalitional antagonism, why should we expect coalitions to reliably signal their formidability when successful territorial defense (or, for that matter, appropriation) might best be accomplished by deceiving rather than informing the enemy?

An alternative function for music in the context of coalitional antagonism is suggested by Sun Tzu: "On the field of battle, the spoken word does not carry far enough: hence the institution of gongs and drums. Nor can ordinary objects be seen clearly enough: hence the institution of banners and flags." In this sense, music may serve to coordinate coalitional members in the context of intergroup antagonism, but music (or, for that matter, flag-waving) does not itself function as a strong *signal to enemies* of the group's *ability* to coordinate and, in turn, to enforce its interests. One does not easily imagine that soldiers of the Chinese army, upon encountering enemy infantry massed before them on the plain, would mutter to themselves, "Oh, shit. They've got gongs and drums." Nor would those soldiers overly concern themselves with the threat of banners and flags. Thus, we suspect that inferences of coalitional formidability from cues of coordination are not made as readily as the authors' coalitional signaling account suggests. Additional empirical evidence is needed, given that the only research cited in support (Fessler & Holbrook, 2016) relied upon indirect measures of perceived coalitional formidability.

Second, with respect to parent–infant interactions, it is once again unclear that music is decisively different from or superior to language in its ability to solve the adaptive problem of assessing caregiver attention. Cognitive mechanisms for inferring the direction and source of vocalizations, for inferring attention from vocal turn-taking, and for associating voice tones and volume with meaning and intent – all appear to be features of language and music alike. Consider motherese (not mentioned by Mehr et al.). Motherese solves the problem of infant-directed attention, but motherese does not represent a clear break from language into a distinctly musical realm. It is linguistic although emphasizing language's ability to exploit pitch and tone. As an available solution to the problem of infant-directed attention, motherese points to the strong overlap between language and music – the latter of which elaborates upon elements of the former. We suggest that the increasing complexity of human social structures over time enabled the production and perception of subtle shades of linguistic expressions, meanings, and intentions that could be variably deployed across an array of relationships, caregiver–infant interactions and coalitional allies being two prime examples. In short, the flexibility of language solves the problem of mental coordination.

To the claim that music represents a separate adaptation, we must therefore echo the words of Galileo, "E pur si quacks" (And yet it quacks).

Conflict of interest. None.

References

Fessler, D. M. T., & Holbrook, C. (2016). Synchronized behavior increases assessments of the formidability and cohesion of coalitions. *Evolution and Human Behavior*, *37*(6), 502–509.

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*(6), 1281–1296.

Signaling games and music as a credible signal

Massimo Lumaca[a], Elvira Brattico[a] and Giosuè Baggio[b]

[a]Department of Clinical Medicine, Center for Music in the Brain, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, 8000 Aarhus, Denmark and [b]Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology, 7941 Trondheim, Norway.
massimo.lumaca@clin.au.dk; elvira.brattico@clin.au.dk; giosue.baggio@ntnu.no

doi:10.1017/S0140525X20001016, e107

Abstract

The argument by Mehr et al. that music emerged and evolved culturally as a credible signal is convincing, but it lacks one essential ingredient: a model of signaling behavior that supports the main hypothesis theoretically and empirically. We argue that signaling games can help us explain how musical structures emerge as population-level phenomena, through sender–receiver signaling interactions.

We welcome the target article's main hypothesis: Music may originally have served as a credible signal, with diverse musical forms and features subsequently emerging via cultural evolution. What may be missing from the authors' analysis is an explicit model of the processes by which musical systems emerged and evolved. Such a model would build on the idea of signaling and would be applicable to simulate and study experimentally how patterns of musical behavior appear and change over time.

A widely-accepted proposal in cultural evolution theory is that population-level patterns partly emerge from individual cognitive biases amplified by cultural transmission (Kirby, Dowman, & Griffiths, 2007). This population-level thinking is now supported by a growing number of studies using theoretical or experimental models of cultural transmission, studied using diffusion chains or more complex societies of interacting agents (Mesoudi, 2011). These methods aim to capture microevolutionary mechanisms at work at small-scale and over short time periods, reproducing data that matches actual historical patterns. Cultural transmission experiments (CTEs) offer valuable means to empirically test theoretical predictions with human participants, while maintaining much of the rigor and factor control of theoretical models (Mesoudi & Whiten, 2008). In section 5.2 of their article, Mehr et al. do not mention this line of research, and they do not advocate using CTEs to test their claims on music evolution. To evaluate some of the hypotheses proposed in their article, Mehr et al. would need *a model of signaling behavior* that lends itself to experimental and computational analyses of how universal and diverse musical structures (Mehr et al., 2019; Savage, Brown, Sakai, & Currie, 2015) emerge from individual behaviors and their neurocognitive underpinnings. A model that meets this requirement, and that embodies the idea of music as a credible signal, is the signaling game (Lewis, 1969; Skyrms, 2010).

Signaling games in dyads or structured populations of senders and receivers have been widely applied to model coordination behaviors (Galantucci, 2009). Recently, our team has adapted two-player signaling games in dyads and linear diffusion chains (Moreno & Baggio, 2015; Nowak & Baggio, 2016) to test hypotheses on the cultural transmission and evolution of music (Lumaca & Baggio, 2016, 2017; Lumaca, Haumann, Vuust, Brattico, & Baggio, 2018, 2019). In two studies (Lumaca & Baggio, 2017; Lumaca et al., 2018), senders and receivers were arranged in linear diffusion chains of several generations each: Each game between two players modeled an interaction between two adjacent generations of learners. The signals were tone sequences, paired with affective meanings: The pairing emerged as a result of player coordination in a game. At the end of each game, the receiver (generation n) became the sender in the next game ($n + 1$). Senders were asked to transmit the musical code (signals and mappings to meanings) learned in the previous game. In cooperative signaling games, there is an incentive to signal honestly and credibly in order to coordinate more rapidly with the other player: In all our studies, receivers eventually learned the mapping that senders used in a game. However, more relevant from the point of view of the target article is evidence of convergent evolution. Small transmission errors, likely driven by individual biases, accumulated in the musical code. Each transmission chain developed its own "musical culture" based on patterns of melodic and rhythmic structure. Thus, we demonstrated experimentally that individual biases, brought out by intergenerational signaling, can lead to convergence toward attested musical patterns.

Modifications of the signaling games paradigm could be useful to test other hypotheses from the target article. One is the cumulative increase in the complexity and diversity of signals, particularly in groups where signalers have conflicting interests (sect. 5.2). Signaling games are flexible enough to accommodate several network structures – from simple dyads to games with many senders and receivers – and payoff structures – from shared to conflicting interests between signalers. To address the former hypothesis, one could organize senders and receivers into "microsocieties" (Baum, Richerson, Efferson, & Paciotti, 2004) of several interacting individuals, where player payoffs would either differ (experimental groups) or not (control groups). The generational progression would be recreated by replacing the longest-standing members of the groups with naive players. Finally, the complexity and diversity of signals could be quantified (Miton & Charbonneau, 2018) and compared between groups and across generations.

Importantly, some of our experimental results diverge from a core proposal by Mehr et al.: the music-specificity of cultural attractors (sect. 3.1). In two studies (Lumaca & Baggio, 2016; Lumaca et al., 2018), we used signaling games in combination with electroencephalogram (EEG) to test the idea that music adapts to auditory perception mechanisms (Trainor, 2015). We recorded participants' brain responses in an auditory oddball task, which evoked an ERP signature of auditory scene analysis (ASA): the mismatch negativity (MMN) (Näätänen, Gaillard, & Mäntysalo, 1978). Another day, participants played in one signaling game as receivers and in a subsequent game as senders of a musical code. We showed that individual MMN latencies, which reflect ASA efficiency, predict the degree of melodic (Lumaca & Baggio, 2016) and rhythmic structures (Lumaca et al., 2018) introduced in the code. These findings trace the origins of core musical structures to neural mechanisms of ASA, which are arguably phylogenetically older than human musicality and are fairly widely conserved across species.

We argue that Mehr et al. could take advantage of the signaling games model to refine, constrain, and empirically test their hypothesis on the origins of music as a credible signal. Our experiments are a highly simplified model of signaling behavior and music transmission, yet they tap into the essential mechanisms which we suspect are at work in the emergence and evolution of music as a cultural symbolic system. Ultimately, the study of music's origins demands a joint effort across different disciplines and methods, including behavior and neuroscience. But, a unification of methods and results is unlikely to happen in the absence of a model and paradigm that can guide research. Signaling games can take on such a unifying role, especially if we accept the idea that human symbolic systems, including music, are systems of culturally transmitted credible signals.

Acknowledgments. We thank Hella Kastbjerg for helping in revising the manuscript.

Financial support. Center for Music in the Brain is funded by the Danish National Research Foundation (DNRF117).

Conflict of interest. None.

References

Baum, W. M., Richerson, P. J., Efferson, C. M., & Paciotti, B. M. (2004). Cultural evolution in laboratory microsocieties including traditions of rule giving and rule following. *Evolution and Human Behavior*, *25*(5), 305–326.

Galantucci, B. (2009). Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, *1*(2), 393–410.

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 5241–5245.

Lewis, D. (1969). *Convention: A philosophical study.* Harvard University Press.

Lumaca, M., & Baggio, G. (2016). Brain potentials predict learning, transmission and modification of an artificial symbolic system. *Social Cognitive and Affective Neuroscience*, *11*(12), 1970–1979.

Lumaca, M., & Baggio, G. (2017). Cultural transmission and evolution of melodic structures in multi-generational signaling games. *Artificial Life*, *23*(3), 406–423.

Lumaca, M., Haumann, N. T., Vuust, P., Brattico, E., & Baggio, G. (2018). From random to regular: Neural constraints on the emergence of isochronous rhythm during cultural transmission. *Social Cognitive and Affective Neuroscience*, *13*(8), 877–888.

Lumaca, M., Kleber, B., Brattico, E., Vuust, P., & Baggio, G. (2019). Functional connectivity in human auditory networks and the origins of variation in the transmission of musical systems. *eLife*, *8*, e48710.

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*, *366*(6468), 957–970.

Mesoudi, A. (2011). *Cultural evolution: How Darwinian theory can explain human culture and synthesize the social sciences.* University of Chicago Press.

Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1509), 3489–3501.

Miton, H., & Charbonneau, M. (2018). Cumulative culture in the laboratory: Methodological and theoretical challenges. *Proceedings of The Royal Society B: Biological Sciences*, *285*(1879), 20180677.

Moreno, M., & Baggio, G. (2015). Role asymmetry and code transmission in signaling games: An experimental and computational investigation. *Cognitive Science*, *39*(5), 918–943.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–329.

Nowak, I., & Baggio, G. (2016). The emergence of word order and morphology in compositional languages via multigenerational signaling games. *Journal of Language Evolution*, *1*(2), 137–150.

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 United States of America*, *112*(29), 8987–8992.

Skyrms, B. (2010). *Signals: Evolution, learning, and information.* Oxford University Press.

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140089.

Why don't cockatoos have war songs?

Cody Moser, Jordan Ackerman, Alex Dayer,
Shannon Proksch and Paul E. Smaldino

Department of Cognitive and Information Sciences, University of California,
Merced, Merced, CA 95343, USA.
cmoser2@ucmerced.edu
jackerman2@ucmerced.edu
adayer@ucmerced.edu
sproksch@ucmerced.edu
psmaldino@ucmerced.edu; smaldino.com

doi:10.1017/S0140525X20001223, e108

Abstract

We suggest that the accounts offered by the target articles could be strengthened by acknowledging the role of group selection and cultural niche construction in shaping the evolutionary trajectory of human music. We argue that group level traits and highly variable cultural niches can explain the diversity of human song, but the target articles' accounts are insufficient to explain such diversity.

It is rare (but not uncommon) that music is performed alone or specifically for oneself. Most functional accounts of music evolution, therefore, rightfully place group functions such as bonding and coordination at the forefront of their hypotheses. The target articles by Savage et al. and Mehr et al. excel at providing functional group accounts for their hypotheses using phylogenetic and comparative accounts of animal vocal behavior. But, a question remains regarding the differentiation of human musicality from the music-like behaviors of animals. We want to ask not what features human music has *in common* with animal vocalizations, but ask instead, "why is human music *unique*?" That is, why do we see an increase in the diversity and flexibility of form–function links in human song compared to birds and nonhuman primates? We propose a simple answer to the dilemma by noting that it is not just human music which is unique, but human sociality, which may have had a fundamental role in the evolution of music. In short, the uniqueness and diversity of human music could be the result of the manner in which human songs are nested within complex and highly variable social and cultural environments.

Both articles compellingly point to comparative examples of music-like behaviors in other species in order to show continuity between nonhuman and human musicality. In the case of Mehr et al., phylogenetic examples from the primate kingdom are highlighted to emphasize the role that territorial calls may have had in shaping group songs. In Savage et al., examples from primates, whales, and birds are used to show that coordination of melodic, harmonic, and complex rhythmic patterns is less developed in these social species. Meanwhile, these coordinated actions serve as a sort of glue for the role of *communitas* in human sociality. Despite this difference, they also note that some birds demonstrate human-like beat perception and rhythmic abilities. One notable example of such a bird is Snowball, a sulfur-crested cockatoo who can famously entrain to a musical beat (Patel, Iversen, Bregman, & Schulz, 2009) and has recently showcased a diversity of spontaneous dance movements in response to music (Keehn, Iversen, Schulz, & Patel, 2019). Additionally, thrush nightingales, like humans, demonstrate cultural evolution of categorical rhythms (Roeske, Tchernichovski, Poeppel, & Jacoby, 2020). These nightingales also produce isochronous rhythms, a pattern important for synchronous coordination in human music and dance (however, nightingale rhythmic coordination is notably different from that of human music-making, Roeske et al., 2020). There is now some evidence of similar coordinated rhythmic abilities in primates (Gamba et al., 2016). In sum, both humans and nonhuman animals share similar, yet not identical, capacities for rhythm and synchrony, both fundamental features of human musicality.

If we share so many important music-like features with primates and birds, as in the case of Snowball, why is it that our *repertoire* is so much more diverse? A general musical toolbox as proposed in Savage et al.'s hypothesis is insufficient for explaining musical diversity, as the authors note themselves, stating, "Each feature may have been initially based on behavioral innovations … each innovation opened a new cognitive/musical niche selecting for independent specialization of relevant neural circuitry." Similarly, the more *specific* territorial defense feature of our primate ancestors as proposed by Mehr et al. is insufficient to explain the plethora of form–function links that are the hallmark of their theoretical approach. In both articles, the role that group selection plays in shaping form–functionality is largely downplayed. This approach is limiting, as group selection is essential when we begin to ask questions such as, "why don't cockatoos have war songs?"

This is where integrating an understanding of selection for group-level traits is critical (Richerson et al., 2016; Smaldino, 2014; Zefferman & Mathew, 2015). Cockatoos lack war songs because cockatoos lack *war* (see Hobson, 2020 on the individualistic nature of bird fights). Unlike the examples from both birds and primates, humans occupy a unique social niche characterized by both its productivity and recombination (cultural evolution) and its ability to create new problems and avenues for these processes (cultural niche construction). Although many birds indeed exhibit cultural evolution of their songs and material culture, as in the case of bowerbirds, and possess the same hallmarks as human song's "unique" features such as its incremental change, learned elements, and social preferences, the application of these features is largely tied to singular and highly specific functions such as mate choice or predator evasion. In the case of humans, form–function links in song are highly varied precisely because our "functions" vary along an extremely diverse social dimension.

Smaldino (2014) refers to many of these unique traits as emergent "group-level traits," which are those traits which "are properly defined only at the level of group organization." A timeline of the evolution of human music should certainly take into account the evolution of group-level traits, all the way from our basal primate origins to what Turchin (2016) has coined our "ultrasociety." Unlike primate and avian societies, human societies exhibit group structures that are both hierarchical and multidimensional, with differentiation within and between levels, and traits distinguishing these structures and levels (Moffett, 2019; Smaldino, 2019). The adaptive significance of these traits almost certainly had an effect on the evolution of human music diversity (related proposals have been suggested for the evolution of language – see Thompson, Kirby, & Smith, 2016). It is not unlikely that as human social life expanded the importance of culture in shaping human behavior did as well, with vocal plasticity both in the forms of speech and music finding its way into our social niches.

We believe that the accounts by both articles greatly expand our understanding of human music evolution and are a long awaited start to a serious conversation on the origins of music. However, both approaches would be enriched by granular attention to the unique social evolution of our species, particularly the way our complex social structure has shaped the cultural evolution of behavior – from kinship, to occupations, to social differentiation. The complex and highly variable social and cultural environments associated with human ultrasociality almost certainly had a functional effect on music evolution.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Gamba, M., Torti, V., Estienne, V., Randrianarison, R. M., Valente, D., Rovara, P., … Giacoma, C. (2016). The indris have got rhythm! Timing and pitch variation of a primate song examined between sexes and age classes. *Frontiers in Neuroscience*, *10*, 249.

Hobson, E. A. (2020). Differences in social information are critical to understanding aggressive behavior in animal dominance hierarchies. *Current Opinion in Psychology*, *33*, 209–215.

Keehn, R. J. J., Iversen, J. R., Schulz, I., & Patel, A. D. (2019). Spontaneity and diversity of movement to music are not uniquely human. *Current Biology*, *29*(13), R621–R622.

Moffett, M. W. (2019). *The human swarm: How our societies arise, thrive, and fall*. Basic Books.

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*(10), 827–830.

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., … Ross, C. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*, e30.

Roeske, T. C., Tchernichovski, O., Poeppel, D., & Jacoby, N. (2020). Categorical rhythms are shared between songbirds and humans. *Current Biology*, *30*(18), 3544–3555.

Smaldino, P. E. (2014). The cultural evolution of emergent group-level traits. *Behavioral and Brain Sciences*, *37*(3), 243.

Smaldino, P. E. (2019). Social identity and cooperation in cultural evolution. *Behavioural Processes*, *161*, 108–116.

Thompson, B., Kirby, S., & Smith, K. (2016). Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*, *113*(16), 4530–4535.

Turchin, P. (2016). *Ultrasociety: How 10,000 years of war made humans the greatest cooperators on earth*. Beresta Books.

Zefferman, M. R., & Mathew, S. (2015). An evolutionary theory of large-scale human warfare: Group-structured cultural selection. *Evolutionary Anthropology: Issues, News, and Reviews*, *24*(2), 50–61.

Sex and drugs and rock and roll

Steven Pinker

Department of Psychology, Harvard University, Cambridge, MA 02138, USA.
pinker@wjh.harvard.edu

doi:10.1017/S0140525X20001375, e109

Abstract

This article is extraordinarily rigorous and rich, although there are reasons to be skeptical of its theory that music originated to signal group quality and infant solicitude. These include the lack of any signature of the centrality of these functions in the distribution or experience of music; of a role for the pleasure taken in music; and of its connections with language.

As someone who accidentally kicked off two decades of theorizing about the evolution of music with a few pages of discussion in one book (Pinker, 1997), I've long been nonplussed at the fervor with which theoreticians have striven to show that music is a Darwinian adaptation. I had included that discussion partly to sharpen readers' conceptions of the criteria for an adaptation. After 500 pages that had argued for the adaptive basis of many psychological faculties – stereo vision, the recency effect, disgust, jealousy, and revenge – I wanted to show that not *everything* is an adaptation. Any rigorous adaptationist hypothesis had to go beyond a trait merely being commonplace and lay out independent signs of engineering design for attaining some goal that was a subgoal of inclusive fitness. Whereas, it's easy to reverse-engineer, say, language, stereo vision, or fear, the function of music is far from obvious. If the concept of adaptation is not to apply to everything, and hence to nothing, we must entertain the possibility that music instead is a pleasure technology: an application of human ingenuity to the evolutionarily dubious but proximately compelling goal of activating our pleasure circuits. In the case of music, these circuits might belong to language, auditory scene analysis, habitat selection, emotional calls, motor control, and perhaps some non-adaptive features of the auditory brain, such as proximity to other systems and their entrainment by the periodicity in auditory signals.

And yet while many scholars despise the idea that psychological traits are adaptations, blowing it off as a bunch of after-the-fact just-so stories, they are equally offended by the idea that music is *not* an adaptation, and so have offered hypotheses that are dead on arrival, such as that music evolved to bond the group or attract females. The reason for the discrepancy, I suspect, is that adaptation is not conceived of as a testable hypothesis from evolutionary biology but as an affirmation of how we value, deplore, or frame features of human nature. To say that music is an adaptation is to exalt its value; to say it is a by-product is a philistine denigration.

For these reasons, it's a pleasure to see Mehr et al. transcend all this wooliness in their superb article. After performing masterful necropsies on the bond-the-group and woo-the-ladies hypotheses, and raising reasonable criticisms of the by-product possibility, they propose a two-part hypothesis – credible signaling of coalition quality and of attention to infants – that satisfies the criteria for an adaptation and has impressive support from phylogenetic, ethnographic, genetic, and behavioral evidence. Maybe the theory is even true, although I think that it has some shortcomings.

First, it's not easy to see how these two very specific functions can be reconciled with the broad range of forms and contexts in which music is produced and enjoyed. If coalition quality and infant care are the two pillars, and everything else a set of cultural embellishments and extensions, we should see signs that those two functions are particularly robust, universal, archetypal, pervasive, and salient in the panoply of musical experience. But, that is exactly what was *not* found in Mehr et al.'s (2019) mammoth cross-cultural survey. It was not the case that music exemplifying the two proposed cores, such as war songs and lullabies, were universal, whereas the supposed extensions, such as love songs, healing songs, dance music, and other genres, were distributed more patchily, followed paths of historical influence rather than species-wide universality, or had less reliable acoustic signatures. Our major conclusion was that the four kinds of music spotlighted in the paper, together with 16 other genres were pretty much equally robust, distinctive, and universal: "Music is not a fixed biological response with a single prototypical adaptive function: It is produced worldwide in diverse behavioral contexts …."

In a similar vein, the contemporary phenomenology of music shows no signs of the core-plus-periphery structure their theory implies. I see no evidence that group-advertising genres such as anthems and team songs, together with lullabies, are the most popular or accessible musical genres, that listeners backslide to pondering formidable cliques or calm babies when they experience other kinds of music, or any other sign of centrality. Both the ethnography and the psychology imply that music involves a broad mapping between acoustic structures and human experience, with no obvious common reaction or instrumental benefit. We enjoy a diversity of musical forms equally, and with no characteristic outcome other than the pleasure we get as we listen.

This leads to my second reservation about the theory. The most blazingly obvious feature of music – people enjoy it – plays no role in the theory. Although the authors criticize the by-product hypothesis in generic terms, they don't focus on the specific version in which humans apply their know-how to tap sources of acoustic and motoric pleasure. Indeed, I find it hard to see how that hypothesis could *not* be true. If we're smart and our brains motivate us with pleasure, what could *stop* us from deploying our intelligence to gratify ourselves, as we do so flagrantly and maladaptively with non-procreative sex and recreational drugs (two examples for those of you, who, like me, don't even like cheesecake)?

My final reservation concerns something else that is conspicuous by its absence: language. Music, like language, lacks close homologs in other apes; has a hierarchical structure of phrases within phrases; has a complex rhythmic structure that matches that of language so uncannily that we can put words to music; feels like it's communicating something even when it isn't; and is composed of harmonically related frequencies found in no commonly experienced natural acoustic stimulus other than the human voice (and, less frequently, animal vocalization). And, as we showed in Mehr et al. (2019), all musical genres include lyrics. In the current theory, these are all coincidences.

Despite these zones of skepticism, I commend the authors on this tour de force, which elevates the topic to new levels of theoretical rigor and empirical richness.

Conflict of interest. None.

References

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*, *366*(6468), 957–970.

Pinker, S. (1997). *How the minds works*. Norton.

Ecological and psychological factors in the cultural evolution of music

Thom Scott-Phillips[a], Atsuko Tominaga[a] and Helena Miton[b]

[a]Department of Cognitive Science, Central European University, Oktober 6 utca 7, Budapest 1051, Hungary and [b]Santa Fe Institute, Santa Fe, NM 87501, USA.
scott-phillipst@ceu.edu
tominaga_atsuko@phd.ceu.edu
helena@santafe.edu
https://thomscottphillips.com
https://atsukotominaga.com
https://www.santafe.edu/people/profile/helena-miton

doi:10.1017/S0140525X20001181, e110

Abstract

The two target articles agree that processes of cultural evolution generate richness and diversity in music, but neither address this question in a focused way. We sketch one way to proceed – and hence suggest how the target articles differ not only in empirical claims, but also in their tacit, prior assumptions about the relationship between cognition and culture.

"Music" refers to and is studied as a set of cultural practices. As such, understanding its nature and origins requires some specification of how social interaction shapes, channels, and scaffolds cultural practices. Yet, both articles only emphasize the role of global, psychological factors in the cultural evolution of music – and in doing so they risk obscuring key aspects of the problem. Savage et al. highlight stress reduction, cooperation, group identification, and other such factors that contribute to social bonding. Mehr et al. remark that features of music "will interact with evolved capacities for nonmusical traits… increasing variability in music." Here, "evolved" means (roughly) global, psychological factors, and indeed such factors clearly help to explain why, for instance, some different song types (dance, lullabies, healing, and love) exhibit cross-cultural similarity and robust form–function relations, discernible by listeners of other cultural backgrounds (Mehr, Singh, York, Glowacki, & Krasnow, 2018). Nevertheless, factors relevant to cultural evolution can also be local rather than global, and ecological rather than psychological. In Table 1, we list a few specific examples of each type relevant to the cultural evolution of music.

As a specific case, consider the evolution of violin design (Tai, Shen, Lin, & Chung, 2018). Early violins, in particular those developed by Andrea Amati (1505–1577), had acoustic properties characteristic of male singers. However, as female voices became increasingly common and popular in Baroque music in the early 1600s, violin designs changed in ways that complemented their sound. In particular, the designs of Antonio Stradivari (1644–1737), widely considered to be the gold standard of violin making, have acoustic properties that closely resemble those of female singers (*ibid.*). It seems, then, that the acoustic properties of the female voice were an important ecological factor in the evolution of violin design (see also Miton, Wolf, Vesper, Knoblich, & Sperber, 2020, for experimentally induced demonstration of how ecological factors can influence the cultural evolution of rhythms).

We expect the authors of both target articles would agree that existing musical behaviours, and associated mental representations, interact not only with universal features of the human mind, but also with non-universal features of minds, and features of the world beyond the mind. Moreover, the sheer number and range of such interactions is how massive cultural variation can emerge from human psychic unity. However, the absence of any clear exposition of these points from their papers is, we believe, an oversight that risks, unnecessarily, downplaying important aspects of the problem.

In any case, it is revealing to ask in what ways, if any, these cultural evolutionary considerations might inform our understanding of the origins of music. Here, the two target articles seem to differ. At one extreme, it may be that music's origins are largely independent of its cultural evolution. This is (more or less) the approach adopted by Mehr et al., who distinguish between the "proper domain" and the "actual domain" of musicality and its associated cognitive capacities, and attempt to distinguish the contents of each (see also e.g., Honing, ten Cate, Peretz, & Trehub, 2015). After all, cultural items frequently fall inside the actual domain but outside the proper domain of the relevant cognitive capacities, such as, for instance, portraits, caricatures, and masks, all of which trigger cognitive processes for facial recognition, but none of which fall within the proper domain of those cognitive processes (Sperber & Hirschfeld, 2004). At the other extreme, it may be that music's origins are irreducibly intertwined with its cultural evolution. By placing significant weight on gene-culture coevolution, in which musical practices impact biological

Table 1. (Scott-Phillips et al.) Factors relevant to cultural evolution, with examples relevant to the cultural evolution of music

	Global	Local
Psychological	(1) Pleasure derived from predictive processes in perception (e.g., Cheung et al., 2019; Salimpoor, Zald, Zatorre, Dagher, & McIntosh, 2015) (2) Dispositions towards particular rhythms, possibly shared with other species (e.g., Koelsch, 2011) (3) Stress reduction, cooperation, group identification, and other such factors that contribute to social bonding (e.g., Savage et al., target article)	(1) Knowledge of already existing musical genres and practices in the local community (2) Prosodic grouping based on a mother's native language (e.g., Abboub, Nazzi, & Gervain, 2016)
Ecological	(1) Natural sounds of the human voice (such as in the cultural evolution of violin design: see main text), and other features of the body (2) Motor and other physical constraints on melodic vocalization, possibly shared with other species (e.g., Savage, Tierney, & Patel, 2017; Tierney, Russo, & Patel, 2011) (3) Materials common and generally available in human environments e.g., animal hides, bones, and stones	(1) Already existing instruments, means of notation, and other non-psychological aspects of music production in the local community (e.g., Strayer, 2013) (2) Locally specific materials e.g., bamboos and seashells (e.g., Wegst, 2008)

The terminology of local and global, ecological and psychological, is taken from Cultural Attraction Theory (Sperber, 1996; Scott-Phillips, Blancke, & Heintz, 2018).

evolution because of their impact on social bonding, which in turn feeds back on culture, Savage et al. tend more towards this pole.

The different empirical claims developed in the two target articles thus seem to reflect, in part, different tacit and prior assumptions about the relationship between cognition and culture. To what extent is any particular aspect of the human cognitive phenotype (e.g., musicality) shaped, by natural selection, *in order to* acquire specific cultural phenomena (e.g., music)? This is a major question for the evolutionary human sciences, relevant not just to music but numerous cultural domains (language, religion, and politics, *inter alia*). A range of answers are possible. In the specific case of music, we do not believe that any answer will achieve strong scientific consensus unless and until there is a more detailed analysis of the factors that shape its cultural evolution, as it proliferates from its proper domain to its (larger) actual domain. It is certain that these factors are not exclusively global and psychological.

Financial support. TSP was financially supported by the European Research Council, under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 609819 (Somics project). AT was supported by European Research Council grant agreement no. 616072, JAXPERTISE. HM was financially supported by a Complexity Postdoctoral Fellowship at the Santa Fe Institute.

Conflict of interest. None.

References

Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain & Language*, *162*, 46–59.

Cheung, V. K., Harrison, P. M., 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.

Honing, H., ten 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*, *370*(1664), 20140088.

Koelsch, S. (2011). Toward a neural basis of music perception – A review and updated model. *Frontiers in Psychology*, *2*, 110.

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

Miton, H., Wolf, T., Vesper, C., Knoblich, G., & Sperber, D. (2020). Motor constraints influence cultural evolution of rhythm. *Proceedings of the Royal Society B*, *287* (1937), 20202001.

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.

Savage, P. E., Tierney, A. T., & Patel, A. D. (2017). Global music recordings support the motor constraint hypothesis for human and avian song contour. *Music Perception*, *34* (3), 327–334.

Scott-Phillips, T., Blancke, S., & Heintz, C. (2018). Four misunderstandings about cultural attraction. *Evolutionary Anthropology*, *27*(4), 162–173.

Sperber, D. (1996). *Explaining culture*. Blackwell.

Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, *8*(1), 40–46.

Strayer, H. R. (2013). From neumes to notes: The evolution of music notation. *Musical Offerings*, *4*(1), 1–14.

Tai, H. C., Shen, Y. P., Lin, J. H., & Chung, D. T. (2018). Acoustic evolution of old Italian violins from Amati to Stradivari. *Proceedings of the National Academy of Sciences*, *115* (23), 5926–5931.

Tierney, A. T., Russo, F. A., & Patel, A. D. (2011). The motor origins of human and avian song structure. *Proceedings of the National Academy of Sciences*, *108*(37), 15510–15515.

Wegst, U. G. (2008). Bamboo and wood in musical instruments. *Annual Review of Materials Research*, *38*, 323–349.

Rapid dissonant grunting, or, but why does music sound the way it does?

Beau R. Sievers[a] and Thalia Wheatley[b]

[a]Psychology Department, Harvard University, Cambridge, MA 02138, USA and [b]Psychological and Brain Sciences, Dartmouth and Santa Fe Institute, Hanover, NH 03755, USA.
beau@beausievers.com, thalia.p.wheatley@dartmouth.edu
www.wheatlab.com
www.beausievers.com

doi:10.1017/S0140525X20001284, e111

Abstract

Each target article contributes important proto-musical building blocks that constrain music as-we-know-it. However, neither the credible signaling nor social bonding accounts elucidate the

central mystery of why music sounds the way it does. Getting there requires working out how proto-musical building blocks combine and interact to create the complex, rich, and affecting music humans create and enjoy.

The social bonding and credible signaling hypotheses share a basic strategy: To identify a set of proto-musical building blocks that, if connected in the right way, could be shaped by cultural evolution into music as-we-know-it. To this end, both hypotheses identify functional fixedness between features of sounds and features of the agents producing them: An uncooperative group cannot produce synchronized sound, a caregiver singing to an infant cannot also be talking to someone else, and so on. We endorse this approach, but argue that neither hypothesis fixes enough functions to explain a central mystery of musical evolution: But, why does music sound like *that?* Furthermore, we note that neither hypothesis explains how to get from the simple signaling of brute facts to the complex semantic playground of music as-we-know-it. We suggest the cognitive capacity for domain-general compositional thinking may have played an important role.

To understand why the account of functional fixedness needs to be elaborated, consider the case of the lullaby. On the credible signaling hypothesis, the form of the lullaby is fixed by the functional requirements of signaling attention, proximity, and responsiveness of the caregiver. But taken alone, this does not explain why lullabies across the world are slow and consonant (Mehr et al., 2019). Given the spare set of requirements, lullabies could just as well be mostly rapid dissonant grunting, yet to our knowledge no culture has adopted this strategy.

To better account for the form of the lullaby, we propose that connections between sound, movement, and emotion are an additional source of functional fixedness. If one function of emotion is to bias agents toward context-appropriate action (Frijda, Kuipers, & Ter Schure, 1989), then the path from action predisposition to functional fixedness is short. If high-arousal states make characteristically low-arousal movement difficult (and vice versa), and if high-arousal movements produce sounds that distinguish them from low-arousal movements, then sounds should credibly signal both the movement that produced them and the sound producer's state of mind. Rapid dissonant grunting is, therefore, bad lullaby material – not because it fails to signal attention, proximity, or responsiveness, but because it signals that the caregiver is in a high-arousal state inappropriate for the context (e.g., bedtime).

Supporting this account, we have shown that music and movement share a dynamic structure, such that the same combinations of music and movement features express the same emotions across cultures (Sievers, Polansky, Casey, & Wheatley, 2013). Furthermore, harsh timbres and spiky movement contours (both quantified using the spectral centroid) are reliably used in the expression and perception of high-arousal emotion (Sievers, Lee, Haslett, & Wheatley, 2019). Interestingly, Filippi et al. (2017) have shown that harsh timbres are used to express high-arousal by many species of terrestrial vertebrate, suggesting the sound–emotion connection is evolutionarily ancient, fitting the proposed timeline of the credible signaling hypothesis.

The social bonding hypothesis faces a similar challenge, as predictability cannot do the job of fixing musical form all on its own. Because predictability is, in principle, separable from sonic features such as loudness and harshness, two pieces of music could be similar in terms of overall predictability but otherwise completely different. Our proposed approach should work here, too: Expand the account of inferential roles and functional fixedness to accommodate more of what makes music matter, even at the cost of weakening the claim that music has a singular function.

The credible signaling and social bonding hypotheses each describe building blocks to be shaped into music by cultural evolution. For the social bonding hypothesis, the building blocks are interlocking neurobiological reward-learning systems. For the credible signaling hypothesis, the building blocks are signaling systems selected to fit specific inferential roles. Both hypotheses face the same problem: If mere exposure to a sonic stimulus can provoke appropriate behavior, why is music so elaborate, such a parade of semantic excess? Although signals communicate simple meanings (the shake of the rattlesnake's tail means you are in danger), music communicates *complexes* of meaning, supporting a semantics with a richness different from but rivaling that of language (Schlenker, 2019). This may also be true of dance (Charnavel, 2019; Patel-Grosz, Grosz, Kelkar, & Jensenius, 2018). What could possibly get us from the mechanistically and computationally simple raw material of reward-learning and signaling to the confoundingly meaningful playground of music as-we-know-it?

Building blocks are meant to be combined. We suggest that the capacity for compositional thought – the ability to "make infinite use of finite means" (von Humboldt, 1836/1999) by recombining parts into novel arrangements – had an important role in the transition from proto-music to music. If so, the "shapes" of the proto-musical building blocks should matter – each block must be interoperable with the others. This interoperability could take many forms, ranging from accessibility to a global workspace (Baars, 1993) or a symbolic representational system (Fodor, 1975) to participation in a network of interacting cognitive maps (Bottini & Doeller, 2020; Parkinson, Liu, & Wheatley, 2014). Hinting at the latter possibility, we have shown that the brain represents emotional music and movement using a similar format, possibly to facilitate comparison across sensory modalities (Sievers et al., 2018).

Critically, the requirement of proto-musical interoperability poses different challenges for each hypothesis. The social bonding hypothesis must avoid the trap of behaviorism, showing how the simplistic stimulus-response characteristics of reward-learning systems could be bootstrapped to build a rich inferential semantics. By contrast, the credible signaling hypothesis must avoid the trap of massive modularity, showing how a motley of signaling systems, each evolved to narrowly serve different inferential roles, could be harmonized, placed in a shared context, and used to express a wider range of meanings.

Both target articles, here, elucidate important proto-music building blocks that functionally constrain music as-we-know-it. But, neither explains why music *sounds the way it does*. Getting from proto-music to music as-we-know-it requires not only knowing what the building blocks are, but also how they fit together, combining and interacting to create the deeply affecting, complex and semantically rich music humans enjoy.

Financial support. This short commentary was not funded by any institution.

Conflict of interest. None.

References

Baars, B. J. (1993). *A cognitive theory of consciousness*. Cambridge University Press.

Bottini, R., & Doeller, C. F. (2020). Knowledge across reference frames: Cognitive maps and image spaces. *Trends in Cognitive Sciences*, *8*, 606–619.

Charnavel, I. (2019). Steps towards a universal grammar of dance: Local grouping structure in basic human movement perception. *Frontiers in Psychology*, *10*, 1364.

Filippi, P., Congdon, J. V., Hoang, J., Bowling, D. L., Reber, S. A., Pašukonis, A., … Newen, A. (2017). Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: Evidence for acoustic universals. *Proceedings of the Royal Society B*, *284*, 1859.

Fodor, J. A. (1975). *The language of thought*. Thomas Y. Crowell.

Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212–228.

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*, *366*(6468), 957–970.

Parkinson, C., Liu, S., & Wheatley, T. (2014). A common cortical metric for spatial, temporal, and social distance. *Journal of Neuroscience*, *34*, 1979–1987.

Patel-Grosz, P., Grosz, P. G., Kelkar, T., & Jensenius, A. R. (2018). Coreference and disjoint reference in the semantics of narrative dance. *Proceedings of Sinn und Bedeutung*, *22*, 199–216.

Schlenker, P. (2019). Prolegomena to music semantics. *Review of Philosophy and Psychology*, *10*, 35–111.

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

Sievers, B., Parkinson, C., Kohler, P. J., Hughes, J., Fogelson, S. V., & Wheatley, T. (2018). Visual and auditory brain areas share a neural code for perceived emotion. *BioRxiv*, 254961.

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

von Humboldt, W. (1836/1999). *On language: On the diversity of human language construction and its influence on the mental development of the human Species*. Cambridge University Press.

Functional and evolutionary parallels between birdsong and human musicality

Kate T. Snyder and Nicole Creanza

Department of Biological Sciences, Vanderbilt University, Nashville, TN 37240, USA.
katherine.t.snyder@vanderbilt.edu
nicole.creanza@vanderbilt.edu
katetsnyder.com
creanzalab.com

doi:10.1017/S0140525X20001661, e112

Abstract

Here, we compare birdsong and human musicality using insights from songbird neuroethology and evolution. For example, neural recordings during songbird duetting and other coordinated vocal behaviors could inform mechanistic hypotheses regarding human brain function during music-making. Furthermore, considering songbird evolution as a model system suggests that selection favoring certain culturally transmitted behaviors can indirectly select for associated underlying neural functions.

What is the function of music, how did it originate and evolve, and how has it changed us over time? The proposed mechanisms underlying the evolution of musicality in both Savage et al. and Mehr et al. are supported with cogent arguments and convincing evidence. Both discuss important parallels between human music and animal models of vocal learning and coordination. Complementing this discourse, we suggest additional insights drawn from birdsong neuroethology and evolution that support the broad ideas presented in these articles.

Birdsong has evolved under selection pressures associated with multiple functions, including sexual selection, within-species interactions, and species recognition. Songbird behaviors may provide insights into neural mechanisms of coordination between individuals during music production that are currently difficult to study in humans. For example, duetting is a relatively rare phenomenon in birds, involving two individuals producing a temporally coordinated song consisting of overlapping or alternating phrases (Hall, 2009). Recently, researchers have simultaneously recorded vocalizations and electrophysiological signals in duetting pairs of birds, finding correlated neural activity between paired individuals. During a duet, individuals modulated their vocalization tempo to fit into the silences between their partner's vocalizations, just as humans can use auditory feedback to match an external tempo. Interestingly, these birds exhibited premotor neural activity both while singing their own portions of the song and during their partner's parts of the song (Hoffmann et al., 2019). A similar phenomenon is observed in zebra finches producing single-note calls to establish social contact: Individuals dynamically modulate the tempo of their calls to fall in the silences between their partner's calls and, again, inhibition of call production is associated with premotor neural activity that anticipates a partner's vocalization (Benichov & Vallentin, 2020; Benichov et al., 2016). In other words, remaining silent to accommodate a vocal partner is an active neural process mediated by both auditory and social stimuli.

Songbirds also provide an example of evolution of auditory preferences based on culturally transmitted (i.e., learned) signals. In nature, juvenile songbirds are exposed to many species' songs. To be recognized as the correct species and successfully attract a mate, birds must identify and selectively learn conspecific songs. We posit that this might occur similarly to how, according to Mehr et al., human babies attune to rhythmic vocalizations. In songbirds, these relatively coarse song-selection filters are thought to be innate, because studied species of songbirds react differently to conspecific vocalizations while still in the egg or nest (Colombelli-Négrel et al., 2012; Hudson, Creanza, & Shizuka, 2020; Hudson & Shizuka, 2017) and selectively learn conspecific song without prior exposure to it (Colombelli-Négrel et al., 2012; Marler & Peters, 1977; Soha & Marler, 2000). Selectivity in song learning varies widely between species and is based on species-specific song features, which can be sound properties of individual syllables (e.g., timbre or pitch modulation), or the pattern of these syllables. For example, white-crowned sparrows attune to a pure-tone whistle at the beginning of songs, whereas swamp sparrows, which sing a trilled song of multiple frequency sweeps, selectively learn conspecific syllables regardless of temporal organization (Marler & Peters, 1977; Soha & Marler, 2000). In contrast, zebra finches raised by Bengalese finches learn Bengalese finch syllables but transpose these syllables to match a typical zebra finch song temporal pattern (Araki, Bandi, & Yazaki-Sugiyama, 2016). These findings suggest that, for zebra finches, there is an innate template for rhythm of syllable production. Understanding how the brains of young songbirds selectively attune to certain auditory patterns may shed light on how human brains attune to rhythm.

In addition to their diversity of innate auditory preferences, songbirds exhibit wide variation in the duration of their song-learning window. We recently modeled and analyzed the evolution of songs alongside the duration of song learning; we found that selection for more elaborate songs can drive the evolution of the capacity to learn throughout life (Creanza, Fogarty, & Feldman, 2016; Robinson, Snyder, & Creanza, 2019). We propose that this evolutionary paradigm in songbirds – that selection on a learned trait can drive evolution of the brain – provides a possible example of the phenomenon depicted in Savage et al. (Fig. 2, left panel): Musical features can act as an intermediary between social functions and their neurobiological underpinnings.

Savage et al. describe musicality as a "cognitive toolkit." How might the framing of musicality as a set of tools affect our understanding of its evolution? Our lab modeled the evolution of birdsong features as culturally transmitted functional traits, similar to tools, wherein learners aim to imitate proficient tutors (Hudson & Creanza, 2021). Like other fitness-altering cultural traits, functional signals based on rhythmicity or pitch modulation could have gradually become more complex if learners preferentially choose tutors with complex signals. Over time, the cultural development of functional signals could elevate the minimum cognitive baseline to recognize and reproduce these signals, thereby influencing brain evolution to favor attention to and learning capacity for these acoustic features. In this context, elements of musicality might have been under selection for purposes other than the umbrella explanation of "social bonding." Savage et al. describe the neural synchronization between auditory and motor brain regions during rhythm perception to explain the origins of dance, but only briefly mention other functions of coordinated behavior. Could rhythmic movement have functioned as a fitness-enhancing tool? Rhythmicity allows for synchronization of actions between individuals and for individuals to accurately predict the actions of others. It is thus conceivable that the development of rhythmicity would have facilitated a large repertoire of coordinated behaviors that could have impacted group survival.

Finally, both target articles discuss the hypothesis that musicality evolved through sexual selection, concluding that it is inadequate to explain the evolution of musicality. However, this hypothesis is framed from an intraspecific mate selection perspective, where females choose males with the most attractive musical displays. Studying the evolution of birdsong and its role in species recognition suggests another perspective: in our evolutionary past, could musicality have served an interspecific function, mediating the interactions between the ancestors of *Homo sapiens* and other hominin lineages? Although musicality appears to be uniquely human among extant species, Mehr et al. conjecture that the basic elements of musicality are ancestral to all primates – just as song is to all songbirds. Did musicality contribute to species recognition when our ancestors formed groups or selected mates, perhaps before the emergence of language? We are unable to know how much musical predisposition we shared with our evolutionary cousins – those we interbred with, and those we didn't. However, considering songbirds as a model system suggests that the evolutionary implications of musicality need not be limited to interactions within our own species.

Financial support. This study was supported by the NSF (NC & KS, grant number BCS-1918824); and Vanderbilt University (KS & NC).

Conflict of interest. None.

References

Araki, M., Bandi, M. M., & Yazaki-Sugiyama, Y. (2016). Mind the gap: Neural coding of species identity in birdsong prosody. *Science (New York, N.Y.)*, *354*, 1282–1287.

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A., & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology*, *26*, 309–318.

Benichov, J. I., & Vallentin, D. (2020). Inhibition within a premotor circuit controls the timing of vocal turn-taking in zebra finches. *Nature Communications*, *11*, 221.

Colombelli-Négrel, D., Hauber, M. E., Robertson, J., Sulloway, F. J., Hoi, H., Griggio, M., & Kleindorfer, S. (2012). Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. *Current Biology*, *22*, 2155–2160.

Creanza, N., Fogarty, L., & Feldman, M. W. (2016). Cultural niche construction of repertoire size and learning strategies in songbirds. *Evolutionary Ecology*, *30*, 285–305.

Hall, M. L. (2009). A review of vocal duetting in birds. *Advances in the Study of Behavior*, *40*, 67–121. doi: 10.1016/s0065-3454(09)40003-2.

Hoffmann, S., Trost, L., Voigt, C., Leitner, S., Lemazina, A., Sagunsky, H., … (2019). Duets recorded in the wild reveal that interindividually coordinated motor control enables cooperative behavior. *Nature Communications*, *10*, 2577.

Hudson, E. J., & Creanza, N. (2021). Ornament, armament, or toolkit? Modelling how population size drives the evolution of birdsong, a functional cultural trait. *bioRxiv*, 2021.04.29.442039.

Hudson, E. J., Creanza, N., & Shizuka, D. (2020). The role of nestling acoustic experience in song discrimination in a sparrow. *Frontiers in Ecology and Evolution*, *8*, 99. doi: 10.1101/756445.

Hudson, E. J., & Shizuka, D. (2017). Introductory whistle is sufficient for early song recognition by golden-crowned sparrow nestlings. *Animal Behaviour*, *133*, 83–88. doi: 10.1016/j.anbehav.2017.09.018.

Marler, P., & Peters, S. (1977). Selective vocal learning in a sparrow. *Science (New York, N.Y.)*, *198*, 519–521.

Robinson, C. M., Snyder, K. T., & Creanza, N. (2019). Correlated evolution between repertoire size and song plasticity predicts that sexual selection on song promotes open-ended learning. *eLife*, *8*, 44454. doi: 10.7554/eLife.44454.

Soha, J. A., & Marler, P. (2000). A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Animal Behaviour*, *60*, 297–306.

Making music: Let's not be too quick to abandon the byproduct hypothesis

Steve Stewart-Williams

School of Psychology, University of Nottingham Malaysia, Jalan Broga, 43500
Semenyih, Selangor Darul Ehsan, Malaysia.
steve.stewart-williams@nottingham.edu.my; http://www.stevestewartwilliams.com/

doi:10.1017/S0140525X20001119, e113

Abstract

It is premature to conclude that music is an adaptation. Given the danger of overextending the adaptationist mode of explanation, the default position should be the byproduct hypothesis, and it should take very strong evidence to drag us into the adaptationist camp. As yet, the evidence isn't strong enough – and the proposed adaptationist explanations have a number of unresolved difficulties.

Mehr et al. and Savage et al. have both put forward interesting and very reasonable adaptationist accounts of music – or more precisely, of certain aspects of musicality and musical behavior. I'm more sympathetic to such accounts than I was before. On balance, though, I think it's still premature to conclude that music is an adaptation, and more plausible to think that it's a byproduct. There are three main reasons for this.

First, a lot of the evidence adduced in favor of adaptationist explanations of music is equally amenable to a byproduct explanation. The cross-cultural universality of music is *consistent* with the claim that music is an adaptation – but it's also consistent with the claim that it's a byproduct of other adaptations that are universal but not music-specific (e.g., emotional responsiveness to the prosody in speech, which cultures might independently "learn" to trigger with melodies). The complex design evident in music *could* come from biological evolution – but it could also come from cumulative cultural evolution; after all, smart phones and bureaucracies exhibit complex design as well, but are clearly not adaptations. Children take to music early and easily – but they also take to iPads and TV; sometimes ease of acquisition is a result of culture evolving for our minds, rather than the other way around. Damage to certain areas of the brain impairs the ability to make or appreciate music – but none of these areas is involved exclusively in music, and it's possible that the areas in question evolved primarily for their non-musical functions (which are presumably also impaired by damage to those areas). Music-like abilities in nonhuman animals show that traits of that kind *can* evolve – but they don't show that they necessarily did evolve in our species, as human culture can sometimes independently discover traits that evolved in other animals: The fact that leaf-cutter ants engage in something akin to agriculture doesn't imply that human agriculture is an adaptation; similarly, the fact that various nonhuman animals produce auditory displays doesn't imply that human *music* is an adaptation. In short, much of the evidence is ambiguous. Given the danger of overextending the adaptationist mode of explanation, the byproduct approach seems like the safer default position in lieu of more decisive evidence.

Second, the byproduct approach has a number of advantages over its adaptationist rivals. Uncontroversial adaptations, such as arms and the basic motivations, are found in all typically developing human beings and are reasonably similar across cultures, subcultures, and historical periods. Music, in contrast, varies greatly from place to place and from time to time, and many people spend little time making or consuming it. These facts are easier to square with a byproduct explanation than an adaptationist one. Even if one argues that certain *core* features of music are found in every culture, it remains the case that plenty of individuals within those cultures devote little time to music, whereas almost every individual has arms and the basic motivations. And even if one argues that, in *traditional* cultures, almost every individual devotes substantial time to music, the fact that many individuals in modern cultures do not is still surprising on an adaptationist account – after all, even in modern cultures, every typically developing human being uses language frequently, and it would be surprising on an adaptationist account of language if this were not the case.

Third and finally, the adaptationist accounts of music proposed in this dual treatment face a number of challenges that byproduct explanations do not. If stronger social bonds are adaptive, as Savage et al. argue, why not select directly for a tendency to bond more strongly, rather than a tendency to make and enjoy rhythmically patterned pitch-sequences and to bond with others who do the same? Regarding Mehr et al.'s account, does it seem plausible that raiding parties would be less inclined to attack a group that kept perfect time than an equivalently fierce group whose rhythms were slightly off, or that such a strategy would be particularly useful? Keeping time isn't important in chimpanzee territorial displays, so the closest animal analogy doesn't support the idea. Is music-making prowess a reliable way to assess a group's potential as allies? People could make beautiful music together but be hopeless at hunting, making tools, or doing anything else that might make an alliance valuable. Why not assess the valuable abilities directly, rather than assessing people's musical chops? If rhythm evolved for territorial signaling, why aren't men notably more rhythmical than women, given that men have historically done the bulk of the territorial displaying and defense? If melody evolved for infant-directed song, why aren't women notably more melodic than men, given that women have historically done the bulk of the infant care? Although some studies suggest such differences (e.g., Miles, Miranda, & Ullman, 2016), the broader literature is mixed and it's certainly not obvious that the sexes differ much in these domains. Is infant-directed song a reliable signal of commitment in any evolutionarily meaningful way? It "tells" the baby that it has the parent's undivided attention at that particular moment, while the parent is singing the song. However, the fact that it has their attention in a context where it isn't especially costly to the parent doesn't guarantee that the parent will prioritize the baby if and when difficult trade-offs need to be made – for example, if the parent has to choose to invest either in the baby or in one of the baby's siblings. A peacock can't grow a decent tail unless it's in good condition; in contrast, it's easy enough to sing a baby a song then withdraw support later on, if one's circumstances change.

I don't claim that these difficulties are necessarily insurmountable, and I concede that some of the evidence presented in favor of an evolved contribution to human musicality is at the very least suggestive. However, the difficulties do hint that it's premature to accept an adaptationist account at this stage – and if I had to make a bet today, my money would be on the byproduct approach.

Financial support. The author received no specific grant for this work from any funding agency.

Conflict of interest. None.

Reference

Miles, S. A., Miranda, R. A., & Ullman, M. T. (2016). Sex differences in music: A female advantage at recognizing familiar melodies. *Frontiers in Psychology*, *7*, 278–278. https://doi.org/10.3389/fpsyg.2016.00278.

Pre-hunt charade as the cradle of human musicality

Szabolcs Számadó[a,b,c]

[a]Department of Sociology and Communication, Budapest University of Technology and Economics, Egry J. u. 1., Budapest, 1111, Hungary; [b]Centre for Social Sciences (TK CSS) "Lendület" Research Center for Educational and Network Studies (CSS-RECENS), Tóth Kálmán u. 4, Budapest, 1097, Hungary and [c]Evolutionary Systems Research Group, Centre for Ecological Research, Klebelsberg Kuno u. 3, Tihany 8237, Hungary.
szamszab@ludens.elte.hu; szamado.szabolcs@tk.mta.hu

doi:10.1017/S0140525X20001077, e114

Abstract

Human language and human music are both unique communication systems that evolved in the human lineage. Here, I propose that they share the same root, they evolved from an ancestral communication system yet to be described in detail. I suggest that pre-hunt charade was this shared root, which helped organize and coordinate the hunt of early hominins.

In a twin submission, Mehr, Krasnow, Bryant, and Hagen and Savage et al. offer two interesting yet very distinct accounts for the evolution of human music and musicality. Before we dive into the evaluation of these different proposals, it is worth to clarify what is music, that is, what needs to be explained. Human music is a unique communication system unmatched in nature. We can communicate the whole range of human emotions with music, we can signal our commitment, our attention, intentions, we can coordinate work with music. Music relates to animal "song" similar to human language relates to the "language of the bees." Interestingly enough, the species that evolved this unique communication system also evolved another unique communication system: human language.

It is hard not to notice the similarities between ideas proposed to explain the origin of human language on the one hand and music on the other hand. All the functions discussed by the two groups were proposed to explain the origin of language, as well. Famously, the "by product" theory was proposed by Chomsky and colleagues (Hauser, Chomsky, & Fitch, 2002), perhaps even more famously the "vocal grooming" theory was originally proposed by Dunbar (1993) in the context of language evolution, the signalling of mate quality was proposed by Miller (2011), bonding of female coalitions was proposed by Knight (1998), and finally bonding between mother and her child was proposed by Falk (2004) as a selective force to explain the evolution of human language (see Számadó & Szathmáry, 2006 for a review).

The similarities do not end here. Both fields aim to explain the origin of a unique communication system found only in humans, both fields aim to recreate a series of events that happened hundreds of thousands if not million years ago and both fields have to cope with the lack of direct evidence (i.e., lack of direct fossil record on language or music). It is clear that because of the scarcity of relevant information there is no agreement how to evaluate these proposals. One group (Mehr et al.) argues in favour of music as a credible signal of cooperative intent, whereas the other group (Savage et al.) argues in favour of music as a general bonding device replacing "ancient," presumably less efficient bonding mechanisms (such as physical grooming). Both proposals have its merit as both groups suggest a function that can be efficiently performed by music. Yet, this is a weakness of these proposals as well. Deacon (1993) criticized such an approach in his response to Dunbar's (1993) original idea: "Language makes X more efficient, therefor selection for X explains the origin of language" (Deacon, 1993, p. 699). The problem is that both music and language are very good at making many aspects of our life more efficient, therefore this cannot be the sole criteria.

How to move forward? One way to cope with this problem is to set up a list of criteria that potential solutions have to fulfil. Such a list was proposed for human language: honesty, groundedness, power of generalization, and uniqueness (Számadó & Szathmáry, 2006). Similar lists were proposed later on (Bickerton, 2009; Hurford, 2011). Although I am not claiming that it solved the problem of language evolution, yet it could bring us closer to the solution by making it clear what needs to be solved in the first place. Another way is formal modelling of evolutionary trajectories (e.g., Zachar, Szilágyi, Számadó, & Szathmáry, 2018 on the origin of mitochondria). Although it is admittedly difficult, and especially so in a situation in the lack of direct evidence, formal models can help us by formalizing the assumptions behind different narratives.

Is this a coincidence that in a lineage where human language evolved music evolved as well? I do not think so. Unfortunately, the groups offer very little discussion on the relation of language and music. Although the second group observes similarities in the proximate level they seem to argue that the two have different functions: "However, the fact that music and language are both found universally in all known societies […] suggests that both music and language independently fulfill more fundamental adaptive functions […]."

I think that the observation made by the second group is open to a different interpretation. Namely, I would like to suggest that language and music share the same roots, they evolved from a shared ancestral communication system, which was neither language nor music in its modern form, just as the shared ancestor of humans and chimps were neither human nor chimp. I propose, following Számadó (2010), that this shared ancestral system was a pre-hunt communication (charade) that served to organize and coordinate hunting effort in early hominins.

Note that this proposal is different from the "what language/music is good for" type of proposals because in our current societies charade is not really good for anything. We play it for fun, it is not a death or life issue. We do not play charade to coordinate anything or to signal coalition strength or to create everlasting bonds. Although the bonding function could be debated, as charade is a favourite party game, but realistically any other collective game could fill the same function. Therefore, it is difficult to think that there would be anything special in charade and that there were times when we played it for real. However, I would like to suggest that not just that we played it for real, but at that time it was everything we had (see Számadó, 2010 for details).

All in all, I would like to urge scientists working on the problem of language and music evolution to join their efforts to identify the ancestral communication system to language and music. It may not be the charade as I described it; however, it is very unlikely that these communication systems evolved independently from each other. Accordingly, the isolated study of both systems may not be as fruitful as the joint study of the two systems.

Financial support. S.S. gratefully acknowledge support from the Hungarian National Research, Development and Innovation Office NKFIH (grant number K 132250). S.S. was supported by the European Structural Investment Funds (ESIF) in Hungary (GINOP) 2.3.2-15-2016-00057.

Conflict of interest. None.

References

Bickerton, D. (2009). *Adam's tongue: How humans made language, how language made humans*. Macmillan.

Deacon, T. W. (1993). Confounded correlations, again. *Behavioral and Brain Sciences*, *16*(4), 698–699.

Dunbar, R. I. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*(4), 681–694.

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese?. *Behavioral and Brain Sciences*, *27*(4), 491.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve?. *Science*, *298*(5598), 1569–1579.

Hurford, J. R. (2011). *The origins of grammar: Language in the light of evolution II*. OUP Oxford.

Knight, C. (1998). Ritual/speech coevolution: A solution to the problem of deception. *Approaches to the Evolution of Language*, 68–91.

Miller, G. (2011). *The mating mind: How sexual choice shaped the evolution of human nature*. Anchor.

Számadó, S. (2010). Pre-hunt communication provides context for the evolution of early human language. *Biological Theory*, *5*(4), 366–382.

Számadó, S., & Szathmáry, E. (2006). Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution*, *21*(10), 555–561.

Zachar, I., Szilágyi, A., Számadó, S., & Szathmáry, E. (2018). Farming the mitochondrial ancestor as a model of endosymbiotic establishment by natural selection. *Proceedings of the National Academy of Sciences*, *115*(7), E1504–E1510.

Beyond “consistent with” adaptation: Is there a robust test for music adaptation?

Parker Tichko[a], Kevin A. Bird[b,c] and Gregory Kohn[d]

[a]Department of Music, Northeastern University, Boston, MA 02115, USA; [b]Department of Horticulture, Michigan State University, East Lansing, MI 48823, USA; [c]Ecology, Evolutionary Biology and Behavior Program, Michigan State University, East Lansing, MI 48823, USA and [d]Department of Psychology, University of North Florida, Jacksonville, FL 32224, USA.
p.tichko@northeastern.edu; birdkevi@msu.edu; gregory.kohn@unf.edu

doi:10.1017/S0140525X20001132, e115

Abstract

In their article, Mehr et al. conclude that the design features of music are consistent with adaptations for credible signaling. Although appealing to design may seem like a plausible basis for identifying adaptations, probing adaptive theories of music must be done at the genomic level and will require a functional understanding of the genomic, phenotypic, and fitness properties of music.

> “… functional information does not ‘validate’ claims of selection. It mostly serves to provide a more interesting and entertaining story. In humans, where controlled experiments and measurements of fitness are difficult or impossible to obtain, the evidence for selection must come directly from the genetic data.”
>
> — Nielsen (2009)

In their target article, Mehr et al. propose a theory of musical adaption as credible signaling, citing a breadth of evidence that spans the disciplines of anthropology, developmental psychology, and comparative psychology. Underlying their claim is the belief that establishing *adaptations* for music requires unequivocal evidence for the functional *design* of music. Writing on “what constitutes evidence for adaption” (sect. 2), the authors opine that “a successful account of music must provide evidence for the design of its principal features” (sect. 3, para. 1) and that “supporting a claim of adaptation therefore requires evidence for design.” (sect. 2, para. 8). Music, in their view, fulfills these criteria and “… exhibits design features consistent with adaptations for credible signaling, which give rise to a universal human psychology of music” (sect. 5.3, para. 2).

Although the authors put forth a well-reasoned perspective on the adaptive origins of music in credible signaling, we believe the authors err in presuming that evidence for design is a tenable basis upon which to infer adaptations for music. Criticisms regarding the inadequacy of appealing to design for adaptive theorizing are certainly not new and have historically been aimed at the adaptationist program, more generally (Gould & Lewontin, 1979). These criticisms have highlighted that evolutionary theorists have a tendency to conflate design and adaptation (Nielsen, 2009), while ignoring or underestimating the role of non-adaptive, evolutionary processes (Jensen et al., 2019) that can produce organismal complexity (Lynch, 2007).

Alongside these criticisms against the adaptationist program and the growing appreciation of non-adaptive processes in evolution, methodological improvements in the evolutionary and biological sciences have helped move the scientific paradigm for identifying adaptations beyond simple appeals to design. Innovations in evolutionary biology, population genetics, and comparative genomics, for instance, have produced a systematic framework for testing for selection, and, by extension, for testing for adaptations by natural selection. Over several decades, quantitative approaches to evolutionary biology have yielded a suite of statistical tests that detect signatures of selection at the level of the genome (Nei, Suzuki, & Nozawa, 2010; Nielsen, 2005). Using these techniques, population-genetic and comparative-genomic analyses have successfully identified selection at the local and global levels. For example, local selection can be detected based on patterns of excessive population differentiation (e.g., Fst) or by patterns of haplotype structure and linkage disequilibrium (Nei et al., 2010). Similarly, global selection is indicated by elevated rates of non-synonymous (Ka) substitutions compared to synonymous substitutions (Ks) (Clark et al., 2003) or by a heighted ratio of non-synonymous and synonymous substitutions between species (Ka/Ks), relative to the ratio of non-synonymous and synonymous polymorphisms within a species (Pn/Ps) (McDonald & Kreitman, 1991). Moreover, lineage-specific gene duplications and expansions of gene families (Demuth & Hahn, 2009) can be investigated in comparative frameworks to test for positive selection at the global level (Hahn, De Bie, Stajich, Nguyen, & Cristianini, 2005). These techniques can even be expanded to study the evolution of higher order features of the genome, such as changes in gene-coexpression networks (Oldham, Horvath, & Geschwind, 2006).

Once identified, genomic signatures of selection can, then, be related to traits proposed to be functional, providing an initial foundation for identifying adaptations. However, simply associating molecular signals of selection with a trait posited to have adaptive origins can still produce erroneous conclusions about adaptation. Associations between genetic signals of selection, although possibly reflecting an adaptive effect, could also reflect pleiotropic effects (Nielsen, 2009). To rule out pleiotropic effects, among other confounds, evolutionary biologists have developed an integrated framework that combines observational, field, and experimental methodologies to study adaptations in a number of organisms. It is only after arriving at a robust understanding of the underlying pathways between the genome, the phenotype, and fitness of the trait under question, using this integrated framework, can conclusions about adaptations confidently be drawn (Barrett & Hoekstra, 2011).

Thus, even with the remarkable methodological progress in the evolutionary and biological sciences, producing unequivocal evidence of selection, and further, adaptation, is extremely difficult, as it requires a mechanistic understanding of the functional links between the genome, the phenotype, and the fitness of the trait under study (Barrett & Hoekstra, 2011). To this end, unraveling the evolution of music may seem like an insurmountable undertaking. Already, however, the methods described above have helped illuminate the evolution of language in humans (Fisher, 2017; Fitch, Huber, & Bugnyar, 2010), echolocation in bats (Zhang et al., 2013), and the evolution of song and vocal learning in birds (Zhang et al., 2014). A similar framework could be adopted to probe adaptive hypotheses of music, particularly at the global level, as music is considered a human universal (Mehr et al., 2019). For example, comparative studies could investigate selection related to music by studying the evolution of genes associated with music and musical ability across the mammalian phylogeny. Candidates for this analysis could include genes associated with music-related disorders (e.g., Angelman and Prader–Willi syndromes as discussed in the target article, congenital amusia; Peretz, Cummings, & Dubé, 2007) and human genes orthologous to those involved in auditory traits in songbirds, which have convergent patterns to humans (Pfenning et al., 2014). Moreover, candidate genes for musical aptitude (Järvelä, 2019) could be investigated for molecular signals of selection in a comparative framework or could be checked for their overlap with gene families experiencing expansion. Findings from such genomic studies would lend more credibility to adaptive theories of music, complement the existing anthropological–psychological data on music, and move the evolutionary science of music toward a more rigorous framework, similar to the one that is currently employed in other evolutionary and biological sciences to identify adaptations. Finally, directly testing for selection using genomic data could motivate future anthropological and psychological research by identifying possible biological pathways and traits of musicality, which could be used to generate novel, testable hypotheses about the evolution of music.

Financial support. This study was supported by the National Science Foundation (KAB, NSF-GRFP DGE-142487.1).

Conflict of interest. None.

References

Barrett, R. D. H., & Hoekstra, H. E. (2011). Molecular spandrels: Tests of adaptation at the genetic level. *Nature Reviews. Genetics*, *12*(11), 767–780.

Clark, A. G., Glanowski, S., Nielsen, R., Thomas, P. D., Kejariwal, A., Todd, M. A.…, (2003). Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science (New York, N.Y.)*, *302*(5652), 1960–1963.

Demuth, J. P., & Hahn, M. W. (2009). The life and death of gene families. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *31*(1), 29–39.

Fisher, S. E. (2017). Evolution of language: Lessons from the genome. *Psychonomic Bulletin & Review*, *24*(1), 34–40.

Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social cognition and the evolution of language: Constructing cognitive phylogenies. *Neuron*, *65*(6), 795–814.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society*, *205* (1161), 581–598.

Hahn, M. W., De Bie, T., Stajich, J. E., Nguyen, C., & Cristianini, N. (2005). Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Research*, *15*(8), 1153–1160.

Järvelä, I. (2019). Genomics approaches for studying musical aptitude and related traits. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 439–458). Oxford University Press.

Jensen, J. D., Payseur, B. A., Stephan, W., Aquadro, C. F., Lynch, M., Charlesworth, D., & Charlesworth, B. (2019). The importance of the neutral theory in 1968 and 50 years on: A response to Kern and Hahn 2018. *Evolution; International Journal of Organic Evolution*.

Lynch, M. (2007). The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(Suppl. 1), 8597–8604.

McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, *351*(6328), 652–654.

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*, *366*(6468), 957–970. https://doi.org/10.1126/science.aax0868.

Nei, M., Suzuki, Y., & Nozawa, M. (2010). The neutral theory of molecular evolution in the genomic era. *Annual Review of Genomics and Human Genetics*, *11*, 265–289.

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, *39*, 197–218.

Nielsen, R. (2009). Adaptionism-30 years after Gould and Lewontin. *Evolution; International Journal of Organic Evolution*, *63*(10), 2487–2490.

Oldham, M. C., Horvath, S., & Geschwind, D. H. (2006). Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(47), 17973–17978.

Peretz, I., Cummings, S., & Dubé, M.-P. (2007). The genetics of congenital amusia (tone deafness): A family-aggregation study. *American Journal of Human Genetics*, *81*(3), 582–588.

Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., … & Jarvis, E. D. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science (New York, N.Y.)*, *346*(6215), 1256846.

Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K. A., Fang, X., … & Wang, J. (2013). Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science (New York, N.Y.)*, *339*(6118), 456–460.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., … & Wang, J. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science (New York, N.Y.)*, *346*(6215), 1311–1320.

Understanding the origins of musicality requires reconstructing the interactive dance between music-specific adaptations, exaptations, and cultural creations

Laurel J. Trainor

Department of Psychology, Neuroscience and Behaviour; McMaster University,
West Hamilton, ON L8S4B2, Canada.
LJT@mcmaster.ca; https://trainorlab.mcmaster.ca/

doi:10.1017/S0140525X20001363, e116

Abstract

The evolutionary origins of complex capacities such as musicality are not simple, and likely involved many interacting steps of musicality-specific adaptations, exaptations, and cultural creation. A full account of the origins of musicality needs to consider the role of ancient adaptations such as credible singing, auditory scene analysis, and prediction-reward circuits in constraining the emergence of musicality.

"Origins of Music in Credible Singing" by Mehr et al. and "Music as a Co-evolved System for Social Bonding" by Savage et al. attempt to explain the evolutionary origins of musicality. Although both papers mention gene–culture interaction, both fall short of an evolutionary account, which would need to consider the complex interactions between adaptations specifically for musicality (i.e., musical capacity itself increased inclusive fitness, leading to adaptive changes), exaptation (evolutionary pressures for adaptation of non-musical traits or capacities that, once evolved, helped enable musicality), and cultural creation (which could be useful for fitness and, although not an evolutionary adaptation, could drive subsequent evolutionary adaptations). The evolution of complex cognitive capacities such as musicality and language are almost certain to involve a complex interplay between these three factors (see Trainor, 2015, 2018 for a detailed discussion).

Both the Mehr et al.'s and Savage et al.'s papers agree in focusing on social interactions as the adaptive functions that drove the evolution of musicality. However, they differ markedly in how they conceptualize the role of musicality in promoting social interaction. For Mehr et al., music is a form of credible signaling that conveys information from sender to receiver about their social cohesion intentions – music can signal to a competitor the coalition strength of one's group, or signal one's attention to an infant. Savage et al. postulate that, independent from musicality, social cohesion within groups has adaptive benefits – and that musicality originated as a human cultural creation, but because music caused increased social cohesion, through adaptation it became more effective in this regard.

One elephant in the room of these papers is a definition of musicality. For Savage et al., musicality arose recently, initially through human culture, and involves a set of design features including predictive hierarchical beat structures and "synchronized, harmonized singing and dancing in groups." For Mehr et al., music is ancient and found in many nonhuman species in the form of vocalizations to signal territory and contact calls. Thus, directly comparing the papers is difficult.

What both papers are missing is an evolutionary approach that not only addresses possible adaptive functions of musicality, but why and how musical capacities evolved in the particular way they did. Similar to other complex capacities, the evolution of musical capacities will likely not consist of one adaptation, but rather a long sequence of adaptive, exapted, and cultural influences that interact in complex ways. Human biology, including brain architecture, has changed from that of our closest genetic relatives (see Savage et al.), so some adaptations (whether music-specific or exaptations) enhancing musicality have evolved fairly recently. At the same time, a full understanding requires considerations of precursor abilities that evolved over longer timeframes.

Both papers describe multiple adaptive effects of musicality, from intimate infant caretaking to social signaling or bonding between large groups of adults. Why would these disparate functions be served by the same musical faculty? Why did we not evolve three systems, for example, for information communication (language), intimate emotional interaction, and group bonding? Because music evolved in the context of a brain and body already adapted for other functions, there were likely severe constraints on possible forms communication systems might take. In other words, music is, in part, an exaptation or perhaps in some aspects a "byproduct" of these prior adaptations. Here, I discuss three important constraints.

First, why is music based in the auditory-motor system? As Mehr et al. elaborated, signaling through sound production and perception involves evolutionarily ancient adaptations found in animals as disparate as insects, birds, and mammals. Many would not call such signaling music *per se*, but it was available to be exapted in the creation of human musicality, thereby greatly constraining musical forms. Thus, musicality is, in part, an exaptation of adaptations for auditory production–perception signaling systems.

Second, why does music have the pitch structure it does? As elaborated by Trainor (2015, 2018; see also Huron, 2001), the perception of pitch itself is most likely a byproduct of auditory scene analysis (Bregman, 1990). The goal of auditory scene analysis is to identify and locate multiple sound sources in a typical environment. An obvious evolutionary advantage is for identifying predators, prey, and offspring. Most animal vocalizations have harmonic structure consisting of a fundamental frequency and harmonics at integer multiples of the fundamental. Harmonics of simultaneous sound sources overlap in frequency range, and the signal reaching the ear consists of their sum in one complex waveform. The neural solution, preserved across many species, is to first separate the complex waveform into its constituent frequencies, and then groups frequencies together that stand in harmonic relations to a fundamental, as these likely originated from the same sound source. At the level of consciousness, individual frequencies are not perceived, only integrated percepts with particular pitches, timbres, and spatial locations. Thus, pitch perception is a byproduct of evolutionary pressures for auditory scene analysis, and was likely exapted for the creation of musicality. There may also have been more recent musicality-specific adaptations to enhance perception and production of pitch and tonal structures especially in group contexts.

Third, why is music rewarding? Accurately predicting the future is crucial for fitness (Huron, 2006) – prediction errors can result in being eaten or a missed mating opportunity. Indeed, the brain is continually predicting the future and adjusting its internal models when incorrect (*predictive coding*) (e.g., Heilbron & Chait, 2018; Trainor & Zatorre, 2015). Across many species, prediction is intimately connected to brain reward centers (Schultz, 2013). The often-noted tonal and rhythmic regularities in music enable prediction of upcoming sounds. Indeed, one argument for musicality being a cultural creation is that it well designed to activate preexisting reward centers (Salimpoor, Zald, Zatorre, Dagher, & McIntosh, 2015). Although subsequent adaptations may have enhanced these effects, musicality is, in part, an exaptation of ancient adaptations for rewarding correct predictions.

In sum, although it is not easy to reconstruct the evolutionary history of complex capacities, I propose that to understand the evolution of musicality we need to seriously consider complex interactions between music-specific adaptations, exaptations, and cultural creation over an extended evolutionary timeframe.

Financial support. This commentary was supported by grants from the Canadian Institutes of Health Research (MOP 153130), the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-05416), the Social Science and Humanities Research Council (435-2020-0442), and the Canadian Institute for Advanced Research.

Conflict of interest. None.

References

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.

Heilbron, M., & Chait, M. (2018). Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience*, *389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception 19*, 1–64. doi: 10.1525/mp.2001.19.1.1.

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

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*, 86–91.

Schultz, W. (2013). Updating dopamine reward signal. *Current Opinion in Neurobiology*, *23*, 229–238. https://doi.org/10.1016/j.conb.2012.11.012.

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal British Society B: Biology*, *370*, 20140089.

Trainor, L. J. (2018). The origins of music: Auditory scene analysis, evolution and culture in musical creation. In H. Honing (ed.), *The origins of musicality* (pp. 81–112). MIT Press.

Trainor, L. J., & Zatorre, R. J. (2015). The neurobiology of musical expectations from perception to emotion. In S. Hallam, I. Cross & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 285–305). Oxford University Press.

Challenging infant-directed singing as a credible signal of maternal attention

Sandra E. Trehub

Department of Psychology, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada.
sandra.trehub@utoronto.ca

doi:10.1017/S0140525X20001442, e117

Abstract

I challenge Mehr et al.'s contention that ancestral mothers were reluctant to provide all the attention demanded by their infants. The societies in which music emerged likely involved foraging mothers who engaged in extensive infant carrying, feeding, and soothing. Accordingly, their singing was multimodal, its rhythms aligned with maternal movements, with arousal regulatory consequences for singers and listeners.

Mehr et al. contend that infant-directed (ID) singing arose as a credible *acoustic* signal of maternal attention and provided an impetus for the emergence of melody. They emphasize the divergent interests of infant and mother – infants wanting more attention than mothers care to provide. The historical record, however, is at odds with this contention. Infant carrying has been central to infant well-being and survival for most of human history, promoting safety, thermoregulation, and arousal regulation. Ubiquitous infant carrying, both in the past and in contemporary foraging cultures (e.g., Aka of Central Africa) co-occurs with frequent feeding, prolonged breastfeeding (yielding greater interbirth intervals), and prompt responsiveness to infant distress, resulting in less infant crying and more sleeping than infants in nearby (Nganda) farming communities whose mothers exhibit less proximal interaction but somewhat more distal interaction such as vocalizing (Hewlett, Lamb, Shannon, Leyendecker, & Schölmerich, 1998). These caregiving differences are not attributable to differences in environmental hazards or infant mortality, which are comparable. Incidentally, the carrying and feeding frequency of Nganda farming mothers greatly exceeds that of Western mothers. In any case, foraging mothers provide considerably more attention than demanded by infants. Moreover, complaints about the burdens of infant care are relatively rare in such societies and more common in affluent Western societies. Also notable is the more egalitarian social structure of foraging societies than farming or industrial societies, with foragers sharing caregiving and provision-seeking in their tightly knit communities.

ID singing likely emerged in conjunction with carrying in the foraging communities of the distant past. Accordingly, such singing would be multimodal rather than the acoustic signal envisioned by Mehr et al. Moreover, one would expect the rhythm of movement to influence the rhythm of singing (Ayres, 1973). Interestingly, rhythmic movement such as brisk walking calms human and other mammalian infants more effectively than stationary holding (Esposito et al., 2014). In a recent cross-cultural study of mother–infant dyads from several developed countries, mothers systematically responded to distressed infants by picking them up and talking (Bornstein et al., 2017). Although these mothers sang to their infants, they did so in playful rather than soothing contexts (Bornstein, personal communication, November 13, 2017). In a laboratory study in which infant distress was induced, and mothers were restricted from picking up infants but allowed to touch them, maternal singing of familiar songs reduced infant distress more effectively than unfamiliar songs, which were more effective than ID speech accompanied by affectionate touch (Cirelli & Trehub, 2020). Repetition and predictability, which are crucial in parent–infant interactions, are also central to music and its appeal (Margulis, 2014).

Western mothers' interactions with infants are often distal, and their singing commonly occurs in face-to-face contexts, with mutual gaze potentially compensating for the absence of physical contact (Konner, 2010). Falk (2004) has proposed that ID vocalizations arose to enable hominin mothers to put their baby down in service of efficient foraging – keeping in touch without tactile or eye contact. Indeed, ID singing is notably expressive and recognizable, with its warm vocal timbre, temporal regularity, and highly stereotyped performances (Bergeson & Trehub, 2002; Trainor, Clark, Huntley, & Adams, 1997; Trehub, Plantinga, & Russo, 2016). Such singing commonly occurs in conjunction with coordinated rhythmic movement (Longhi, 2009) and smiling (Trehub et al., 2016).

In the absence of carrying, mutual gaze, or rhythmic movement, ID singing is not energetically costly. Songs for infants are melodically, rhythmically, and textually simple and repetitive (Sakata, 1987; Unyk, Trehub, Trainor, & Schellenberg, 1992), and they are often sung with meaningless syllables (Brown, 1980; Finnegan, 1977). In fact, it is not unusual for mothers to engage in other demanding activities (e.g., driving a car) while singing to infants (e.g., in out-of-view car seats), so the notion of ID singing as a costly and credible signal of maternal investment or attention remains unconvincing.

ID singing may be as important for mothers as for infant listeners. For example, singing in a soothing manner (lullaby style) reduces arousal levels in mothers as well as infants (Cirelli, Jurewicz, & Trehub, 2020), and maternal calming may accelerate infant calming. Similarly, lively play songs may have energizing or euphoric consequences for the dyad. In situations when other emotional outlets are unavailable, the one-on-one context enables singers to vent their grievances by improvising the lyrics of songs to prelinguistic listeners (e.g., Masuyama, 1989).

The rewarding nature of musical engagement was notably absent from the maternal-attention signaling and coalition-signaling aspects of the Mehr et al. account. According to Merker, Morley, and Zuidema (2015), any theory of the origins of music must account for its motivational core, as reflected in the universal tendency to gather for group singing and dancing unless there are societal prohibitions for doing so. These group musical contexts, including dyadic musical interactions, provide opportunities for coordinating arousal levels and emotions and reinforcing the solidarity of participants.

If ID singing evolved, at least in part, as an acoustic signal of maternal attention, as claimed by Mehr et al., one would expect it to show greater efficacy in attenuating infant distress than rhythmic but non-melodic vocalizations (e.g., rhythmic shushing), rhythmic movement (e.g., rocking and jiggling), holding with or without affective touch (e.g., stroking), and transport (e.g., carrying while walking). These comparisons, which are readily testable, should reveal whether ID singing is a primary component of infant care and therefore a possible signal of parental investment or whether it is an embellishment superimposed on other, more basic forms of infant care.

In contrast to Mehr et al., the Savage et al. companion paper is broadly applicable to musical functions, past and present, which typically involve affiliative or social bonding activities. Savage et al. are also flexible about the timeline, precedence, and relative contributions of cultural and biological evolution, according primacy to the former. Moreover, in characterizing music as multimodal, incorporating sound and movement rather than sound alone, and according a central role to music-making, sociality, and their rewarding consequences, their account is more aligned with the perspectives outlined in this commentary.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Ayres, B. (1973). Effects of infant carrying practices on rhythm in music. *Ethos 1*, 381–404.

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, *13*, 72–75.

Bornstein, M. H., Putnick, D. L., Rigo, P., Esposito, G., Swain, J. E., & … Venuti, P. (2017). Neurobiology of culturally common maternal responses to infant cry. *Proceedings of the National Academy of Sciences*, *114*, E9465–E9473.

Brown, M. J. E. (1980). Lullaby. In S. Sadie (Ed.), *The new grove dictionary of music and musicians* (pp. 313–314). Macmillan.

Cirelli, L.K., Jurewicz, Z.B., & Trehub, S. E. (2020). Effects of maternal singing style on mother-infant arousal and behavior. *Journal of Cognitive Neuroscience*, *32*, 1213–1220.

Cirelli, L. K., & Trehub, S. E. (2020), Familiar songs reduce infant distress. *Developmental Psychology*, *56*, 861–868.

Esposito, G., Yoshida, S., Ohnishi, R., Tsuneoka, Y., Rostagno, M., & … Kuroda, K. O. (2014). Infant calming responses during maternal carrying in humans and mice. *Current Biology*, *23*, 739–745.

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, *27*, 491–541.

Finnegan, R. H. (1977). *Oral poetry: Its nature, significance, and social context.* Cambridge University Press.

Hewlett, B. S., Lamb, M. E., Shannon, D., Leyendecker, B., & Schölmerich, A. (1998). Culture and early infancy among Central African foragers and farmers. *Developmental Psychology*, *34*, 653–661.

Konner, M. J. (2010). *The evolution of childhood: Relationships, emotion, mind.* Harvard University Press.

Longhi, E. (2009). "Songese": Maternal structuring of musical interaction with infants. *Psychology of Music*, *37*, 195–213.

Margulis, E. H. (2014). *On repeat: How music plays the mind.* Oxford University Press.

Masuyama, E. E. (1989). Desire and discontent in Japanese lullabies. *Western Folklore*, *48*, 169–177.

Merker, B., Morley, B., & Zuidema (2015). Five fundamental constraints on theories of the origins of music. *Philosophical Transactions of the Royal Society B*, *370*, 20140095.

Sakata, H. L. (1987). Hazara women in Afghanistan: Innovators and preservers of a musical tradition. In E. Koskoff (Ed.), *Women and music in cross-cultural perspective* (pp. 85–95). Greenwood.

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*, 383–396.

Trehub, S. E., Plantinga, J., & Russo, F. A. (2016). Maternal vocal interactions with infants: Reciprocal visual influences. *Social Development*, *25*, 665–683.

Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music*, *20*, 15–28.

The evolutionary benefit of less-credible affective musical signals for emotion induction during storytelling

Caitlyn Trevor[a] and Sascha Frühholz[a,b]

[a]Department of Psychology, University of Zurich, Cognitive and Affective Neuroscience Unit, Binzmuehlestrasse 14, 8050 Zurich, Switzerland and
[b]Department of Psychology, University of Oslo, Blindern, 0317 Oslo, Norway
caitlyn.trevor@psychologie.uzh.ch
sascha.fruehholz@uzh.ch
https://www.psychology.uzh.ch/en/areas/nec/kaneuro.html

doi:10.1017/S0140525X20001004, e118

Abstract

The credible signaling theory underexplains the evolutionary added value of less-credible affective musical signals compared to vocal signals. The theory might be extended to account for the motivation for, and consequences of, culturally decontextualizing a biologically contextualized signal. Musical signals are twofold, communicating "emotional fiction" alongside biological meaning, and could have filled an adaptive need for affect induction during storytelling.

Although we generally agree with the credible signaling hypothesis and provide evidence for credible signaling in contemporary music, an important issue remains unaddressed by the theory. Recent research suggests that contemporary credible musical signals are less emotionally impactful than their vocal counterparts. Musical signals take far more time and energy to manufacture compared to vocal ones. The theory falls short of explaining the evolutionary added value of these more taxing and less affective musical signals. The credibility hypothesis should be extended to account for this counterintuitive observation by including a component regarding the motivation for, and consequences of, culturally decontextualizing a biologically contextualized signal (Frühholz, Trost, & Kotz, 2016). Specifically, we hypothesize that these affectively weaker musical signals communicate "emotional fiction" alongside their biological meanings and may have been motivated by the adaptive need for emotionally impactful storytelling.

Although we agree with the authors' claim that today's actual domain of music is far removed from its proper domain, recent findings of credible signals in contemporary music show that some ancient, vocal-inspired signals have resiliently persisted throughout the diverse cultural metamorphoses that music has undergone over centuries across the world. For example, one contemporary credible signal feature in music to convey affective meaning is *roughness*, a harsh, buzzing, raspy sound quality (Vassilakis & Kendall, 2010). Roughness has a long evolutionary trajectory in human and animal alarm calls (Arnal, Flinker, Kleinschmidt, Giraud, & Poeppel, 2015; Engelberg & Gouzoules, 2019; Schwartz, Engelberg, & Gouzoules, 2019) and has been found to be present in terrifying excerpts from horror film music (Trevor, Arnal, & Frühholz, 2020). Another contemporary credible signal in music is the *sigh*, a vocal signal generated by 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.

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.

Smith, D., Schlaepfer, P., Major, K., Dyble, M., Page, A. E., Thompson, J., … Astete, L. (2017). Cooperation and the evolution of hunter-gatherer storytelling. *Nature Communications*, *8*(1), 1–9.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645–665.

Teigen, K. H. (2008). Is a sigh "just a sigh"? Sighs as emotional signals and responses to a difficult task. *Scandinavian Journal of Psychology*, *49*(1), 49–57.

Trevor, C., Arnal, L. H., & Frühholz, S. (2020). Terrifying film music mimics alarming acoustic feature of human screams. *The Journal of the Acoustical Society of America*, *147*(6), EL540–EL545. https://doi.org/10.1121/10.0001459.

Trevor, C., & Huron, D. (2018). Are humoresques humorous? On the similarity between laughter and staccato. *Empirical Musicology Review*, *13*(1–2), 66. https://doi.org/10.18061/emr.v13i1-2.5608.

Vassilakis, P. N., & Kendall, R. A. (2010). Psychoacoustic and cognitive aspects of auditory roughness: Definitions, models, and applications. *Proceedings of Human Vision and Electronic Imaging XV*, *7527*, 1–7. https://doi.org/10.1117/12.845457.

Walton, K. L. (1990). *Mimesis as make-believe: On the foundations of the representational arts*. Harvard University Press.

Musical bonds are orthogonal to symbolic language and norms

Connor Wood

Center for Mind and Culture, Boston, MA, 02215, USA.
connorpw@bu.edu, https://www.researchgate.net/profile/Connor_Wood5

doi:10.1017/S0140525X20001272, e119

Abstract

Both Mehr et al.'s credible signaling hypothesis and Savage et al.'s music and social bonding hypothesis emphasize the role of multilevel social structures in the evolution of music. Although empirical evidence preferentially supports the social bonding hypothesis, rhythmic music may enable bonding in a way uniquely fitted to the normative and language-based character of multilevel human societies.

Mehr et al. hypothesize that the ultimate evolutionary function of musical rhythm in its proper domain is credible social signaling of coalition quality, whereas Savage et al. propose instead that musicality developed via iterated niche construction into a gene-culture coadaptation broadly for social bonding. Interestingly, both teams of authors locate important functions and selection pressures for musicality in the multilevel social structures of behaviorally modern humanity. After highlighting several weaknesses in Mehr et al.'s argument, I will conclude by focusing on this issue.

Although Mehr et al. provide suggestive evidence of a phylogenetic association between social complexity and vocal flexibility among primates, their scenario for competitive signaling using music and dance in multilevel societies is unrealistic for several reasons. First, within such societies, intergroup allegiances and cooperative pacts are only rarely freely chosen in the sort of real-time biological marketplace that Mehr et al. envision; instead, reciprocal cooperation typically reflects some prestanding, socially normative, or obligatory relationship, such as between nested clans or incumbents of ritualized aid roles (Hill, Wood, Baggio, Hurtado, & Boyd, 2014; cf. Tomasello, 2020). Dances typically reinvigorate or trigger these commitments rather than enabling participants to choose them; when partner choice does occur, it is likely to be between co-participants in the same dance (e.g., Rappaport, 1968). Although sexual attraction between observers and dance participants can lead to new coalition memberships through marriage (e.g., Bovin, 2001), this sexual display function does not match what the authors have in mind, given their emphasis on the lack of sexual differentiation in music.

Second, Mehr et al.'s argument for a basic coalition-signaling function of rhythmic music partly hinges on the amount of time and shared commitment successful group dances putatively require (sect. 4.2.1). Although rehearsal is indeed crucial for many musical performances, the authors' argument on this point sits uncomfortably with the fact that many mechanistic features of rhythm perception and entrainment serve to enhance temporal predictability and perception-action coupling, and so minimize the effort required to achieve mutual synchronization (cf. Savage et al., sect. 4). For example, motor entrainment facilitates endogenous beat prediction in audiomotor brain networks, which in turn facilitates accurate motor timing – a self-reinforcing cycle (Morillon & Baillet, 2017; Su & Pöppel, 2012). Simple forms of synchrony (e.g., clapping to a shared rhythm) can thus enable dynamic physical coordination to an isochronous or metrical beat with minimal rehearsal. It is unlikely that complex, choreographed dance performances – which require rehearsal and thus index group commitment, as the authors propose – phylogenetically preceded these simpler forms of music and dance, suggesting that choreographed group displays are a secondary function.

Finally, Mehr et al.'s claim that rhythmic music's proper domain is overtly signaling covert, prior cooperative intentions rather than catalyzing such intentions requires the rejection of wide-ranging empirical and experimental evidence that musical and rhythmic performances do, in fact, causally effect cooperation and prosociality (Mogan, Fischer, & Bulbulia, 2017). Indeed, research subjects may put particular effort into synchrony with outgroup members in an attempt to establish relationships – again, as co-participants in the same rhythm, not solicitations to observers (Fujiwara, Kimura, & Daibo, 2020). The authors offer theoretical arguments against music's social bonding role in section 3.2, but additional research poses problems for at least two of these arguments. Deep functional interconnections between audio, motor, and reward systems during rhythm perception and entrainment indicate fundamental design constraints on isolating motor entrainment from reward processing (Todd & Lee, 2015), thus limiting the potential for free rider mutations. Meanwhile, language's ability to convey displaced, propositional information enhances the possible scope of shared action while simultaneously increasing the potential for disagreement and self-other differentiation (Fitch, 2006; Knoblich & Sebanz, 2008). Music thus has a distinct advantage over language for specific kinds of social bonding (Cross, 2009), contrary to the authors' claims.

What kinds of bonding can music better accomplish than language? Savage et al. point out that larger-scale, complex societies are more likely to feature audience–performer divides, whereas smaller-scale, politically egalitarian societies tend to exhibit participatory musical modes (often heavily featuring rhythmic dance). A parsimonious explanation both for rhythmic music's role in forging or certifying bonds between distinct groups or individuals and for its association with egalitarian political formations is that synchrony tends to reduce the salience of categorical role or group boundaries, allowing participants to bond with one another without significant regard to rank or subgroup affiliation. Synchrony thus facilitates new or temporary – often phenomenologically egalitarian – subjective in-group identities (Cross, Turgeon, & Atherton, 2019).

This suggestion points to an important question implicit but not addressed in Savage et al. (e.g., sect. 2.5) and, to a lesser extent, in Mehr et al.: What is the relationship between music and *normative* and *symbolic* social categories and structure? Multilevel human societies – and therefore the social-bonding and signaling problems humans must solve – largely rest on normative, nested categories (e.g., clans and phratries) rather than simply biological kinship. These normative roles and structural divisions are typically complementary, often hierarchical, and invariably defined by differentiated obligations. They depend on language, because language uniquely enables the creation and transmission of abstract, prescriptive concepts such as "mother-in-law" or "fellow moiety member."

Importantly, conversational language necessarily takes a complementary, turn-taking temporal structure (Pika, Wilkinson, Kendrick, & Vernes, 2018), in direct contrast to the convergence and simultaneity of mutual rhythmic entrainment. An important implication is, therefore, that language underlies the establishment of socially patterned behavior in the form of complementary, enduring roles and their (often hierarchically) differentiated obligations, whereas music and dance appear to facilitate social bonding in a different, orthogonal mode characterized by precise motor coordination with immediate, patently available physical stimuli. Rhythmic musical bonding may thus be ideally suited for addressing a need specific to multilevel human societies: Re-establishing bonds at intervals between individuals and subgroups who otherwise are at a physical but also symbolic, normative, and hierarchical remove from one another. Normatively differentiated roles, then, were likely a crucial feature of the cultural niche that drove iterated selection for musicality.

Acknowledgments. I thank Paul Seabright and the Institute for Advanced Study at Toulouse for organizing the "Origins of Music in Non-State Societies" workshop at the Abbaye de Royaumont in 2017 that generated both target articles. This commentary draws on ideas from my contribution to that workshop.

Financial support. This study was enabled by the John Templeton Foundation (Grant no. 61157).

Conflict of interest. None.

References

Bovin, M. (2001). *Nomads who cultivate beauty: Woḏaaḇe dances and visual arts in Niger*. Nordic Africa Institute.

Cross, I. (2009). The evolutionary nature of musical meaning. *Musicae Scientiae*, *13* (2_Suppl.), 179–200. https://doi.org/10.1177/1029864909013002091.

Cross, L., Turgeon, M., & Atherton, G. (2019). How moving together binds us together: The social consequences of interpersonal entrainment and group processes. *Open Psychology*, *1*(1), 273–302. https://doi.org/10.1515/psych-2018-0018.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*(1), 173–215. https://doi.org/10.1016/j.cognition.2005.11.009.

Fujiwara, K., Kimura, M., & Daibo, I. (2020). Rhythmic features of movement synchrony for bonding individuals in dyadic interaction. *Journal of Nonverbal Behavior*, *44*(1), 173–193. https://doi.org/10.1007/s10919-019-00315-0.

Hill, K. R., Wood, B. M., Baggio, J., Hurtado, A. M., & Boyd, R. T. (2014). Hunter-gatherer inter-band interaction rates: Implications for cumulative culture. *PLoS ONE*, *9*(7), e102806. https://doi.org/10.1371/journal.pone.0102806.

Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: From entrainment to joint action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1499), 2021–2031. https://doi.org/10.1098/rstb.2008.0006.

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. https://doi.org/10.1016/j.jesp.2017.03.009.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(42), E8913–E8921. https://doi.org/10.1073/pnas.1705373114.

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: Bridging the gap between human and animal communication. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880), 20180598. https://doi.org/10.1098/rspb.2018.0598.

Rappaport, R. A. (1968). *Pigs for the ancestors: Ritual in the ecology of a new Guinea people*. Yale University Press.

Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, *76*(3), 373–382. https://doi.org/10.1007/s00426-011-0346-3.

Todd, N. P. M., & Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: A new synthesis and future perspectives. *Frontiers in Human Neuroscience*, *9*, 1–25. https://doi.org/10.3389/fnhum.2015.00444.

Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, *43*, e56. https://doi.org/10.1017/S0140525X19001742.

Social bonding and credible signaling hypotheses largely disregard the gap between animal vocalizations and human music

Marcel Zentner

Department of Psychology, University of Innsbruck, 6020 Innsbruck, Austria.
marcel.zentner@uibk.ac.at

doi:10.1017/S0140525X2000165X, e120

Abstract

Mehr et al. propose a theory of the evolution music that can potentially account for most animal vocalizations as precursors to human music. Therein lies its appeal but also its Achilles' heel, for the wider the range of animal vocalizations treated as premusical expressions, the wider the gap to human music. Here, I offer a few critical observations and constructive suggestions that I hope will help the authors strengthen their case.

Philosopher of science Paul Feyerabend once noted that "the more popular an idea, the less one thinks about it and the more important it becomes to examine its limitations" (Feyerabend, 1984, p. 111). This remark could be applied to idea of the evolutionary origins of music in social bonding proposed by Savage et al. in their target article. The idea has risen in popularity and acceptance over the years, partly by virtue of its intuitive appeal and partly because it ennobles music. A theory of the "evolutionary origins of music in self-isolation," even if well-supported, might have enjoyed considerably less acclaim. In this context, Mehr et al.'s contribution offers a much-needed critical examination of prominent views on the evolution of music. By exposing a number of weaknesses in the view of the evolutionary origins of music in social bonding, it invites us to revisit some long-cherished assumptions and open up to alternative hypotheses. Not an evolutionary biologist or psychologist, I will offer a few observations and comments that are inspired by my own research on music and emotion, musical ability, and on infants' music perception rather than assessing the strengths and weaknesses of the views in question from an evolutionary standpoint.

Although mainly focusing on the view of the evolutionary origins of music in social bonding, Mehr et al. also touch on the theory of the origins of music in sexual selection. They see it as poorly supported, in part because "If music evolved to signal mate quality, then adaptations for music production should be more developed in men and adaptations for music perception should be more developed in women" (Mehr et al., target article). Although this particular prediction is difficult to test because the capacities for music production and music perception are related (e.g., Dalla Bella et al., 2017), sex differentiation in musical behaviors and aptitudes between the sexes ought to be expected from a sexual selection point of view. Our laboratory has examined musical abilities and other music related dispositions in thousands of participants and we have found little evidence for sex differences so far (Law & Zentner, 2012; Zentner & Strauß, 2017). Using a scale that captures two components of what we term "music-mindedness," namely liking for music and musical competence (somewhat reminiscent of the distinction music perception and production), we similarly failed to find any sex differences (Zentner, in preparation).

Mehr et al. offer a thought-provoking alternative view to the social bonding hypothesis by proposing that music evolved because it functioned as a credible signal of coalition strength, size, and coordination ability. Much of the literature on the evolution of music has focused on select animal vocalizations, notably vocal learners. A theory of the evolution music that accounts for most animal vocalizations would be both more parsimonious and more powerful than partial, species-specific accounts. One of the strengths of Mehr's proposition is that, by including territorial calls, it can integrate a broader array of animal vocalizations than the social bonding hypothesis and other prior evolutionary accounts of music. Yet one difficulty with this idea is that it hinges on the merits of the claim that "that territorial vocalizations are an evolutionary precursor to music, especially rhythmic music." Given the importance of this claim for the authors' hypothesis on the origins of music, it receives relatively little attention. It would strengthen their position to provide a specification of the types of evidence needed to connect animal vocalizations to human music.

This is not an easy task given how little is known about the earliest forms of human music. If the way Homer's sirens sounded is anyone's guess, the picture becomes even more muddied when we go further back in the history of human music making. Somewhat ironically, both target articles provide relatively circumscribed definitions of music that may make their case more difficult to sustain than it needs to be. For instance, conventions encoded in our language, such as to say that birds sing, but not donkeys or dogs (Nettl, 2015), constitute biases in the understanding of music that may hamper the search for substantive connections between early human, evolved, contemporary forms of music and animal vocalizations.

A relevant but hardly discussed aspect of animal vocalizations is that they appear to be often emotionally charged. For example, tamarin calls can be broadly subdivided into those used in an affiliative context and those relating to fear and threat. In a fascinating effort to particularize such calls, Snowdon and Teie (2010) identified tritones, minor seconds, and noise as musico-acoustical markers for threat related calls. Positing a connection between animal territorial calls and music, Mehr et al. note that, in humans, vocal and instrumental music "appears in political and military contexts with analogues to territorial signaling (…); is generally not sexually differentiated (…); and, of course, is often loud." Although music is occasionally used in bellicose contexts, the vast majority of music is not. Equally rare in human music are the harsh and dissonant features observed in many threat-related territorial calls. Relatedly, territorially relevant emotions such as fear and anger are among the most rarely experienced emotions in response to music (Juslin & Laukka, 2004; Zentner, Grandjean, & Scherer, 2008).

These findings do not necessarily undermine the authors' hypothesis of the origins of music. Still, offering a compelling explanation for why human music is structurally and emotionally so divorced from animal territorial calls would strengthen their hypothesis. Similarly, if music had a social purpose, this purpose seems to have largely vanished. Today, music is predominantly consumed in solitary contexts, and used for self-serving emotional rather than social, societal, or communicative purposes (e.g., Randall & Rickard, 2017). As with many evolutionary theories, it will be difficult to gather unequivocal evidence to support the origins of music in credible signaling hypothesis. Nevertheless, as scholars embrace the idea that music evolved to promote social bonding and cohesion with increasing enthusiasm, a challenge to its general sway may be just what music evolutionists need to keep their inquisitive pulse up.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflict of interest. None.

References

Dalla Bella, S., Farrugia, N., Benoit, C. E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities. *Behavior Research Methods*, *49*(3), 1128–1145.

Feyerabend, P. (1984). *Wissenschaft als Kunst.* [Science as art]. Suhrkamp Verlag.

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*, 217–238.

Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *PLoS ONE*, *7*(12), e52508.

Nettl, B. (2015). *The study of ethnomusicology. Thirty-three discussions* (pp. 463–472). University of Illinois Press.

Randall, W. M., & Rickard, N. S. (2017). Reasons for personal music listening: A mobile experience sampling study of emotional outcomes. *Psychology of Music*, *45*(4), 479–495.

Snowdon, C. T., & Teie, D. (2010). Affective responses in tamarins elicited by species-specific music. *Biology Letters*, *6*(1), 30–32.

Zentner, M. (in preparation). Music-mindedness: The musical dimension of personality.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion (Washington, D.C.)*, *8*(4), 494–521.

Zentner, M., & Strauß, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*, *1400*(1), 33–45.

Authors' Responses

Toward inclusive theories of the evolution of musicality

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

[a]Faculty of Environment and Information Studies, Keio University, Fujisawa 252-0882, Japan; [b]College of Arts, Media and Design, Northeastern University, Boston, MA 02115, USA; [c]Department of Experimental Psychology, Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford OX2 6PN, UK; [d]Department of Psychology, University of California San Diego, La Jolla, CA 92093, USA; [e]Department of Anthropology, Boston University, Boston, MA 02215, USA; [f]Department of Archaeology, University of Reading, Reading RG6 6AB, UK and [g]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, bronwyntarr01@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

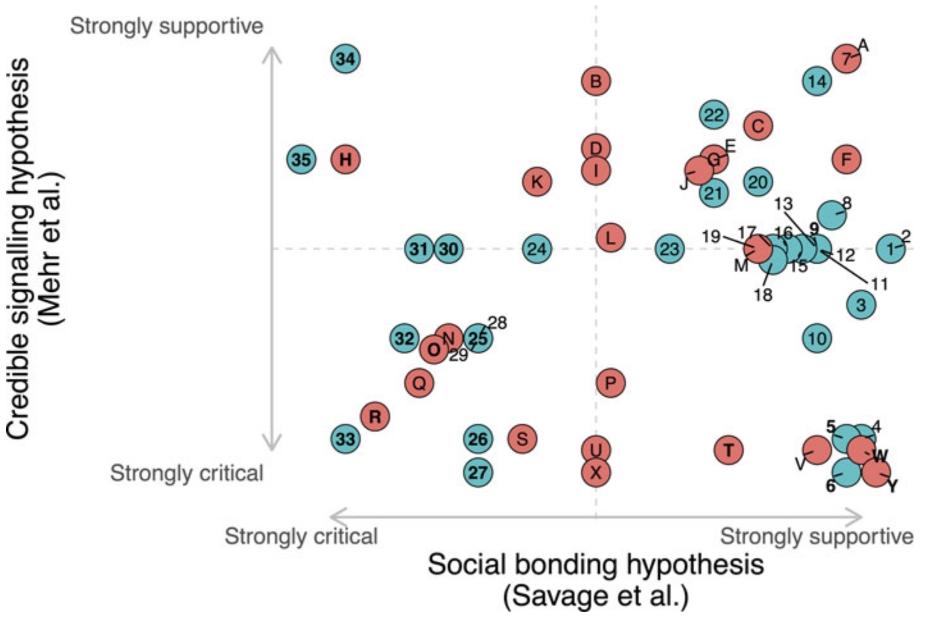


Figure R1. Visual comparison of the 60 commentaries responding to the pair of target articles, based on our subjective evaluation of the degree to which they are supportive or critical of each target article. Figure R1 plots the average of subjective ratings by PES and PL on a scale from −10 ("strongly 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 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. **Popescu, Oesch, & Buck**	E. Snyder & Creanza
6. **Rendell, Doolittle, Garland, & South**	F. Kasdan, Gordon, & Lense
7. Morrison	G. Levitin
8. Ravignani	H. **Kennedy & Radford**
9. **Eirdosh & Hanisch**	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. **Trainor**
16. Ashley	P. Scott-Phillips, Tominaga, & Miton
17. Brown	Q. Számadó
18. Hansen & Keller	R. **Stewart-Williams**
19. Pfordresher	S. Tichko, Bird, & Parker
20. Dissanayake	T. **Bowling, Hoeschele, & Dunn**
21. van Mulukom	U. Levaida
22. Gingras	V. Trehub
23. Fritz	W. **Wood**
24. Zhang & Shi	X. Lieberman & Billingsley
25. **Honing**	Y. **Killin, Brusse, Currie, & Planer**
26. **Wald-Fuhrmann, Pearson, Roeske, Grüny, & Polak**	
27. **Cross**	
28. Hannon, Crittenden, Snyder, & Nave	
29. Iyer	
30. Kraus & Hesselmann	
31. Atzil & Abramson	
32. **Patel & von Rueden**	
33. **Harrison & Seale**	
34. **Merker**	
35. **Verpooten & Eens**	

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.[1] 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 signalling 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 **Grahn, Bauer, & Zamm (Grahn 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).[2]

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).[3] 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, Llona, & 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 (**Szamado),** 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 BRECVEMA 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 Grahn, 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., Llona, 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*, 451–545.

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.

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., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., … Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468), 957–970. https://doi.org/10.1126/science.aax0868.

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.

Moon, S.-J., & Lindblom, B. (2003). Two experiments on oxygen consumption during speech production: vocal effort and speaking tempo. *Proceedings of the 15th International Congress of the Phonetic Sciences*, Barcelona, pp. 3129–3132.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences, 39*(e1). https://doi.org/10.1017/S0140525X14001356.

Oberweger, K., & Goller, F. (2001). The metabolic cost of birdsong production. *Journal of Experimental Biology*, *204*, 3379–3388.

Patel, A. D. (2010). Music, biological evolution, and the brain. In M. Bailar (Ed.), *Emerging disciplines* (pp. 91–144). Rice University Press.

Patel, A. D. (2008). *Music, language and the brain*. Oxford University Press.

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.

Penn, D. J., & Számadó, S. (2020). The handicap principle: How an erroneous hypothesis became a scientific principle. *Biological Reviews*, *95*, 267–290.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, *33*(2), 185–191.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences*, *7*, 362–336.

Peters, G. (2002). Purring and similar vocalizations in mammals. *Mammal Review*, *32*, 245–271.

Pinker, S. (1997). *How the mind works*. Norton.

Poggi, J. (2013). Interview: Sean "Diddy" Combs says you'll soon tune into Revolt TV, and so will brands. *AdAge*. https://adage.com/article/special-report-music-and-marketing/interview-sean-combs-revolt-tv/244350.

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.

Rousseau, J.-J. (1760/1998). *Essay on the origin of languages and writings related to music* (J. T. Scott [ed. & trans.]). University Press of New England.

Ryan, M. J. (1985). *The túngara frog, A study in sexual selection and communication*. University of Chicago Press.

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., 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.

Sewall, K. B., Young, A. M., & Wright, T. F. (2016). Social calls provide novel insights into the evolution of vocal learning. *Animal Behaviour*, *120*, 163–172.

Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature*, *479*(7372), 219–222.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Harvard University Press.

Számadó, S. (2011). The cost of honesty and the fallacy of the handicap principle. *Animal Behaviour*, *81*, 3–10.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Für Tierpsychologie*, *20*, 410–433.

Thompson, W. F., & Olsen, K. N. (Eds.). (2021). *The science and psychology of music: From Beethoven at the office to Beyoncé at the gym*. Greenwood.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*(3), 248–251.

Titze, I. R. (1994). *Principles of voice production*. Prentice Hall.

Tomlinson, G. (2015). *A million years of music: The emergence of human modernity*. MIT 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.

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.

Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2000). *The origins of music*. MIT Press.

Ward, S., Speakman, J. R., & Slater, P. J. B. (2003). The energy cost of song in the canary, *Serinus canaria*. *Animal Behaviour*, *66*, 893–890.

Washington State University. (2020). War songs and lullabies behind origins of music. *Science Daily*. www.sciencedaily.com/releases/2020/10/201026095422.htm.

West, S. A., Griffin, A. S., & Gardner, A. (2008). Social semantics: How useful has group selection been? *Journal of Evolutionary Biology*, *21*, 374–385.

Toward a productive evolutionary understanding of music

Samuel A. Mehr[a,b,c,*], Max M. Krasnow[a,*], Gregory A. Bryant[d,e,*] and Edward H. Hagen[f,*]

[a]Department of Psychology, Harvard University, Cambridge, MA 02138, USA; [b]Data Science Initiative, Harvard University, Cambridge, MA 02138; [c]School of Psychology, Victoria University of Wellington, Wellington 6012, New Zealand; [d]Department of Communication, University of California Los Angeles, Los Angeles, CA 90095, USA; [e]Center for Behavior, Evolution, and Culture, University of California Los Angeles, Los Angeles, CA 90095, USA and [f]Department of Anthropology, Washington State University, Vancouver, WA 98686, USA.

sam@wjh.harvard.edu, https://mehr.cz
krasnow@fas.harvard.edu, https://projects.iq.harvard.edu/epl
gabryant@ucla.edu, http://gabryant.bol.ucla.edu
edhagen@wsu.edu, https://anthro.vancouver.wsu.edu/people/hagen

doi:10.1017/S0140525X21000030, e122

Abstract

We discuss approaches to the study of the evolution of music (sect. R1); challenges to each of the two theories of the origins of music presented in the companion target articles (sect. R2); future directions for testing them (sect. R3); and priorities for better understanding the nature of music (sect. R4).

In our target article, we proposed that human music evolved from territorial advertisements and contact calls, two types of vocal signals that are widespread in primates and other species. Many nonhuman primates signal territory ownership with loud vocal calls, and in some ape species these involve synchronized vocalizations. We suggested that group musical performances evolved in this context, to credibly signal group size and quality to enemies and allies. Furthermore, we suggested that musicality evolved in a second domain that also involves credible signaling: the contact calls found in nonhuman species that serve to keep mothers and infants in close proximity. In humans, we think that two complementary signals are implicated: aversive infant cries to attract attention and soothing infant-directed songs (e.g., lullabies) that credibly signal parental attention to the infant.

We argued that selection in these two domains of signaling behaviors shaped mechanisms for producing and perceiving melodic and rhythmic signals that constitute core components of musicality, providing a foundation for musical creation that was subsequently elaborated by cultural-evolutionary processes. Our theory, in which music encodes contentful information transmitted between individuals and groups, contrasts with the hypothesis in the companion article by Savage et al., that is, that music evolved in order to strengthen social bonds via the joint behavior of musicians.

Here, we address critiques of both theories along with ideas for future tests of the theories and future priorities for research on the evolution of music. We focus, in part, on disambiguating between proximate- and ultimate-level explanations for complex behaviors, emphasizing our view that the social bonding effects proposed by Savage et al. constitute proximate mechanisms that serve ultimate functions of credible signaling.

---

*All authors contributed to this response and are listed in reverse order of seniority.

R1. Adaptations and byproducts: What they are, what they aren't, and why it matters

Although the commentaries included diverse perspectives on the questions of *how* to study the evolution of music and *why* to do so, our approach draws most on the adaptationist framework (Darwin, 1859; Williams, 1966). Williams argued that adaptations are characterized by the form–fit connection between evolved *design features* and recurrent adaptive problems that those features solve. This relationship results, over time, from natural selection removing relatively worse alternative designs from a reproductive lineage.

Not all features are design features. Positive selection for a design feature also creates *byproducts*, that is, features associated with an adaptation but not directly selected for. For example, human bones look whitish because of their high concentration of hydroxyapatite, a mineral that facilitates the effective operation of muscles on rigid bones. The whiteness of bones is a human universal and appears in other species, but it is not an adaptation: it is a byproduct of design for bone rigidity.

Consider another example: many humans ride bicycles, an enjoyable (to some) and functional (to some) activity. That cycling is enjoyable or functional is not evidence for a "cycling adaptation": bicycles did not exist in our ancestry, so the mechanisms underlying our ability to ride them cannot be because of past selection for cycling. Cycling-ability must be a byproduct of other evolved traits (e.g., adjusting balance to a moving center of gravity). Adaptations and byproducts constitute the features that characterize a species' nature.[1]

The majority of traits in any species are byproducts: structural concomitants of adaptations (e.g., bone-whiteness), new uses of adaptations (e.g., cycling), reliable ways that adaptations fail (e.g., prosopagnosia/face-blindness), and so forth. Thus, an appropriate null hypothesis is *always* that a feature is a byproduct: the prior on adaptation is low, or in Williams's terms, adaptation is a "special and onerous concept" (p. 4). Therefore, we agree with **Bowling, Hoeschele, & Dunn** (**Bowling et al.**); **Dissanayake; Harrison & Seale**; **Killin, Brusse, Currie, & Planer** (**Killin et al.**); **Leivada**; **Harrison & Seale**; **Lieberman & Billingsley**; **Moser, Ackerman, Dayer, Proksch, & Smaldino** (**Moser et al.**); **Pinker**; **Stewart-Williams**; **Tichko, Bird, & Kohn** (**Tichko et al.**); **Trainor**, who reference the difficulty of demonstrating music-specific adaptations.

Unlike byproducts, adaptations have reliable effects that *explain their structure*. An adaptationist approach focuses on the fit between the *structure of a particular adaptive problem* posed by the environment (including the organism itself) and *the particular design features predicted to solve it*. To us, the key open questions about the evolution of music are what those adaptive problems were in human ancestry, if any, and what design features in our psychology of music, if any, solved them.

One complication that arose immediately is the assumption that adaptations exist in a vacuum. **Killin et al.** write "to ask whether…cognition [is] 'adapted'…implies a causal simplicity which overlooks music's likely complex, niche-constructed, coevolutionary path." **Trainor** argues "…the evolution of musical capacities will likely not consist of one adaptation, but rather a long sequence of adaptive, exapted, and cultural influences that interact…." **Tichko et al.** argue "…evolutionary theorists have a tendency to conflate design and adaptation, while ignoring or underestimating the role of non-adaptive evolutionary processes, that can produce organismal complexity."

We think this position is a red herring. Traits evolved by natural selection because they reliably caused certain effects, which, through various causal pathways, increased fitness. **Killin et al.** rightly question the idea of explaining the evolution of the human hand via its role in tool-making, but wrongly imply that this undermines an adaptationist approach. Their mistake is to conflate the direct causal effects of a putative adaptation with its (possibly numerous) fitness-increasing consequences. The human hand shows evidence of design to grasp and manipulate objects (its evolved function), a capability that increased fitness via numerous causal pathways (e.g., grabbing tree limbs, making tools, and throwing projectiles). We proposed that key features of music evolved to reliably cause particular inferences in the brains of observers by overtly transmitting covert information (e.g., parental attention; the willingness and ability of individuals to cooperate). Those inferences would have increased fitness via multiple pathways involving cooperative and agonistic relationships among individuals and groups.

The points raised by **Killin et al.**, **Trainor**, and **Tichko et al.** are widely understood, and do not undermine adaptationism, as every adaptationist theory is a theory of coevolution. Understanding the heart as an adaptation for pumping blood, for instance, does *not* imply that its evolution was simple or uncomplicated, did not depend on coevolution with the circulatory system, was unconstrained by fluid dynamics, and so forth. To confront the argument that natural selection is responsible for the form of a particular trait may reliably conjure such notions, but they are false.

Furthermore, natural selection's actions on ancestral populations produce the design of traits in a contemporary species *in a temporally unidirectional fashion*: the past explains the present. As emphasized by **Bowling et al.** and **Pinker**, it is only this directional effect that licenses evidence for design, and not, for example, the functions that a trait is useful for today; functions for which a trait might in principle be used; or functions that are intricate, extraordinary, enjoyable, fascinating, worthy of study, or otherwise interesting (despite claims to the contrary by **Bowling et al.**; **Cross**; **Dubourg, André, & Baumard** (**Dubourg et al.**); **Hannon, Crittenden, Snyder, & Nave** (**Hannon et al.**); **Patel & von Rueden; Számadó; Trevor & Frühholz**). Such characteristics play no causal role in an evolutionary theory. **Scott-Phillips, Tominaga, & Miton** (**Scott-Phillips et al.**) argue this point well, contrasting the social bonding and credible signaling theories in their treatment of culture in the proposed musical adaptations' proper versus actual domains (Sperber, 1994), respectively. This distinction is essential.

In particular, as **Pinker** notes, when an adaptation's proper domain is to motivate "ancestrally rational" action (choosing high-calorie foods, finding mates, communicating social intentions, etc.) the resulting actual domain includes cases where the "ancestrally rational" cue is hijacked by a technology that satisfies the cue, without actually solving the adaptive problem. Such "hijacked" cases, such as the sweet taste of artificial sweeteners, do not jeopardize a theory of the adaptation's proper domain; they should be expected. Because the actual domain of an adaptation in our modern-day environment can differ substantially from its proper domain, confusion between the effects of modern-day music on listeners with their effects in ancestral conditions should be avoided.

This is one sense in which we think the byproduct hypothesis (see sect. 3.1 in **Mehr et al.**) is correct: Once the human mind evolved some basic properties of a music faculty, these properties would be hijacked and shaped by cultural evolution (see **Scott-Phillips et al.**). Just as the language faculty's evolved design enables the cultural evolution of languages, the music faculty's evolved design enables the cultural evolution of different songs and musical traditions.

R2. Challenges to theories of the origins of music

Several commentators agreed with our critique of the social bonding hypothesis and/or provided new critiques. **Juslin** notes that predictions of the social bonding theory are "…either too trivial or too vague to distinguish between rival hypotheses" (see also **Pinker**; **Popescu, Oesch, & Buck** (**Popescu et al.**); **Zentner**). **Fritz** calls it "so broad and sweeping it will be challenging to test, prove, or falsify…." **Zhang & Shi**'s cross-species and neural evidence support our suggestion that language is a more plausible mechanism for social bonding than is music. **Verpooten & Eens** point out that singing is not associated with social complexity across species, contra **Savage et al.**'s prediction.

One consensus that emerged from the commentaries, which we also alluded to, is the idea that social bonding – which we and **Savage et al.** agree is associated with music production – is a plausible *outcome* of credible signaling. **Kennedy & Radford** suggest that two components of the social bonding effects predicted by Savage et al. rely on music acting as a credible signal (see also **Gardiner**). Similarly, **Sachs, FeldmanHall, & Tamir** (**Sachs et al.**) suggest that coalition formation is a likely point of social bonding in music (a primary context that we described for credible signaling via music). Indeed, in our target article we proposed that musical behavior provides information to the musicians: "Within groups, musical performances might also create common knowledge of decisions to cooperate, which could serve group coordination and cooperation." Making music carries probabilistic information about the coordination of mental states and intentions of the music makers, changing the social affordances they represent (i.e., the sense of social connection highlighted by **Gabriel & Paravati**). Manipulating others' impressions of these social affordances is an example of music functioning as a credible signal.

If social bonding is a plausible outcome of credible signaling via music, what of the evidence that music evolved as a credible signal? A variety of critiques of the theory arose in the commentaries.

First, some authors misunderstood or mistook our theory for more than what we intended it to be. **Hansen & Keller** and **Harrison & Seale**'s commentaries imply that we argued for a unitary mechanism underlying the evolution of music. We didn't: As **Gingras** summarized, there are many musical contexts to explain and a credible signaling account only explains some of them. Similarly, **Wald-Fuhrmann, Pearson, Roeske, Grüny, & Polak** (**Wald-Fuhrmann et al.**) imply that our theory of adaptive problems shaping particular features of human music discounts the existence of other features, contexts, or uses of music. It doesn't. **Pinker** asks whether the two contexts for music we focus most on (coalition signaling and infant care) are "more universal" than other contexts,[2] but the answer is not necessarily relevant: we cannot explain *all* behavioral contexts for music. One theory is unlikely to explain every instantiation of a complex psychological phenomenon; ours is no exception.

The narrow scope of the credible signaling theory is a virtue. Contrary to a throwing-up-of-the-hands (e.g., **Savage et al.**'s statement, "We may never know with certainty the precise ancestral adaptive conditions or specific genetic mutations involved in the evolution of musicality"); and the open-ended flexibility of the social bonding theory "about the timeline, precedence, and relative contributions of cultural and biological evolution," as **Trehub** describes it, a narrower scope enables the generation of testable predictions that are distinct from and/or in opposition to the predictions of other theories (social bonding or otherwise). Our scope was disappointing to some commentators, but not others (we agree, for instance, with **Killin et al.** that there is no "unitary proper function" of music). Another virtue of limiting the scope is that it more clearly delineates the areas of human psychology that are best explained by cultural evolution, including how cultural processes apply evolved mechanisms to new functions (**Dubourg et al.**); providing an explanation of the origins of the proper domains of music is a step toward understanding attractor spaces guiding the formation of new actual domains (**Scott-Phillips et al.**).

A variety of commentators agreed with the premise of the credible signaling theory, but suspected the selective pressures imposed by coalitional signaling and/or parent–offspring conflict over parental attention were not strong or reliable enough to produce adaptations. **Trehub** and **Dissanayake** suggest that the safety problem for infants, which we proposed was solved via parental attention elicitations (Mehr & Krasnow, 2017), is less of a problem than we think (because infants were carried and fed on-demand more than is currently typical). We agree that ancestral parenting differed from modern parenting, but three findings undermine this criticism. First, infants in traditionally-living societies are neither exclusively carried nor carried exclusively by mothers (Fouts, Hewlett, & Lamb, 2001); as Lozoff and Brittenham (1979) put it, "When not held, the hunter-gatherer baby has complete freedom of movement except in emergencies, both in early infancy and after crawling" (p. 480), implying a link between parental attention and infant safety. Second, hunter-gatherer infants are likely carried more because the risks of injury or death are elevated. Even a rare lapse of attention over years of care could result in a large fitness cost (e.g., an infant's death), causing the evolution of risk-averse strategies (Hintze, Olson, Adami, & Hertwig, 2015) such as continual mother–infant contact calls. Finally, whether carried or not, infant mortality was far higher for our ancestors than it is for present-day humans (Kramer & Greaves, 2007). Thus, in human ancestry infants could be safer from harm than they were and additional parental attention could help.[3]

Lieberman & Billingsley argue that infant-directed song has no advantage over infant-directed speech. But, infant-directed singing is less compatible with maintaining conversation with others, increasing the credibility of its attentional signal. Indeed, as **Trehub** and colleagues have found, song is a better soother of infant distress than speech (Corbeil, Trehub, & Peretz, 2016).

Moser et al. argue that our analysis of the social selection pressures was too limited: that adaptive problems at the group level, once hominins transitioned to multi-level social organization, "almost certainly had an effect on the evolution of human music diversity." We agree. We argued that "complex forms of social organization likely set the stage for the evolution of complex credible signals," outlining the implications of the human transition to a multi-level society (see Hagen & Bryant, 2003). Indeed, in line with Moser et al.'s emphasis on "group-level traits," Hagen and Hammerstein (2009) sketch the central role of music in the evolution of agent-like properties of human groups. Regarding the perennial debate over group/multilevel selection: Most theorists seem to agree that both the bottom-up gene's eye view and the top-down "group" view provide insights, but in most cases are mathematically equivalent (although the debate continues; Birch, 2019; Queller, 2020).

Patel & von Rueden argue that cross-cultural data might not support the credible signaling theory, providing examples of small-scale societies, such as the Amazonian Tsimane, where group music production is limited. Tsimane music perception shares at least some traits with other cultures, however (e.g., mental scaling of pitch that is logarithmically organized; Jacoby et al., 2019), and evidence for the universality of group music-making across societies is substantial (Mehr et al., 2019). But, we agree with Patel & von Rueden's call for studies explaining variability in music-making. That variability is likely to be related to whether or not the costs of time- and energy-intensive group music/dance performances exceed their benefits; they might not among the Tsimane, who experience substantial nutritional constraints (Blackwell, Urlacher, Beheim, & Kaplan, 2017) and, as Patel & von Rueden note, have likely not experienced violent intergroup conflict for centuries.

Finally, several commentators raised questions concerning the fit between the adaptive problems posed by territorial signaling and music. **Lieberman & Billingsley** find it "unclear why signals of formidability need be credible," because "Predators don't signal prey from afar." But, competitors are not predators and the asymmetric war of attrition is not a predator-prey model. As we explained in our target article, there is a well-documented "prior-residence" effect favoring owners over intruders (e.g., Kokko, Jennions, & Brooks, 2006) that selected for credible signals of occupancy in countless species, and, for group-defended territories, credible signals of coalition size, and quality. Relatedly, several commentators (Lieberman & Billingsley, **Pinker**, and **Zentner**) assumed that territorial signals are usually aggressive; because music usually is not, this would seem to weaken our case. The function of territorial advertisements, however, is to credibly advertise occupancy so as to *avoid* aggressive encounters and fights (Kokko et al., 2006). **Stewart-Williams** argued that the territorial signals of our ape relatives, the chimpanzees, are not synchronized, so those of ancestral humans probably weren't either (see also **Killin et al.**), but as we discussed, the territorial signals of many other ape relatives *are* highly synchronized, and convergent evolution is common.

Lieberman & Billingsley also note that, historically, music was used to coordinate large armies; historical evidence also suggests that music was used to instill fear in enemies (Swope, 2009). Moreover, the use of drums, gongs, flags, and trumpets to coordinate large military operations is conceptually close to their symbolic use to signal such coordination to others that we propose for small prehistoric coalitions. **Stewart-Williams** argued that subtle differences in temporal synchrony carry little information about coalition quality: Why not evaluate dimensions of coalition quality directly? We argue that such direct evaluation would, impractically, require extensive observation of a coalition, whereas a music/dance display that took extensive practice to perfect, encodes substantial information about willingness and ability to cooperate, and could be evaluated rapidly, as in a feast (Hagen & Bryant, 2003). Finally, **Wood** argues that, contrary to our model of competition for allies, "cooperative pacts are only rarely freely chosen" but instead reflect "some pre-standing, socially normative or obligatory relationship." But, his citations do not support his claims and the literature on feasting and alliance formation emphasizes competition among both individuals and groups (Hayden, 2014; Hayden & Villeneuve, 2011).

R3. Future directions for testing the idea that music is a credible signal

Many commentators raised interesting avenues for testing the credible signaling theory. For example, **Lumaca, Brattico, & Baggio** (**Lumaca et al.**) suggest that signaling games could help test the ways credible signals operate in music. We appreciate this approach and agree that multi-player signaling games have been useful in explaining the evolution of cooperation. Applying signaling games to music may be complicated, however, by the fact that participants already have the ability being studied (e.g., mapping tone sequences to affective meanings), which may cloud inferences about the ability's evolution. **Akkermann, Can Akkaya, Dermiel, Pflüger, & Dresler** (**Akkermann et al.**) propose a new methodological application from another field ("sleep wearables"), which may provide a means to explore the biophysical mechanisms of effects of music on affect, emotion, and psychophysiology (see also Bainbridge & Bertolo et al., 2021). And **Sievers & Wheatley** raise interesting questions concerning the degree to which universal forms of lullabies reflect basic properties of arousal in the vocalizations of many species; we are eager to test this hypothesis directly, in particular via the combination of corpus work with citizen science approaches (as in Hilton & Crowley-de Thierry et al., 2021; Mehr et al., 2019).

Tichko et al. suggest applying tools from population genetics and comparative genomics to directly test for the presence (or absence) of adaptations for musicality. Although we evidently disagree on the tenability of evidence for design (see sect. R1), no matter: This is an entirely reasonable program of research to which two of us have contributed (Kotler, Mehr, Egner, Haig, & Krasnow, 2019; Mehr, Kotler, Howard, Haig, & Krasnow, 2017; Mehr & Krasnow, 2017). But, much more can be done in this area, as **Honing**, **Trehub**, and others have previously suggested (Honing, ten Cate, Peretz, & Trehub, 2015). Indeed, **Kasdan, Gordon, & Lense** (**Kasdan et al.**) propose testing musical interventions in genetically informative populations, which we also endorse.

Finally, several commentators suggested that cross-species analyses can test predictions of the credible signaling hypothesis. For example, **Snyder & Creanza** suggest a comparison between culturally transmitted songs in birds and infant-directed songs in humans. As in songbirds, species-specific songs might have had a role in mate selection and other inter-species interactions in hominins, an idea that is supported by the increasing fossil and genetic evidence that the human lineage overlapped spatially and temporally with multiple hominin lineages, and that hybrids had reduced fertility (e.g., Ackermann, Arnold, Baiz, & Zinner, 2019; Sankararaman, Mallick, Patterson, & Reich, 2016). **Ravignani** proposes that a cross-species comparison of honest signaling via vocalization might help to identify core features of musicality.

In principle, we agree with these views, although we caution that interpreting music-like behavior in nonhumans risks anthropomorphism and loose evolutionary logic. For instance, the examples **Hattori** raises of "rhythmic body movements" in nonhumans may have nothing to do with music (Bertolo, Singh, & Mehr, 2021). Cross-species comparisons are inherently difficult, as we pointed out in our target article. Comparative analyses, however, can provide valuable clues regarding pre-existing mechanisms that potentially inform the effort to identify music-related adaptations in humans, and so we look forward to the results of further cross-species work.

R4. Priorities and open questions on the nature of musicality

The discussion of both target articles revealed that fundamental questions about the human psychology of music have yet to be answered. We hope that one productive outcome of the present discussion is to spark new investment in basic research on musicality (see **Honing**), in several areas.

R4.1 Musical esthetics

Why are humans so motivated to seek out and to produce music in the first place? **Dubourg et al.**, **Kraus & Hesselmann**, **Pinker**, **Sievers & Wheatley**, **Trainor**, and others are right to ask how evolutionary theories of music can explain the role of esthetics in music – its "most blazingly obvious feature" (Pinker).

Measuring esthetic value in music is a substantial challenge. The recommendation engines[4] of the world's largest music streaming platforms often use minimal musical information in their attempts to predict whether a given user will enjoy listening to a particular song, instead modeling listener preferences using other information about the similarity of users, such as the particular clusters of songs or artists in common across users' playlists, regardless of musical content (Jacobson, Murali, Newett, Whitman, & Yon, 2016). This approach is consistent with experimental work demonstrating the value of social information in musical preferences (Salganik, Dodds, & Watts, 2006), and, in real-world Spotify data, the fact that musical preferences and microgenres are predictable from users' age, sex, language, and geographical proximity (Schedl, Bauer, Reisinger, Kowald, & Lex, 2021; Way, Garcia-Gathright, & Cramer, 2020). Therefore, although we agree with the commentators that developing an understanding of esthetic preferences in music is a high priority for musicality research, we do not expect it to be easy.

Three considerations of the credible signaling hypothesis are relevant. First, as a broad generalization, humans evolved to enjoy engaging in activities that increased our biological fitness. The credible signaling hypothesis posits fitness benefits to music, so it should then be no surprise that producing and listening to music is pleasurable.

Second, aspects of the credible signaling theory are evident in modern musical activities: beyond the daily use of music in families (Custodero, Rebello Britto, & Brooks-Gunn, 2003; Mehr, 2014; Mendoza & Fausey, 2019; Trehub, Hill, & Kamenetsky, 1997), child- and infant-directed music are also highly successful commercial enterprises (e.g., Raffi); the related genre of "relaxation music" is also popular in adults (**Akkermann et al.**). Popular music is also commonly incorporated into group sporting events of all sorts, suggesting a link between music and coalitional competitions. For example, Eurovision, a competition among pop music groups representing each European country, who perform all genres of music, attracts close to 200 million viewers a year.

Moreover, music industry marketing tactics are permeated with elements of coalitional signaling. Music in the top five musical genres by global sales (Hip-Hop/R&B, Rock, Pop, and Country; MRC/Billboard, 2021) is typically produced by small groups that adopt many elements of coalitional or ethnic identities in their performances, such as distinctive clothing, accessories, dialects, tattoos, and, importantly, political goals. This is especially apparent in enormously popular sub-genres (e.g., k-pop, gangsta rap, and grunge); such signals of group identity may even be detectable in music for infants (Mehr, Song, & Spelke, 2016; Mehr & Spelke, 2017; Xiao et al., 2017).

Third, a key characteristic of musical esthetics is its balancing of predictability and surprise: The music of both industrialized and small-scale societies contains acoustic elements that are patterned according to power laws or other Zipfian-like distributions (Levitin, Chordia, & Menon, 2012; Manaris, Roos, Krehbiel, Zalonis, & Armstrong, 2012; Mehr et al., 2019; Zipf, 1949). How do these and other general principles of musical esthetics arise, should they be reliable features of music across cultures?[5] **Sievers & Wheatley** are right that the credible signaling theory does not fully explain "why music sounds the way it does," but we do argue that patterned variability in recurrent acoustic forms present in music would be essential to convey the content of a credible signal (Hagen & Bryant, 2003; Hagen & Hammerstein, 2009), and we suggest a mechanism for the elaboration of that content (i.e., arms-race coevolution; Mehr & Krasnow, 2017). The details are still murky, however; hierarchical perception of the constituent parts of music (**Hilton, Asano, & Boeckx** [**Hilton et al.**]) could in principle facilitate signal transmission, and draw on other forms of vocal signaling, such as emotional expression (see **Zentner**; Sievers & Wheatley). Indeed, music may be considered the group-level analog of emotional signaling (Bryant, 2013; Hagen & Bryant, 2003; Hagen & Hammerstein, 2009).

R4.2 Music and language

The relationship between music and language figured prominently among the commentaries. For example, **Leivada** presented features of music she argued are derived from language (see also **Lieberman & Billingsley**) and **Számadó** stressed the importance of developing accounts of music and language coevolution. Music and language clearly share several computational principles, many of which are related to auditory processing (see **Trainor**). Similar to others (Doelling & Poeppel, 2015; Jacoby, Margulis, Clayton, & Wald-Fuhrmann, 2020; Patel, 2008), including **Savage et al.**, we think that developing a deeper understanding of the similarities and differences between music and language, and the evolution of those similarities and differences, is a priority, especially insofar as music and speech are directly intertwined (e.g., lyrical music is a universal; Mehr et al., 2019).

Although many distinct cognitive and perceptual traits share processing principles (e.g., statistical learning is important for both vision and speech), these connections do not necessarily imply a causally related evolutionary history. Shared processing principles should be evaluated according to whether a general principle underlies the functional organization of the respective systems. For example, hierarchical organization is a principle of language with clear analogs in music, and as **Hilton et al.** describe, across multiple other domains (including metacognition, action planning, and auditory scene analysis). Although Hilton et al. propose that such similarities are byproducts of domain-general cognitive mechanisms (see also **Sievers & Wheatley**), shared principles across domains can manifest independently in specialized devices as a result of selection converging on similar efficient solutions to distinct adaptive problems.

Furthermore, many manifestations of music incorporate linguistic phenomena in different ways – a major challenge for theories of music and language origins is to distinguish between shared evolutionary history and the effects of cultural evolution. For example, **Levitin** proposed knowledge songs as a mechanism for information transfer prior to written language, citing the well-established effect of musical enhancement of verbal encoding. This is a good example of cultural-evolutionary forces acting on a pre-existing musical capacity. Others stressed the role of music in inducing emotions for storytelling, another cultural-evolutionary effect (**Trevor & Frühholz**). And **Cross** describes deep integrations of music and language in cultural traditions, although we note that such connections may co-exist with form–fit relationships between adaptive problems and music-specific adaptations.

Although the credible signaling theory is agnostic regarding the relative timeline of language and music evolution, we speculate that the two communicative systems evolved in tandem, and some shared processing resources could reflect that fact. That being said, we see language and music as also having distinct computational and behavioral properties that solve different adaptive problems: in our view, language is a "cheap" communication system for cooperative signalers and receivers whose interests generally align (with exceptions, like indirect speech; Pinker et al., 2008), whereas music is a credible communication system for signalers with cooperative intent but who have conflicts of interest with receivers.

Moreover, as we described in our target article, in numerous taxa, including apes and other primates, song-like vocal signals have evolved that, like music and unlike language, comprise repetitive sequences of acoustic events (analogous to "notes") that are loud and directed at physically distant receivers. The convergent evolution of similar acoustic signals in diverse, often distantly related taxa are evidence for common selection pressures, such as territorial signaling, mate quality signaling, and contact calls. Testing the distinct predictions of theories of the biological and cultural evolution of music and language should, therefore, be a priority.

R4.3 Synchrony and other rhythmic phenomena

Several commentators (**Gabriel & Paravati; Grahn, Bauer, & Zamm; Hattori; Pfordresher; Wood**) understood our position as one that denies prosocial effects of synchrony, and any role for synchrony in the evolution of musicality. We remain open to the possibility of such effects, but are concerned that previously reported causal effects of musical behavior on prosociality (e.g., many papers cited in Savage et al.; Gabriel & Paravati; and others) are undermined by demand and/or expectancy characteristics (Atwood, Mehr, & Schachner, 2020). Furthermore, experiments have not yet disentangled the potential uncontaminated direct effects of synchrony on bonding from effects whereby experiencing or anticipating synchrony influences bonding by altering perceptions of social affordances. Put another way, synchrony may be a proximate mechanism employed in the service of signaling a bond, but not actually creating it. We see this as a central point of departure between our account and that of Savage et al. (see also sect. R2). If synchrony serves as a proximate means for engaging in credible signaling, then we should expect connections between musical behavior and reward systems which are detectable even once improvements in experimental design are made (Atwood et al., 2020). Indeed, many other behaviors that cause endogenous reward also enhance bonding, such as laughter; proposing that laughter evolved ultimately for social bonding, for example, would lack an explanation of laughter's communicative functions (Bryant, 2020).

R4.4 Understanding musical diversity via cultural-evolutionary approaches to music

Several commentaries pointed to the substantial effects of cultural evolution on music production and music perception worldwide (**Lumaca; Moser et al., Scott-Phillips et al.**). We find this topic to be one of the most exciting and interesting areas for research on the psychology of music: It is obvious that cultural evolution plays a deep role in the diversity of music's manifestations in contemporary society and across cultures.

How can this role be explained? A first step is to explain the selection pressures that lead to specific core competencies in proper domains of music production and perception. As we described in our target article, we expect cultural-evolutionary processes to have acted on these capacities, as well as many other related, nonmusical abilities (e.g., auditory scene analysis and language) to produce the diversity of musical behavior that exists today (via social learning, including horizontal and vertical transmission, cumulative cultural phenomena, etc.; see sect. 5.2 of Mehr et al.). In our view, understanding the cultural evolution of music is a complementary, but separate task from characterizing the aspects of our psychology of music that were shaped by natural selection. The complexity in music introduced by cultural evolution makes the identification of proper mechanisms difficult, to say the least, and arguably has contributed to the confusion and disagreement that characterizes many theoretical treatments of the evolution of music. Disentangling the effects of the biological and cultural evolution of music is a productive strategy, we think.

With this in mind, we highlight two brief points. First, the credible signaling theory identifies at least two music-specific capacities that map onto cultural attractors (see **Scott-Phillips et al.; Dubourg et al.**), namely, pitched and rhythmic vocalizations used in reliably occurring signaling contexts. One immediate question is whether it is possible, even in principle, that music production in these limited contexts was elaborated via cultural-evolutionary processes to produce *other* musical contexts (which might or might not involve credible signals themselves). Cross-cultural studies, especially those that account for the relatedness of cultures (in a fashion similar to ideas mentioned by **Tichko et al.**) would provide evidence for or against this idea.

Second, the credible signaling theory highlights a possible mechanism for cross-population variation. Credible signals incur opportunity costs, and in the case of coalitional signaling, substantial energetic costs. The psychology of music may be designed to only pay these costs if they are outweighed by the benefits. As discussed above in the Tsimane example (raised by **Patel & von Rueden**), the benefits almost certainly vary across socio-ecological contexts. Hence, via various psychological mechanisms, including individual and social learning, individual and population frequencies of lullabies might depend on, for example, local risks to infants, when and how they are carried throughout the day, and the availability of alloparental care. The frequency and complexity of group musical performances might depend on the local intensity of competition for allies and territory, as well as the extent to which groups can subsidize musically talented individuals (division of labor).

R4.5 The basic facts of music

Perhaps the largest open question about the psychology of music lies at the intersection of evolutionary science and cognitive psychology: How are human minds built for music? This question was alluded to in several commentaries. For example, **Trainor** asks "why does music have the pitch structure it does?," arguing that human pitch perception is a byproduct of auditory scene analysis. Although many aspects of pitch perception are relevant to music perception, we suspect that the phenomenon to explain in music perception has more to do with the perception of meaningful musical units in a hierarchical context: the melodic and rhythmic structures that turn up with surprising regularity across cultures, and which are readily perceived by naïve listeners (Mehr et al., 2019). **Hilton et al.** agree, although they posit that this hierarchical structure is a byproduct of domain-general cognitive mechanisms, such as action-planning. Or perhaps hierarchical structure in music and language are both derived from a common system, with language evolving much greater hierarchical complexity in grammatical structures.

We hesitate to make a strong claim here but wonder whether hierarchical structures of tonality and meter could provide an effective platform for the transmission of credible signals. In our view, before investigating a possible link between the credible signaling theory and the evolution of hierarchical music perception, however, it is important to first test whether such hierarchies are indeed the structural components of music perception that need to be explained – insofar as understanding them can lead to a deeper understanding of how "music-as-we-know-it" is functionally constrained (as **Sievers & Wheatley** put it).

In this sense, we agree with **Honing**'s call for further research into musicality so as to "identify the core constituent components of musicality." To us, studying music production across cultures is a prerequisite for a comprehensive understanding of the psychology of music perception, simply because one needs to know what it is in music production that should be examined in music perception. This is especially so given the preponderance of WEIRD research in the psychology of music (Jacoby et al., 2020). We would not presume, however, that the credible signaling theory can explain all aspects of music perception (just as it cannot explain all contexts of music). The present discussion makes it clear that deep questions on the nature of music perception remain open. We eagerly anticipate research answering them.

R5. Concluding remarks

In the course of reading and thinking about theories of the evolution of musicality, or the evolution of any trait, it seems prudent to step back and ask: who cares? In human evolution we can rarely observe the counterfactual of an adaptationist hypothesis. Why bother?

This question seems to undergird some commentaries. **Wald-Fuhrmann et al.** feel that music is "a contemporary concept of European heritage without direct equivalent in many other cultures and eras," making it fruitless to study from an evolutionary perspective. **Margulis** disagrees, granting the use of studying the evolution of music, but prefers that evolutionary theories hail from researchers in multiple academic fields, so that they "…end up with conclusions that are resilient, and do not easily break down…." **Iyer** is not "invested in the research question of why music might have evolved," suggests that the entire question is irrelevant, and argues that instead scientists should study what "feels like music."

We have two views on these issues. First, in an intellectual community, we believe there's "room enough in the sandbox for everyone."[6] Although reasonable people may disagree over the interpretation of one datum or another, we would not presume to judge other scholars' research priorities, and we prefer to evaluate theories on their supporting evidence, not on the academic affiliations of their authors. Just as the eventual clinical application of basic science is difficult to predict, who is to say what approach is best?

Second, as the saying goes, talk is cheap. Heady questions of the evolution of psychological traits can only be resolved via programmatic empirical research, without which evolutionary theorizing is interesting, but unproductive. Therefore, to those who raised testable questions, we say: let's get to work. Measurement is hard, but not impossible; **Iyer** could improve on his Twitter survey by measuring "what feels like music" in representative samples of humans (as **Levitin** does in a re-analysis of *Natural History of Song* ethnographies). **Tichko et al.** could measure the genetic architecture of musicality via genome-wide association studies, comparative phylogenetic methods, and, as **Kasdan et al.** suggest, in studies of people with neurodevelopmental disorders. Needless to say, many of the commentators have already spent years designing careful experiments that are essential to understanding the evolution of music, as evidenced by the breadth and depth of engagement with scientific and humanistic literatures in both target articles.

As for "why bother?," our view is simple. The goal of the science of music should be to explain music. By testing competing hypotheses of the evolution of musicality, we can hone the reasonable hypothesis space of the functions and mechanisms of the psychology of music, yielding questions, experiments, and entire research programs that are generative, and hopefully, robust.

Financial support. S.A.M. is supported by the Harvard Data Science Initiative and the National Institutes of Health Director's Early Independence Award DP5OD024566.

Conflict of interest. None.

Notes

1. An *exaptation* (see **Trainor**) refers to a feature designed for one adaptive problem but that subsequently came to be used in some other way. We consider such features to be byproducts (Buss, Haselton, Shackelford, Bleske, & Wakefield, 1998). If selection further shapes a trait, the new design features should be considered adaptations.
2. The degree of universality across behavioral contexts of music is not yet known and was not studied in Mehr et al. (2019); the analysis therein that Pinker refers to tests the evidence for or against universality of music in a particular context, but does not compare across contexts.
3. **Cross** suggests that infant attachment could play a role in the evolution of music. Note that the evolution of human parental care is characterized by the interplay of cooperation and conflict (Haig, 2000): parent–offspring conflict and parental attachment can and do coexist.
4. *Recommendation*, a central topic in the field of music informatics, refers to a collection of technologies used by music streaming companies (e.g., Spotify and Pandora) that predict what music a given listener will enjoy. Because this topic is of substantial commercial interest, the tools involved are often proprietary, and direct evidence on the topic can be difficult to obtain.
5. The scarcity of empirical cross-cultural studies of esthetics in music makes it hard to know what such general principles might be. In addition to the previously mentioned study of power laws in music across cultures (Mehr et al., 2019), a notable exception is the finding that Tsimane people do not show a Western-typical preferences for consonance over dissonance in isolated tones (McDermott, Schultz, Undurraga, & Godoy, 2016). The generality of that finding to more explicitly musical esthetics (e.g., in songs) is unknown, however (see, e.g., Bowling, Hoeschele, Gill, & Fitch, 2017).
6. Thanks to Mina Cikara for sharing this quotation, which is attributed to Susan Fiske.

References

Ackermann, R. R., Arnold, M. L., Baiz, M. D., Cahill, J. A., Cortés-Ortiz, L., Evans, B. J., … Zinner, D. (2019). Hybridization in human evolution: Insights from other organisms. *Evolutionary Anthropology: Issues, News, and Reviews*, *28*(4), 189–209.

Atwood, S., Mehr, S., & Schachner, A. (2020). Expectancy effects threaten the inferential validity of synchrony-prosociality research, *PsyArXiv*. https://doi.org/10.31234/osf.io/zjy8u.

Bainbridge, C. M., Bertolo, M., Youngers, J., Atwood, S., Yurdum, L., Simson, J., … Mehr, S. A. (2021). Infants relax in response to unfamiliar foreign lullabies. *Nature Human Behaviour*, *5*(2), 256–64. doi: 10.1038/s41562-020-00963-z.

Bertolo, M., Singh, M., & Mehr, S. A. (2021). Sound-induced motion in chimpanzees does not imply shared ancestry for music or dance. *Proceedings of the National Academy of Sciences*, *118*(2), e2015664118. doi: 10.1073/pnas.2015664118.

Birch, J. (2019). Are kin and group selection rivals or friends? *Current Biology: CB*, *29*(11), R433–R438.

Blackwell, A. D., Urlacher, S. S., Beheim, B., von Rueden, C., Jaeggi, A., Stieglitz, J., … Kaplan, H. (2017). Growth references for Tsimane forager-horticulturalists of the Bolivian Amazon. *American Journal of Physical Anthropology*, *162*(3), 441–461.

Bowling, D. L., Hoeschele, M., Gill, K. Z., & Fitch, W. T. (2017). The nature and nurture of musical consonance. *Music Perception: An Interdisciplinary Journal*, *35*(1), 118–121.

Bryant, G. A. (2013). Animal signals and emotion in music: Coordinating affect across groups. *Frontiers in Psychology*, *4*, 990.

Bryant, G. A. (2020). Evolution, structure, and functions of human laughter. In K. Floyd & R. Weber, eds., *The handbook of communication science and biology*, Routledge, pp. 63–77.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, *53*(5), 533–548.

Corbeil, M., Trehub, S. E., & Peretz, I. (2016). Singing delays the onset of infant distress. *Infancy*, *21*(3), 373–391.

Custodero, L. A., Rebello Britto, P., & Brooks-Gunn, J. (2003). Musical lives: A collective portrait of American parents and their young children. *Journal of Applied Developmental Psychology*, *24*(5), 553–572.

Darwin, C. (1859). *On the origin of species by means of natural selection*, J. Murray.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45), E6233–E6242.

Fouts, H. N., Hewlett, B. S., & Lamb, M. E. (2001). Weaning and the nature of early childhood interactions among Bofi foragers in Central Africa. *Human Nature*, *12*(1), 27–46.

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*(1), 21–51.

Hagen, E. H., & Hammerstein, P. (2009). Did Neanderthals and other early humans sing? Seeking the biological roots of music in the territorial advertisements of primates, lions, hyenas, and wolves. *Musicae Scientiae*, *13*(2 Suppl.), 291–320.

Haig, D. (2000). The Kinship theory of genomic imprinting. *Annual Review of Ecology and Systematics*, *31*(1), 9–32.

Hayden, B. (2014). *The power of feasts*. Cambridge University Press.

Hayden, B., & Villeneuve, S. (2011). A century of feasting studies. *Annual Review of Anthropology*, *40*(1), 433–449.

Hilton, C. B.,, Crowley, L., Yan, R., Martin, A., & Mehr, S. A. (2021). Children infer the behavioral contexts of unfamiliar foreign songs. *PsyArXiv*. https://doi.org/10.31234/osf.io/rz6qn.

Hintze, A., Olson, R. S., Adami, C., & Hertwig, R. (2015). Risk sensitivity as an evolutionary adaptation. *Scientific Reports*, *5*(1), 8242.

Honing, H., ten 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*(1664), 20140088.

Jacobson, K., Murali, V., Newett, E., Whitman, B., & Yon, R. (2016). Music Personalization at Spotify. *Proceedings of the 10th ACM Conference on Recommender Systems*, Boston Massachusetts USA: ACM, pp. 373–373.

Jacoby, N., Margulis, E. H., Clayton, M., … Wald-Fuhrmann, M. (2020). Cross-cultural work in music cognition: Challenges, insights, and recommendations. *Music Perception: An Interdisciplinary Journal*, *37*(3), 185–195.

Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & McDermott, J. H. (2019). Universal and non-universal features of musical pitch perception revealed by singing. *Current Biology*, *29*(19), 3229–3243.

Kokko, H., Jennions, M. D., & Brooks, R. (2006). Unifying and testing models of sexual selection. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 43–66.

Kotler, J., Mehr, S. A., Egner, A., Haig, D., & Krasnow, M. M. (2019). Response to vocal music in Angelman syndrome contrasts with Prader-Willi syndrome. *Evolution and Human Behavior*, *40*(5), 420–426.

Kramer, K. L., & Greaves, R. D. (2007). Changing patterns of infant mortality and maternal fertility among Pumé foragers and horticulturalists. *American Anthropologist*, *109*(4), 713–726.

Levitin, D. J., Chordia, P., & Menon, V. (2012). Musical rhythm spectra from Bach to Joplin obey a 1/*f* power law. *Proceedings of the National Academy of Sciences*, *109*(10), 3716–3720.

Lozoff, B., & Brittenham, G. (1979). Infant care: Cache or carry. *The Journal of Pediatrics*, *95*(3), 478–483.

Manaris, B., Roos, P., Krehbiel, D., Zalonis, T., & Armstrong, J. R. (2012). Zipf's law, power laws, and music aesthetics. In T. Li, M. Ogihara, & G. Tzanetakis (Eds.), *Music data mining* (pp. 169–216). CRC Press.

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*(7613), 547–550.

Mehr, S. A. (2014). Music in the home: New evidence for an intergenerational link. *Journal of Research in Music Education*, *62*(1), 78–88.

Mehr, S. A., Kotler, J., Howard, R. M., Haig, D., & Krasnow, M. M. (2017). Genomic imprinting is implicated in the psychology of music. *Psychological Science*, *28*(10), 1455–1467.

Mehr, S. A., & Krasnow, M. M. (2017). Parent-offspring conflict and the evolution of infant-directed song. *Evolution and Human Behavior*, *38*(5), 674–684.

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*, *366*(6468), 957–970.

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

Mehr, S. A., & Spelke, E. S. (2017). Shared musical knowledge in 11-month-old infants. *Developmental Science*, *21*(2), e12542.

Mendoza, J. K., & Fausey, C. M. (2019). Everyday music in infancy. *PsyArXiv*. doi: 10.31234/osf.io/sqatb.

MRC/Billboard. (2021). *Year-end report*. Retrieved from https://www.musicbusinessworldwide.com/files/2021/01/MRC_Billboard_YEAR_END_2020_US-Final.pdf.

Patel, A. D. (2008). *Music, language, and the brain*, Oxford University Press.

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech.. *Proceedings of the National Academy of Sciences,*, *105*(3), 833–838.

Queller, D. C. (2020). The gene's eye view, the Gouldian knot, Fisherian swords and the causes of selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1797), 20190354.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science (New York, N.Y.)*, *311*(5762), 854–856.

Sankararaman, S., Mallick, S., Patterson, N., & Reich, D. (2016). The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology: CB*, *26*(9), 1241–1247.

Schedl, M., Bauer, C., Reisinger, W., Kowald, D., & Lex, E. (2021). Listener modeling and context-aware music recommendation based on country archetypes. *Frontiers in Artificial Intelligence*, *3*, 508725. doi: 10.3389/frai.2020.508725.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In *Mapping the mind: Domain specificity in cognition and culture*, Cambridge University Press, pp. 39–67.

Swope, K. M. (2009). The beating of drums and clashing of symbols: Music in Ming dynasty military operations. *The Chinese Historical Review*, *16*(2), 147–177.

Trehub, S. E., Hill, D. S., & Kamenetsky, S. B. (1997). Parents' sung performances for infants. *Canadian Journal of Experimental Psychology*, *51*(4), 385–396.

Way, S. F., Garcia-Gathright, J., & Cramer, H. (2020). Local trends in global music streaming. *Proceedings of the International AAAI Conference on Web and Social Media*, *14*, 705–714.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*, Princeton University Press.

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2017). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science*, 2018; 21:e12537. doi:10.1111/desc.12537.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*, Addison-Wesley Press.

Behavioral and Brain Sciences

cambridge.org/bbs

Treatment Index

Treatment Index

TARGET ARTICLES

COMMENTARIES

CAMBRIDGE
UNIVERSITY PRESS

AUTHORS' RESPONSES