

UCLA

UCLA Previously Published Works

Title

Toward a productive evolutionary understanding of music

Permalink

<https://escholarship.org/uc/item/23c05169>

Authors

Mehr, Samuel A

Krasnow, Max M

Bryant, Gregory A

et al.

Publication Date

2021

DOI

10.1017/s0140525x21000030

Peer reviewed



Published in final edited form as:

Behav Brain Sci. ; 44: e122. doi:10.1017/S0140525X21000030.

Toward a productive evolutionary understanding of music

Samuel A Mehr^{a,b,c,*}, Max M Krasnow^{a,*}, Gregory A Bryant^{d,e,*}, Edward H Hagen^{f,*}

^aDepartment of Psychology, Harvard University, Cambridge, MA 02138 USA

^bData Science Initiative, Harvard University, Cambridge, MA 02138 USA

^cSchool of Psychology, Victoria University of Wellington, Wellington 6012, New Zealand

^dDepartment of Communication, University of California Los Angeles, Los Angeles, CA 90095 USA

^eCenter for Behavior, Evolution, and Culture, University of California Los Angeles, Los Angeles, CA 90095 USA

^fDepartment of Anthropology, Washington State University, Vancouver, WA 98686 USA

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).

R1. Adaptations and byproducts: what they are, what they aren't, and why it matters

While 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 whiteish 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.

sam@wjh.harvard.edu <https://mehr.cz>.

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

Conflicts of interest: None

⁶Thanks to Mina Cikara for sharing this quotation, which is attributed to Susan Fiske.

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 due to 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), etc. 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). So we agree with **Killin et al.; Pinker; Trainor; Tichko et al.; Stewart-Williams; Dissanayake; Moser et al.; Lieberman & Billingsley; Leivada; Bowling et al.; Harrison & Seale**, 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, 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

¹An *exaptation* (see **Trainor**) refers to a feature designed for one adaptive problem but that subsequently came to be used in some other way; which we consider such features to be byproducts (Buss et al. 1998). If selection further shapes a trait, the new design features should be considered adaptations.

increased fitness via multiple pathways involving cooperative and agonistic relationships among individuals and groups.

The points raised by **Killing 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 co-evolution 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 produces the design of traits in a contemporary species *in a directional fashion*. As emphasized by **Bowling et al.**, **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 **Trevor & Frühholz**; **Duborg et al.**; **Bowling et al.**; **Cross**; **Patel & von Rueden**; **Hannon et al.**; **Számadó**). Such characteristics play no causal role in an evolutionary theory. **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 vs 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, like 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 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 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 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 contexts², 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

²The 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.

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 elicitation (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 et al. 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 et al. 2015) such as continual mother-infant contact calls. Last, 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 et al. 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

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.

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 (though 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 et al. 2017) and, as **Patel & von Rueden** note, have likely not experienced violent intergroup conflict for centuries.

Last, several commentators raised questions concerning the fit between the adaptive problems posed by territorial signaling and music. **Lieberman and 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 et al. 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 and Billingsley, Pinker, Zentner**) assumed that territorial signals are usually aggressive; since 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.

Lieberman and 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 suspect 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 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 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 et al. 2020). 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 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. While 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 et al. 2019; Mehr et al. 2017; Mehr & Krasnow 2017). But much more can be done in this area, as **Honing, Trehub**, and others have previously suggested (Honing et al. 2015). Indeed, **Kasdan et al.** propose testing musical interventions in genetically informative populations, which we also endorse.

Last, 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 et al. 2019; Sankararaman et al. 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, though we caution that interpreting music-like behavior in non-humans risks anthropomorphism and loose evolutionary logic. For instance, the examples **Hattori** raises of "rhythmic body movements" in non-humans may have nothing to do with music (Bertolo et al. 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 aesthetics

Why are humans so motivated to seek out and to produce music in the first place? **Pinker, Kraus & Hesselmann, Sievers & Wheatley, Trainor, Dubourg et al.**, and others are right to ask how evolutionary theories of music can explain the role of aesthetics in music — its “most blazingly obvious feature” (**Pinker**).

Measuring aesthetic value in music is a substantial challenge. The recommendation engines⁴ 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 et al. 2016). This approach is consistent with experimental work demonstrating the value of social information in musical preferences (Salganik et al. 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 et al. 2021; Way et al. 2020). So, while we agree with the commentators that developing an understanding of aesthetic 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 et al. 2003; Mehr 2014; Mendoza & Fausey 2019; Trehub et al. 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.

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, 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.

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, grunge); such signals of group identity may even be detectable in music for infants (Mehr et al. 2016; Mehr & Spelke 2017; Xiao et al. 2017).

Third, a key characteristic of musical aesthetics 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 et al. 2012; Manaris et al. 2012; Mehr et al. 2019; Zipf 1949). How do these and other general principles of musical aesthetics 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 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 co-evolution. Music and language clearly share several computational principles, many of which are related to auditory processing (see **Trainor**). Like others (Doelling & Poeppel 2015; Jacoby et al. 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).

While 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

⁵The scarcity of empirical cross-cultural studies of aesthetics 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 et al. 2016). The generality of that finding to more explicitly musical aesthetics (e.g., in songs) is unknown, however (see, e.g., Bowling et al. 2017).

organization of the respective systems. For example, hierarchical organization is a principle of language with clear analogues in music, and as **Hilton et al.** describe, across multiple other domains (including metacognition, action planning, and auditory scene analysis). While **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.

Further, 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, though we note that such connections may co-exist with form-fit relationships between adaptive problems and music-specific adaptations.

While 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 is 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 (**Wood; Grahn, Bauer, & Zamm; Pfordresher; Gabriel & Paravati; Hattori**) 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 et al. 2020).

Further, 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 deeply 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, e.g., 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. While 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 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, though 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 utility 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 While 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. So, 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.

Funding statement

S.A.M. is supported by the Harvard Data Science Initiative and the National Institutes of Health Director’s Early Independence Award DP5OD024566.

References

- Ackermann RR, Arnold ML, Baiz MD, ... 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. doi:10.31234/osf.io/zjy8u
- Bainbridge CM, Bertolo M, Youngers J, ... Mehr SA (2020). Infants relax in response to unfamiliar foreign lullabies. *Nature Human Behaviour*. doi:10.1038/s41562-020-00963-z
- Bertolo M, Singh M, & Mehr SA (2021). Sound-induced motion in chimpanzees does not imply shared ancestry for music or dance. *Proceedings of the National Academy of Sciences*, 118(2). doi:10.1073/pnas.2015664118
- Birch J (2019). Are kin and group selection rivals or friends? *Current Biology: CB*, 29(11), R433–R438. [PubMed: 31163152]
- Blackwell AD, Urlacher SS, Beheim B, ... Kaplan H (2017). Growth references for Tsimane forager-horticulturalists of the Bolivian Amazon. *American Journal of Physical Anthropology*, 162(3), 441–461. [PubMed: 28218400]
- Bowling DL, Hoeschele M, Gill KZ, & Fitch WT (2017). The Nature and Nurture of Musical Consonance. *Music Perception: An Interdisciplinary Journal*, 35(1), 118–121.
- Bryant GA (2013). Animal signals and emotion in music: Coordinating affect across groups. *Frontiers in Psychology*, 4, 990. [PubMed: 24427146]
- Bryant GA (2020). Evolution, Structure, and Functions of Human Laughter. In Floyd K & Weber R, eds., *The Handbook of Communication Science and Biology*, Routledge, pp. 63–77.
- Buss DM, Haselton MG, Shackelford TK, Bleske AL, & Wakefield JC (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, 53(5), 533–548. [PubMed: 9612136]
- Corbeil M, Trehub SE, & Peretz I (2016). Singing delays the onset of infant distress. *Infancy*, 21(3), 373–391.
- Custodero LA, 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*, London: J. Murray.
- Doelling KB, & Poeppel D (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, 112(45), E6233–E6242.
- Fouts HN, Hewlett BS, & Lamb ME (2001). Weaning and the nature of early childhood interactions among Bofi foragers in Central Africa. *Human Nature*, 12(1), 27–46. [PubMed: 26191818]
- Hagen EH, & Bryant GA (2003). Music and dance as a coalition signaling system. *Human Nature*, 14(1), 21–51. [PubMed: 26189987]
- Hagen EH, & 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.
- Hintze A, Olson RS, Adami C, & Hertwig R (2015). Risk sensitivity as an evolutionary adaptation. *Scientific Reports*, 5(1), 8242. [PubMed: 25649757]
- Honing H, Cate C. ten, Peretz I, & Trehub (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. In *Proceedings of the 10th ACM Conference on Recommender Systems*, Boston Massachusetts USA: ACM, pp. 373–373.

- Jacoby N, Margulis EH, 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 EA, McPherson MJ, Valdés J, Ossandón T, & McDermott JH (2019). Universal and non-universal features of musical pitch perception revealed by singing. *Current Biology*, 29(19), 3229–3243. [PubMed: 31543451]
- Kokko H, Jennions MD, & Brooks R (2006). Unifying and testing models of sexual selection. *Annual Review of Ecology, Evolution, and Systematics*, 37, 43–66.
- Kotler J, Mehr SA, Egnér A, Haig D, & Krasnow MM (2019). Response to vocal music in Angelman syndrome contrasts with Prader-Willi syndrome. *Evolution and Human Behavior*, 40(5), 420–426. [PubMed: 32655274]
- Kramer KL, & Greaves RD (2007). Changing Patterns of Infant Mortality and Maternal Fertility among Pumé Foragers and Horticulturalists. *American Anthropologist*, 109(4), 713–726.
- Levitin DJ, 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. [PubMed: 381622]
- Manaris B, Roos P, Krehbiel D, Zalonis T, & Armstrong JR (2012). Zipf's law, power laws, and music aesthetics. In *Music data mining*. Boca Raton: CRC Press.
- McDermott JH, Schultz AF, Undurraga EA, & Godoy RA (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, 535(7613), 547–550. [PubMed: 27409816]
- Mehr SA (2014). Music in the home: New evidence for an intergenerational link. *Journal of Research in Music Education*, 62(1), 78–88.
- Mehr SA, Kotler J, Howard RM, Haig D, & Krasnow MM (2017). Genomic imprinting is implicated in the psychology of music. *Psychological Science*, 28(10), 1455–1467. [PubMed: 28857689]
- Mehr SA, & Krasnow MM (2017). Parent-offspring conflict and the evolution of infant-directed song. *Evolution and Human Behavior*, 38(5), 674–684.
- Mehr SA, Singh M, Knox D, ... Glowacki L (2019). Universality and diversity in human song. *Science*, 366(6468), 957–970. [PubMed: 31753988]
- Mehr SA, Song LA, & Spelke ES (2016). For 5-month-old infants, melodies are social. *Psychological Science*, 27(4), 486–501. [PubMed: 26917211]
- Mehr SA, & Spelke ES (2017). Shared musical knowledge in 11-month-old infants. *Developmental Science*, 21(2).
- Mendoza JK, & Fausey CM (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 AD (2008). *Music, language, and the brain*. Oxford; New York: Oxford University Press.
- Queller DC (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 MJ, Dodds PS, & Watts DJ (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, 311(5762), 854–856. [PubMed: 16469928]
- 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. [PubMed: 27032491]
- 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. 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*, New York, NY, US: Cambridge University Press, pp. 39–67.

- Swope KM (2009). The Beating of Drums and Clashing of Symbols: Music in Ming Dynasty Military Operations. *The Chinese Historical Review*, 16(2), 147–177.
- Trehub SE, Hill DS, & Kamenetsky SB (1997). Parents' sung performances for infants. *Canadian Journal of Experimental Psychology*, 51(4), 385–396. [PubMed: 9606951]
- Way SF, 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 GC (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*, Princeton, N.J.: Princeton University Press.
- Xiao NG, Quinn PC, 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*, n/a–n/a.
- Zipf GK (1949). *Human behavior and the principle of least effort: an introd. to human ecology*, Cambridge, Mass.: Addison-Wesley Pr.