

INNATENESS, LEARNING, AND THE DIFFICULTY OF DETERMINING WHETHER MUSIC IS AN EVOLUTIONARY ADAPTATION

A COMMENTARY ON JUSTUS & HUTSLER (2005) AND McDERMOTT & HAUSER (2005)

LAUREL J. TRAINOR
McMaster University

McDERMOTT AND HAUSER (2005) and Justus and Hutsler (2005) argue that in order for music to be an evolutionary adaptation, and not an exaptation, music must be constrained by innate factors that are specific to music and that evolved because music conferred survival advantages. I argue that the dichotomy between adaptation and exaptation is not very clear for higher cognitive functions such as music and language, because genes set up general neural architectures and learning mechanisms, rather than specifying the details of what they will represent, and that such higher cognitive functions are therefore dependent on interactions between genes and experience. Furthermore, because higher cognitive functions depend on a common set of mechanisms such as sensory encoding strategies, working memory capacity, working memory processes, long-term memory encoding and retrieval, and attentional focusing, it is difficult to make arguments about innate specificity. Thus, the question of whether music is an evolutionary adaptation appears to depend on whether or not music conferred survival advantages, a question that is difficult to answer.

Key words: music, evolution, adaptation, pitch, development

McDERMOTT AND HAUSER (2005) and Justus and Hutsler (2005) provide thoughtful papers on the origins of musical behavior that go beyond most previous approaches in attempting to define rigorously the necessary criteria by which to judge whether music arose as an evolutionary adaptation (i.e., music itself conferred survival advantages that caused genetic changes that enhanced musical abilities) or whether it arose as an exaptation (i.e., music was

enabled through traits adapted for other purposes). Both papers agree that two critical conditions are necessary for music to qualify as an evolutionary adaptation. First, music must be *constrained by innate factors* and not be simply a cultural invention. Second, these innate constraints must be *specific to music* and not involve only domain-general mechanisms that could have evolved through other survival pressures, and these constraints *must have evolved because music conferred survival advantages*. Both papers are optimistic that we will be able to answer these questions by collecting more data in the areas of (1) neural encoding and modularity, (2) early infant abilities and cognitive development in humans, (3) cross-cultural comparisons of music across societies past and present, (4) studies of animal capabilities, and (5) computational approaches. I will argue that, especially in the case of higher-level cognitive domains such as language and music, the complex interplay between genes and experience in development and in evolution renders it difficult to determine whether or not the critical conditions for an adaptation are met. Indeed, for complex cognitive abilities, the distinction between exaptation and adaptation may not be completely clear.

In traditional evolutionary theory there is a focus on establishing the adaptive value of traits, but less on determining the series of steps involved in the evolution of complex cognitive abilities such as music or language. As pointed out by Justus and Hutsler (2005), the evolution of all higher cognitive functions must involve exaptation, because more complex neural circuits must have evolved from simpler neural systems. Furthermore, all higher cognitive functions, such as music, language, and mathematical processing depend critically on a general set of processing mechanisms, including sensory encoding strategies, working memory capacity, working memory processes, long-term memory encoding and retrieval, attentional focusing, and so on. This is seen clearly in recent fMRI studies, in which (1) cognitive tasks engage networks of brain areas and (2) many areas

are not specifically activated only by tasks in one domain. In the music domain, Koelsch and colleagues have demonstrated that processing musical syntax engages a very similar network of brain regions to that activated by language (Koelsch et al., 2002). Thus, the genetic processes involved in setting up and controlling these brain regions must affect both music and language. The situation is complicated further by the interaction between genes and culture. For example, although reading and writing, which developed in the recent history of our species, and linguistic and musical behaviour, which developed some time ago, may all have originated as cultural inventions that relied on exaptation (e.g., using visual and motor mechanisms that evolved for other purposes in reading and writing; using auditory abilities that evolved for sound-source separation and auditory object identification in the case of music and language), their importance for fitness in linguistic, musical societies might well cause certain genes to be preferentially passed on. In this way, the boundary between exaptation and adaptation becomes fuzzy.

In making evolutionary arguments about complex cognitive functions, it is essential to examine the relation between genes and learning in the development of the brain. The more complex and adaptable the behavioral possibilities of an organism, the more dependent brain development will be on learning, and the less dependent it will be on genetic programs that specify details of neural circuits and what information is represented in what neural tissue, if for no other reason than that there is a limit on the number of genes that will fit into a cell nucleus (Elman et al., 1996). Thus, complex behaviors emerge through experience-dependent wiring of neural circuits in interaction with architectural genetic constraints. For example, while musical pitch processing relies on tonotopic (frequency) maps in sub-cortical areas and primary auditory cortex, it is not necessary to specify these maps in detail in genetic programs. The experience of hearing sounds with pitch in combination with the physical structure of the cochlea in the inner ear (whereby sounds of different frequencies preferentially enervate different hair cells in an orderly manner along the length of the basilar membrane) automatically leads to the formation of tonotopic maps at higher brain areas through basic processes of experientially driven neural synapse formation. In fact, animals exposed to white noise with no spectral organization do not develop normal tonotopic maps (e.g., Zhang, Bao, & Merzenich, 2002), demonstrating that the maps are not genetically determined. Thus, because it can be assumed that people will hear sounds with

pitch in normal environments, it is not necessary for genes to specify the fine neural circuitry of tonotopic maps. Indeed, such maps are plastic to some extent even in adults, and can be modified by experience, or lack of experience, with particular frequencies (Bosnyak, Eaton, & Roberts, 2004), an ability that allows us to adjust to different environments.

What, then, do genes contribute to the cortical processing of musical pitch? To begin to answer this question, it is instructive to consider what genes do in general. It is highly unlikely that genes encode specific entities or concepts such as the tonal hierarchy in music or the past tense in language. Rather, genes likely specify architectural constraints that allow or prohibit certain representations or types of complex processing to be learned. For example, genes are involved in the formation of brain regions and the initial number and characteristics of neurons in each region. They play a large role in controlling the chemical processes that regulate synaptic formation and how these processes change across development. In the case of pitch processing, then, the genes must specify an architecture that is able to learn to represent pitch with particular precision and in a hierarchical manner. Indeed, Zatorre (2003) has described how the architecture of the left and right auditory cortices differ in terms of the spacing of neurons and the number of long-range connections, such that the left has an advantage for fine temporal processing and the right an advantage for fine spectral processing. Thus, because of the intimate interrelation between gene expression and learning, it is not really possible to study these processes separately.

Given that for complex cognitive functions, genes contribute general architectural constraints that limit learning mechanisms and the types of representations that are possible in the neural tissue, how can we evaluate the two conditions outlined above that have been proposed for determining whether or not music is an evolutionary adaptation? Let us first consider the condition that music must be *constrained by innate factors*. While the approaches of McDermott and Hauser (2005) and Justus and Hutsler (2005) both pit genetic constraints against learning, according to the perspective of the present commentary, learning and innate constraints work together. The early development of certain musical abilities has often been used as evidence that music is innate (Justus & Hutsler, 2005; McDermott & Hauser, 2005). For example, 2-month-olds prefer consonance over dissonance (Trainor, Tsang, & Cheung, 2002), 6-month-olds process relative pitch (Plantinga & Trainor, 2005; Trehub, Bull, & Thorpe, 1984) and musical

meter (Hannon & Trehub, 2005), and 7-month-olds show multisensory interactions between auditory rhythm and movement (Phillips-Silver & Trainor, 2005). On the other hand, the later development of knowledge about key membership by 5 years of age (Trainor & Trehub, 1992, 1994), implied harmony by 7 years (Trainor & Trehub, 1994), and the full tonal hierarchy even later (Costa-Giomi, 2003; Cuddy & Badertsher, 1987; Krumhansl & Keil, 1982) has been interpreted as evidence that these aspects of music depend primarily on learning (Justus & Hutsler, 2005). However, late developments, such as the hormonal changes seen at puberty, can also be powerfully driven by genetic factors. In the music domain, the late emergence of knowledge about key membership and implied harmony might critically depend on the size of working memory or the maturation of frontal brain areas. On the other side, even at a few months of age, infants have already had a lot of exposure to music. For example, infants might prefer consonance because of their experience with complex tones whose lower harmonics stand in consonant relations (Terhardt, 1974). Thus, it is difficult to draw definitive conclusions as to whether aspects of music are innate or learned on the basis of such developmental data.

There is, however, more compelling evidence for innate constraints. Neuroimaging work clearly shows that music is encoded in specific brain networks, and that brain damage can result in severe impairments of specific aspects of musical behaviour (Peretz, 2003). Thus, although there may not be genes that determine musical representations, there are at least genetic constraints that allow music to be learned. Cross-cultural evidence shows commonalities across musical systems, both past and present, in, for example, octave and small-integer ratio primacy, the importance of pitch contours, optimal tempos, and hierarchical rhythmic structures. According to the perspective of the present commentary, this suggests that there are general genetic constraints that limit the types of musical structures that are easily learned. Furthermore, computational models provide an existence proof that musical structure is learnable. The final type of evidence often cited for the existence of genetic constraints concerns differences between humans and other animals. For example, despite the fact that monkeys can categorize consonant and dissonant intervals (Izumi, 2000), they show no preference for consonance over dissonance (McDermott & Hauser, 2004). Given that these animals did not develop in a musical environment, these studies do not definitively indicate whether humans and other animals are differentially genetically constrained with respect to

music; however, the fact that these species have not spontaneously developed music, in contrast to all known human societies, suggests fundamental genetic differences.

The second proposed critical condition for the determination of whether music is an adaptation is that *genetics must dictate constraints that are specific to music and that led to selection pressures*. This condition is difficult to evaluate for complex cognitive abilities such as musical behaviour. For example, McDermott and Hauser (2005) argue that because infants prefer lullabies from different cultures over adult songs, infants have an innate preference for a particular musical style. However, the preference for infant-directed songs (Trainor, 1996) is likely due to their simple structure (Unyk, Trehub, Trainor, & Schellenberg, 1992) and the emotional manner in which they are sung (Trainor, Clark, Huntley, & Adams, 1997), neither of which is a feature that is specific to music. The preference for simple structure likely arises from general cognitive limitations of infants, and the preference for positive emotion is found for speech as well as music (Trainor, Austin, & Desjardins, 2000). Justus and Hutsler (2005) conclude that the evidence for music-specific constraints in general is quite limited. However, while the processing and learning mechanisms may well overlap across different cognitive domains, each domain has unique content. Aspects of music content that appear to be unique to music include a preference for consonance, scales based on octave equivalence that contain unequal intervals, and the tonal pitch hierarchy. According to the argument set up by Justus and Hutsler (2005), the critical question is to determine whether these unique aspects of music are innate or whether they are learned. If they are innate, then music is likely an adaptation and if they are learned, it is an exaptation. The problem with this dichotomy is that, although it is unlikely that there is a genetic program to set up, for example, a part of the brain to perform operations to extract the tonal hierarchy, neural circuits are set up that can learn to do this given the appropriate input, but can also learn to perform other computations as well. Thus, these unique aspects of music depend on both innate constraints and learning, and thereby seem to fall in between an adaptation and an exaptation. This appears to leave the determination of whether music is an adaptation to the direct examination of whether the innate constraints involved in musical behaviour developed at least in part from *selection pressures specific to musical behaviour*. It is not clear how evidence from neural encoding and modularity, early infant abilities, cross-cultural studies, studies of animal

capabilities, and computational approaches will answer this question.

In summary, the distinction between adaptation and exaptation is not very clear for higher cognitive functions such as music and language, where genes set up general neural architectures and learning mechanisms rather than determining the details of neural circuits and the contents of what they will represent, and where individual learning and culture play a large role in the development of abilities. A determination of whether music is an evolutionary adaptation will depend on new ways of approaching the question of whether it actually exerted selection pressures independently from other cognitive abilities. Whatever the answer to this question, music is a prominent, universal human activity that depends on our genetic

heritage, and we should continue to explore its origins and its relation to other cognitive and social functions.

Author Note

The writing of this paper was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes for Health Research. Thanks to Sigal Balshine for comments on an earlier draft.

Address correspondence to: Laurel J. Trainor, Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON L8S 4K1. E-MAIL LJT@mcmaster.ca

References

- BOSNYAK, D. J., EATON, R. A., & ROBERTS, L. E. (2004). Distributed auditory cortical activations are modified when nonmusicians are trained at pitch discrimination with 40-Hz amplitude modulated tones. *Cerebral Cortex*, *14*, 1088–1099.
- COSTA-GIOMI, E. (2003). Young children's harmonic perception. In G. Avanzini, C. Faienze, D. Minciocchi, L. Lopez, & M. Majno (Eds.), *The neurosciences and music: Annals of the New York Academy of Sciences* (Vol. 999, pp. 477–484). New York: New York Academy of Sciences.
- CUDDY, L. L., & BADERTSHER, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception and Psychophysics*, *41*, 609–620.
- ELMAN, J. L., BATES, E. A., JOHNSON, M. H., KARMILOFF-SMITH, A., PARISI, D., & PLUNKETT, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- HANNON, E. E., & TREHUB, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, *16*, 48–55.
- IZUMI, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America*, *108*, 3073–3078.
- JUSTUS, T., & HUTSLER, J. J. (2005). Fundamental issues in evolutionary psychology of music: Assessing innateness and domain specificity. *Music Perception*, *23*, 1–27.
- KOELSCH, S., GUNTER, T. C., V. CRAMON, D. Y., ZYSSET, S., LOHMANN, G., & FRIEDERICI, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage*, *17*, 1956–1966.
- KRUMHANSL, C. L., & KEIL, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition*, *10*, 243–251.
- 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.
- PERETZ, I. (2003). Brain specialization for music: New evidence from congenital amusia. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 192–203). Oxford: Oxford University Press.
- PHILLIPS-SILVER, J., & TRAINOR, L. J. (2005). Feeling the beat in music: Movement influences rhythm perception in infants. *Science*, *308*, 1430.
- PLANTINGA, J., & TRAINOR, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, *98*, 1–11.
- TERHARDT, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, *55*, 1061–1069.
- TRAINOR, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development*, *19*, 83–92.
- 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*, 188–195.
- TRAINOR, L. J., CLARK, E. D., HUNTLEY, A., & ADAMS, B. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, *20*, 383–396.
- TRAINOR, L. J., & TREHUB, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure.

- Journal of Experimental Psychology: Human Perception and Performance*, 18, 394–402.
- TRAINOR, L. J., & TREHUB, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics*, 56, 125–132.
- TRAINOR, L. J., TSANG, C. D., & CHEUNG, V. H. W. (2002). Preference for consonance in 2- and 4-month-old infants. *Music Perception*, 20, 187–194.
- TREHUB, S. E., BULL, D., & THORPE, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55, 821–830.
- 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.
- ZATORRE, R. (2003). Neural specialization for tonal processing. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 231–246). Oxford: Oxford University Press.
- ZHANG, L. I., BAO, S., & MERZENICH, M. M. (2002). Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. *Proceedings of the National Academy of Sciences*, 99, 2309–2314.

