

THOUGHTS ON AN EMPIRICAL APPROACH TO THE EVOLUTIONARY ORIGINS OF MUSIC

JOSH McDERMOTT
*Department of Brain and Cognitive Sciences,
Massachusetts Institute of Technology*

MARC D. HAUSER
*Departments of Psychology, Organismic and
Evolutionary Biology, and Human Evolutionary Biology,
Harvard University*

ONE YEAR AGO WE PUBLISHED a review article in *Music Perception* surveying and summarizing the empirical literature relevant to the evolutionary origins of music. Seven pieces commenting on our review have appeared in the present issue. In this reply we discuss some of the arguments made by the commentators and clarify some of the points made in our original review.

Key words: music, evolution, perception, cognition, comparative, monkey

THESE SEVEN COMMENTARIES, KINDLY SOLICITED by the editors of *Music Perception*, aptly reflect the enthusiasm, interest, and debate that have marked the emergence of the evolution of music as a field of study. It was this emergence, and our own efforts to contribute to the field, that prompted us to summarize and critique the relevant empirical literature in our review one year ago (McDermott & Hauser, 2005). It is clear that the issues engender a wide-ranging set of opinions, and healthy debate; this is no surprise given the natural inclination to wonder about the origins of a phenomenon that is so important to so many people. Here we reply to the commentaries and clarify some of the main points made in our review.

Innateness and Adaptation

We began our review by stressing the importance of innateness, uniqueness, and specificity to music. We argued that candidates for adaptations selected to enable

musical behavior would be found in aspects of music that are innate, uniquely human, and specific to music—innate because adaptations must involve changes to the genetic code, uniquely human because nonhuman animals do not make music, and specific to music because otherwise some other function could have fueled the evolutionary process. Empirical results can push the debate forward, we argued, by providing evidence for or against these three properties of a music-related trait.

Several of the commentaries (Livingstone & Thompson, 2006; Trainor, 2006) contend that we are overly optimistic, particularly with regard to the tasks of determining innateness and of distinguishing adaptations from exaptations (a distinction which we largely did not address in our review). With their caveats we by and large agree. Complex cognitive behaviors presumably have complex evolutionary histories, which will not be straightforward to uncover. We never claimed the research project would be easy, or that it will be completed to the level of detail present in traditional accounts of some biological adaptations. However, we continue to maintain that appropriate experiments have the potential to reveal some clues, and that our guidelines provide a useful framework for these experiments. In reply to the commentators we would like to address two points in general.

First, there is no question that cognitive and neural processes such as those associated with music develop through an interplay between genes and experience. Trainor argues that this makes it difficult to assess whether a trait is innate, noting that even basic auditory perceptual processes seem not to develop normally in the absence of input. This means that most cognitive or neural traits of any interest will be determined by both genetic and environmental factors. How then, is it meaningful to speak of innateness? In our view the critical issue for music and other high-level cognitive phenomena involves the sort of input needed. Clearly, music cannot happen without normal auditory function, which depends on auditory experience to some extent. However, many of the cognitive and perceptual mechanisms involved in music might be able to develop

without prior experience with music *per se*. This would indicate that the trait, although dependent on auditory processes that may require auditory input to develop, is not itself learned from exposure to music. To be clear, this obviously cannot be true of all facets of musical behavior, as music comprehension and production in their full complexity are uncontroversially shaped by exposure to music and culture. But many elementary musical traits might well be present independent of musical input. The abilities observed in young infants, such as meter perception (Hannon & Trehub, 2005), melody recognition (Chang & Trehub, 1977), and preferences for consonance (Trainor, Tsang, & Cheung, 2002) are all plausible candidates in this regard.

Experiments thus have the potential to reveal the extent to which aspects of our capacity to perceive and appreciate music depend on specific musical input as opposed to more general auditory input. Trainor reiterates the commonly cited problems associated with interpreting experiments on infants' musical abilities as evidence for innateness, but this is a practical issue rather than a theoretical one. The musical exposure of human infants could, in principle, be manipulated to test the extent to which effects measured at very young ages are dependent on experience. An effect of deprivation or altered input on some musical trait is consistent with a variety of possible gene/environment interactions, but if deprivation has no effect, it suggests that the trait is not learned from musical input. Given the difficulty of eliminating or systematically altering human musical input, at present we really have little idea of what the effects of this might be, but it is not out of the question to test it. For instance, normally hearing infants born to deaf parents likely have drastically altered musical exposure, and might provide one way to test the effects of music deprivation. Experiments can also in principle test how much input is needed for musical traits to emerge in cases where they are input-dependent. As in language, the development of music cognition processes with minimal input might indicate a biological component (Chomsky, 1986; Crain & Pietroski, 2001).

Moreover, while it is true that apparent species differences in consonance perception (McDermott & Hauser, 2004) and perhaps melody perception (D'Amato, 1988) could be due to differences in musical experience between animal and human subjects, animals' musical experience can in principle be manipulated to be comparable in some ways to that of human infants, potentially testing whether any differences observed are due in part to genetic differences. We are not claiming that these experiments will be easy to carry out, but they are certainly not impossible. Indeed, we have just completed some experiments

manipulating the musical environment of a colony of monkeys, to test its effect on their musical preferences (McDermott and Hauser, forthcoming).

As another example, Trainor discusses the neural circuits that learn and represent the tonal hierarchy. She argues that it is unlikely that these circuits are completely genetically specified, but rather again likely result from some interaction between genetic constraints and auditory input. We would again contend, however, that this interaction does not make it impossible to learn anything about the evolutionary origins of such circuitry, or to discuss the extent to which it has an innate basis. As before, the critical issue is whether the mechanisms that represent the tonal hierarchy depend on exposure to music. This is again an empirical issue; one can in principle test whether prior musical experience, and how much of it, is needed to represent tonal pitch hierarchies. This could be bolstered by testing for representations of tonal hierarchy in listeners with dramatically different cultural backgrounds, an approach that has already had some success (Kessler, Hansen, & Shepard, 1984). Second, if animals can extract tonal dependencies and learn to recognize violations as they can other statistical regularities (Newport, Hauser, Spaepen, & Aslin, 2004), a general-purpose statistical learning mechanism would be implicated, rather than something that could have evolved to enable music perception. Third, brain imaging and neuropsychology can test whether the mechanisms involved in extracting and storing the tonal hierarchy are music-specific as opposed to shared with language, mathematics, or other domains. If such mechanisms are not music-specific, they are less likely to be adaptations for music, as many other domains have rather clear and compelling putative selective pressures. Each of these steps is non-trivial, and might in practice be quite difficult to carry out, but even a few successful experiments would reveal a great deal about the origins of this particular component of music. Accumulating similar experimental results about other components as well will place strong constraints on the evolutionary history of music. Thus while it is true that culture and biology are entangled in music and other complex human behaviors, there is nonetheless reason to think that their relative contributions might be teased apart in some cases.

A second concern voiced by several of the commentators involves the difficulty of distinguishing between adaptations and exaptations, and the general notion that such labels may be oversimplified and ill-suited as descriptors of music. With this we are also sympathetic. Music is a complex, multifaceted phenomenon, consisting of many different behaviors each of which involves

many component processes. Although this complexity probably means that the question “is music an adaptation?” is not itself well posed, it does not necessarily preclude arriving at relatively clear answers for some of the component processes, particularly if these seem to be unique to music. However, our commentators fear that even for manageable component processes of music, it may be difficult to distinguish whether they are adaptations, exaptations, or some combination thereof. We fully agree that determining adaptive function is particularly difficult for cognitive processes, and may prove all but impossible for some aspects of music. The three criteria that we have emphasized—innateness, uniqueness to humans, and specificity to music—serve to identify plausible *candidates* for adaptations. We largely avoided discussion of how to proceed further, precisely because it is not obvious how one should do so. Determining specificity to music can be of some help in this context, as traits that are specific to music and that lack functions in other domains are more likely to be adaptations for music (in that there are fewer obvious alternative accounts of how they might have evolved). But determining this specificity is admittedly not always straightforward, and even if it is, there is no assurance that the trait in question did not have some other function at the time of its evolution that has since become irrelevant and disappeared. Thus determining which if any function of a trait was subject to selection is indeed difficult for cognitive domains, and we do not claim to have a magical solution. However, we would emphasize that whether a musical trait is an adaptation or not is by no means the only question of interest regarding its evolutionary history. Even if the adaptation question lacks a simple or attainable answer for a given musical trait, we may still be able to gain valuable insights into its origins by investigating whether the trait is innate and unique to humans. Other methods (looking at neural overlap with other domains, manipulating musical exposure to probe importance of musical input) can also provide useful information even if the adaptation question is not answered definitively. We also must note that the question of whether a trait is an adaptation for music can be rendered moot if empirical evidence shows that the trait is not heritable (via developmental, cross-cultural or other methods), or that it is shared with nonhuman animals, as they do not make music. The qualifications and caveats emphasized by our commentators are thus valuable reminders of the difficulties inherent in probing the evolution of human cognition and behavior but are not grounds for discouragement. A complete story about music’s origins is by no means guaranteed to fully emerge, but we might nonetheless learn a great deal.

Perception, Production, and Animal Song

Several of the commentators also discussed our focus (Cross, 2006; Merker, 2006). They correctly observed that our review dealt mainly with perception, and within that emphasized pitch over rhythm. As we noted in our review, the emphasis on pitch was, in our view, in rough proportion to what is available in the literature. While there is an ample literature on rhythm perception, there is little work with the methodologies most relevant to the evolution of our rhythmic abilities (although the developmental literature in this area is picking up (Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005)). Indeed, Patel (2006) proposes some interesting experiments on rhythm perception and production in nonhuman primates that could help fill this gap. Our own research indicates that marmosets do not prefer metrical to nonmetrical music, perhaps suggesting that meter is not aesthetically significant to them to the extent that it is for us (McDermott & Hauser, unpublished results), but it remains to be seen whether they or other animals could make metrical discriminations or otherwise perceive aspects of meter. Rhythm thus remains an important area for future research. Incorporation of additional insights from the music of other cultures could also help to shift the research away from being so focused on Western pitch structures, which many would argue are not the strongest candidates for having an innate basis.

Our focus on perception was also reflective of the sense that our productive abilities (e.g., vocal learning), while impressively distinct from those of other primates, seem to be shared with other domains, and thus less likely to have evolved to enable music. Merker clearly begs to differ, suggesting that what we describe as a lack of evidence for music-specific adaptations is unsurprising given that auditory perceptual mechanisms are found in humans and nonhuman animals alike. He is also fond, as are we, of accounts that explain aspects of music as optimal features of a generative system working within certain constraints (e.g., Balzano, 1982); such accounts appear to obviate the need for evolutionary explanations. He suggests that productive and motivational traits are more likely to be adaptations specifically for music, as these traits are not shared with nonhuman animals.

We think it curious that production would be considered a more likely locus for a music-specific adaptation than perception, given that there are at least a few examples of perceptual traits that seem to be specific to music. Consonance, tonality, diatonicity, and meter, to name a few, all lack relevance for other domains apart from music. In contrast, production

does not obviously have music-specific components, apart, perhaps, from the ability to drum rhythmically (Patel, 2006). The existence of multiple functions for productive mechanisms does not rule out the possibility that they evolved to enable music, but they do make it less likely. Moreover, there is ample evidence from the comparative literature to suspect that perhaps there are aspects of music perception that are uniquely human. Most obviously, animals do not seem to readily extract relative pitch, for instance, as discussed extensively in our review. The oft-cited study by Wright and colleagues (Wright, Rivera, Hulse, Shyan, & Neiworth, 2000) is intriguing, but is but one report, one whose results differ from what one would expect given a number of other studies (Brosch, Selezneva, Bucks, & Scheich, 2004; D'Amato, 1988; Hulse & Cynx, 1985), as well as anecdotal evidence from several other animal behavior labs. And as mentioned earlier, our own research has raised the possibility that the preference for consonance may be uniquely human (McDermott & Hauser, 2004). It thus remains quite possible that aspects of the way humans perceive music are uniquely human, and thus candidates for music-specific adaptations.

We certainly would not deny that many other perceptual components of musical behavior likely derive from preexisting, general-purpose features of the auditory system; this idea was a main theme of our review. However, we regard this as an empirical matter rather than a foregone conclusion. It must also be noted that we remain in the early stages of investigating these matters, so the fact that we have little evidence for a music-specific adaptation should not be viewed as disappointing or in any way conclusive. We think it is as much a reflection of how little work has been done in the area as anything else.

Similarly, observing that aspects of music can be viewed as optimal features of a generative system leaves open the question of whether this is the result of cultural coevolution, natural selection, or some combination thereof. Parsimony might favor learning and culture but ultimately this is an empirical question. By the same logic, we should not exclude features of production from consideration as adaptations for music. From our perspective, however, it is unclear where to begin, as so much of production seems mediated by general-purpose machinery.

Several of the commentators contend that insights into the evolution of our productive abilities can be gained from the consideration of animal song as a metaphor for human music (Fitch, 2006; Merker, 2006). We have maintained that for the purpose of understanding the origins of human music, investigations of animal song are a less

promising avenue for research than are tests of how animals perceive and respond to human musical stimuli (Hauser & McDermott, 2003; McDermott & Hauser, 2005). This stance was motivated by three main observations. First, animal song is clearly not homologous to any aspect of human music, given that our closest animal relatives, in contrast to birds and whales, do not sing. The case for homologies in music perception on the other hand, is still open, as these might be present despite the lack of musical behavior in other primates. Second, animal song as observed in birds and whales differs in many ways from human music. For instance, whereas animals almost exclusively sing in the service of territorial defense and mate attraction, humans often do not, and instead employ music in a variety of other contexts, many of which lack obvious functions. In our view, functional differences such as these, as well as differences in structure, complexity, and variety, weaken any analogies between animal song and human music. Third, animal song has in many instances been quite extensively studied, whereas there are only a few reports of experiments testing how nonhuman animals perceive musical stimuli. We argued on the basis of these observations that tests for perceptual homologies might be more likely to yield insight.

As the commentaries by Merker and Fitch reveal, the relevance of animal song for human music remains a controversial issue, and that is fine with us. We are content to await any insights that might come from comparing animal and human song, while pursuing a comparative research program focusing on testing animals for human traits. Indeed, if animal song is similar to human music in important respects, one might expect to find aspects of music *perception* in common between humans and singing animals, an interesting direction for future research. It is worth noting, however, that a substantial literature in songbirds has yet to reveal much in the way of perceptual homoplasies between songbirds and humans. To the contrary, songbirds seem to have great difficulty extracting relative pitch, and therefore do not readily recognize transpositions of melodies, as alluded to earlier (Hulse & Cynx, 1985).

One point on which we agree with Merker is that motivation for music is one of its most striking features, and seems a strong candidate for an adaptation. However, we deem it possible that at least some of this motivation is embedded in music perception; we suspect we are motivated to engage in music in part because we like it. We discussed aesthetic responses at length in our review, and have in fact attempted to study their origins empirically by testing nonhuman animals for rudimentary music preferences (McDermott & Hauser, 2004; McDermott & Hauser, 2006). To return to our original

framework, the pleasurable response to music would be a candidate for an adaptation if it were innate, unique to humans, and specific to music. While there are no doubt many factors contributing to our enjoyment of music, one contributing factor may be the preference for consonance found in most adults. It is worth noting that this preference may satisfy all three preconditions—it appears to be potentially innate (being present even in very young infants (Trainor et al., 2002)), unique to humans (McDermott & Hauser, 2004), and specific to music. As such it is at least conceivable that it could have evolved to motivate us to engage in music. Similarly, the link between music and emotion remains another strong candidate for an adaptation. Specific cues to emotion may not be innate (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001), but the tendency to associate emotion with music probably is, as it is found cross-culturally. The extent to which we perceive emotion in music also does not appear to be mirrored in other domains; music is profoundly and routinely emotional to an extent unparalleled even by other forms of art. Given that the link between emotion and music is plausibly unique to humans, this is a candidate for an adaptation, and speaks to the need to look at the interfaces between different systems as possible centers for species-unique adaptations.

Definitions

Another theme to emerge from the commentary pieces involves the definition of music and the heterogeneity of musical behavior. We explicitly put off the issue of defining music in our review, mainly because we deemed it nonessential for our purposes, but also because we think music in its widely varied forms is probably ill-suited to a concise definition. The diversity of music impedes the formation of a comprehensive definition, but it need not impede our research. We would rather dive in and start studying phenomena of interest rather than worry about how to fit a definition of music around them. Cross maintains that the lack of a definition “limits the power” of our arguments; we remain unconvinced. From our

perspective, all that matters is that readers know what we are talking about, and we think there is little doubt that they do, at least as evidenced by the lack of confusion among our commentators.

New Empirical Work

The three commentary pieces we have not discussed in detail (Dean & Bailes, 2006; Livingstone & Thompson, 2006; Patel, 2006) all propose new lines of empirical work. We will let these speak for themselves, and merely state our belief that given the volume of theoretical speculation on our chosen topic, such new empirical work is particularly critical and welcome.

Conclusion

The evolution of music remains a topic of great interest and passionate debate, and these commentaries along with the original review articles will hopefully generate more of it. We will conclude by reiterating our central theme, that conducting experiments to test whether aspects of musical behavior are innate, uniquely human, and specific to music will place useful constraints on our understanding of the evolutionary history of music. These experiments are often difficult to design and carry out, and their interpretation will doubtlessly fuel more debate, but we are hopeful that they will help shape the way we think about music’s origins in the years to come.

We thank the other authors for their provocative comments, and the editors of *Music Perception* for inviting our response.

Author Note

Address correspondence to: Josh McDermott, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 46-4115, 77 Massachusetts Avenue, Cambridge MA 02139. E-MAIL jhm@mit.edu

References

- BALZANO, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind and brain: The neuropsychology of music* (pp. 321–351). New York: Plenum.
- BROSCH, M., SELEZNEVA, E., BUCKS, C., & SCHEICH, H. (2004). Macaque monkeys discriminate pitch relationships. *Cognition*, 91, 259–272.
- CHANG, H. W., & TREHUB, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, 24, 324–331.
- CHOMSKY, N. (1986). *Knowledge of language: Its nature, origins, and use*. New York: Praeger.
- CRAIN, S., & PIETROSKI, P. (2001). Nature, nurture and Universal Grammar. *Linguistics and Philosophy*, 24, 139–185.

- CROSS, I. (2006). The origins of music: Some stipulations on theory. *Music Perception*, 24, 79–82.
- DALLA BELLA, S., PERETZ, I., ROUSSEAU, L., & GOSSELIN, N. (2001). A developmental study of the affective value of tempo and mode. *Cognition*, 80, B1–B10.
- D'AMATO, M. R. (1988). A search for tonal pattern perception in cebus monkeys: Why monkeys can't hum a tune. *Music Perception*, 5, 453–480.
- DEAN, R. T., & BAILES, F. (2006). Toward a sociobiology of music. *Music Perception*, 24, 83–84.
- FITCH, W. T. (2006). On the biology and evolution of music. *Music Perception*, 24, 85–88.
- HANNON, E. E., & TREHUB, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, 16, 48–55.
- HAUSER, M. D., & MCDERMOTT, J. (2003). The evolution of the music faculty: A comparative perspective. *Nature Neuroscience*, 6, 663–668.
- HULSE, S. H., & CYNX, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus*, *Molothrus*, and *Sturnus*). *Journal of Comparative Psychology*, 99, 176–196.
- KESSLER, E. J., HANSEN, C., & SHEPARD, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, 2, 131–165.
- LIVINGSTONE, S. R., & THOMPSON, W. F. (2006). Multimodal affective interaction: A comment on musical origins. *Music Perception*, 24, 89–94.
- MCDERMOTT, J., & HAUSER, M. D. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition*, 94, B11–B21.
- MCDERMOTT, J., & HAUSER, M. D. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception*, 23, 29–59.
- MCDERMOTT, J., & HAUSER, M. D. (in press). Nonhuman primates prefer slow tempos but dislike music overall. *Cognition*.
- MERKER, B. (2006). The uneven interface between culture and biology in human music. *Music Perception*, 24, 95–98.
- NEWPORT, E., HAUSER, M. D., SPAEPEN, G., & ASLIN, R. N. (2004). Learning at a distance. II. Statistical learning of non-adjacent dependencies in a nonhuman primate. *Cognitive Psychology*, 49, 85–117.
- PATEL, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, 24, 99–104.
- PHILLIPS-SILVER, J., & TRAINOR, L. J. (2005). Feeling the beat in music: Movement influences rhythm perception in infants. *Science*, 308, 1430.
- TRAINOR, L. J. (2006). Innateness, learning, and the difficulty of determining whether music is an evolutionary adaptation. *Music Perception*, 24, 105–110.
- TRAINOR, L. J., TSANG, C. D., & CHEUNG, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception*, 20, 187–194.
- 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, 291–307.