

## Music: Specialized to Integrate?

PAULO ESTÊVÃO ANDRADE

*Department of Pedagogical Studies, Colégio Criativo, Marília, Brazil  
Laboratory of Investigation of Learning Deviations (LILD), Faculty of Philosophy and Sciences, São Paulo  
State University - Marília (SP), Brazil*

JOYDEEP BHATTACHARYA[1]

*Department of Psychology, Goldsmiths, University of London*

**ABSTRACT:** In her paper Schaefer (2014) provides a relevant amount of behavioral and neuroimaging evidence within and outside the realm of music favoring the notion that predictive processing plays a prominent role in the coupling of perception, cognition and action, and further, that imagery and active perception are closely associated with each other. Central to this review is that research into music imagery is exceptionally suitable and informative since prediction has a prominent role in music processing. In this commentary we suggest that it could be useful to investigate the role of working memory in this context since imagery and memory are inextricably associated processes. In addition to neuroimaging we also highlight that anthropological and developmental evidence could be relevant in showing that music is possibly unique in the coupling of perception, cognition and action. However, we believe that greater caution is needed regarding the author's assumption that perception and interpretation of music is uniquely determined by the listening biography of the listener.

Submitted 2014 November 15; accepted 2014 December 21.

**KEYWORDS:** *predictive coding, imagery, action, perception, universality*

### INTRODUCTION

AS pointed out by McClelland (1996, p. 633) “The sciences of mind and brain have seen an alternation between global approaches and more modular approaches”. Indeed, during the first half of the 20th century, global approaches were more popular to explain cognition in terms of a few general principles, such as the Gestalt psychology (Köhler, 1969), genetic epistemology (Piaget, 1971), and behaviorism (Skinner, 2011). The second half of that century, in contrast, witnessed the emergence of the notion that some important aspects of cognition, like language, can arise from the independent activity of neural populations specialized for representing different types of information referred to as autonomous modules. Among these modular approaches are those proposed for language by Noam Chomsky (Chomsky, 1980) and for vision by David Marr (Marr, 1982), which were subsequently systematized as a general approach for brain functioning by the philosopher Jerry Fodor (Fodor, 1983). According to Fodor (1983), cognitive abilities are fast (automatic), hardwired or localized (mediated by dedicated neural systems), domain-specific (only activated by, for example, language-specific information), informationally encapsulated (do not interfere with other modules), and innate (pre-programmed).

For instance, according to the modularity view, language is a large computational module, mainly subserved by the left perisylvian cortical areas, which are autonomous from other cognitive functions and comprised of different submodules each with its own functional and neural architecture, such as phonology, syntax and semantics, which operate serially (phonology precedes syntax which precedes semantics) and independently from each other (encapsulated) so that grammar is not influenced by meaning (Chomsky, 1980; Fodor, 1983; Pinker, 1994). A modular view of music has also been proposed since neural networks in the right hemisphere appear to be crucial for the perception and recognition of melodies (Peretz & Coltheart, 2003).

In his reflections, however, McClelland (1996) gave prominence to those theoretical trends falling in between the global and modular approaches, i.e. those who argue for the importance of recognizing evidence of relatively separate cognitive systems but also recognize the importance of understanding how these specialized areas work together in order to give rise to cognition. Among these intermediate views, McClelland highlighted Luria's contribution (1966) to cognitive neuroscience that the cerebral cortex is organized according to three separate and conceptually distinct types of areas. According to Luria (1966) the primary and secondary areas are populated exclusively by modality-specific neurons (including specific aspects or subdomains of visual, auditory, tactile and motor stimuli) and whose responses occur primarily for simple and local features and for the more complex conjunction of features of the stimulus, respectively; tertiary areas are populated by non-modality specific neurons that compute spatial and temporal aspects of the stimulus, its semantic content and communicative purposes. Inspired by Luria, interactive approaches assume bidirectional interactions among neurons of specialized areas representing different types of information, so that instead of being viewed in isolation, these areas should be seen as working together in an integrated/interactive way in order to understand how they give rise to system-level functions such as perception, communication, and action (McClelland, 1996, p. 634). The interactive view of brain functioning has recently been championed by many authors (Mesulam, 1998; Rumelhart, McClelland, & PDP Research Group, 1995). Indeed, particularly in the last fifteen years, a significant number of behavioral and neurological studies favor the notion that cognitive outcomes arise from mutual, bidirectional interactions among neuronal populations representing different types of information (McClelland, 2001), indicating that many primary brain areas (e.g., primary visual or auditory cortices), traditionally viewed as modality specific, appear to participate in the representation of the global structure of a stimulus or response situation and are modulated by mutual influences from each other. The discovery of various cross-modal interactions suggests that these interactions are the rule rather than exception in human perceptual processing (Driver & Noesselt, 2008; Senkowski, Schneider, Foxe & Engel, 2008).

In the above summarized scenario comes the insightful contribution by Schaefer (2014) in which the author aligns herself with the notion of an interactive neurofunctional architecture of the brain, providing a relevant amount of behavioral and neuroimaging evidence within and outside the realm of music suggesting that perception, cognition and action are intricately coupled, being modulated by mutual influences. Central in Schaefer's review is the notion of predictive coding, "which is a unified framework for understanding redundancy reduction and efficient coding in the nervous system" (Huang & Rao, 2011, p. 580); here, the brain is essentially considered as a predicting machine that is in predicting mode perpetually as it makes the prediction from learned models of previous sensory-cognitive experiences concerning what will come next, and compares this with incoming sensory data. This predictive coding is supposed to play a prominent role in the coupling of perception, cognition and action and, further, in the claim that imagery and active perception are closely associated to each other (Clark, 2013). Finally, within this framework, Schaefer highlights how exceptionally suitable and informative research into music imagery is to investigate these interactive mechanisms because of the crucial importance of predictive mechanisms and statistical learning in music processing in general, including music perception. In the present commentary we emphasize the importance of working memory processes in imagery and discuss anthropological and developmental evidence in order to argue that music is possibly unique in involving and integrating several cognitive domains.

### **IMAGERY: TYPES AND STAGES**

Schaefer offers definitions of four types of imagery as described by Moore (2010), namely, sensory imagery, creative imagery, propositional imagery and constructive imagery. Schaefer has selected only two of these four types of imagery as the most suitable for her paper: sensory imagery and constructive imagery. Sensory imagery is the most studied type of imagery by psychologists and refers to the mental recreation of a sensory experience. Constructive imagery occurs when percepts are organized for coherent cognition and interpretation. The author also details some differences between both types of imagery by noting that "sensory imagery is a deliberate task that requires some effort or concentration, whereas constructive imagery happens automatically, as a part of making sense of incoming information" (p.163), and she adds that constructive processes "may *also* be present in deliberate, effortful sensory imagery, in some sense closing the gap with other uses of the term 'constructive imagery'". One important claim by Schaefer is that constructive imagery is included in the sensory imagery since "a vividly imagined stimulus comprises structural and temporal features that need to be perceptually organized" (p.167) and, in this

sense, the notion of constructive imagery embraces the overlap between sensory imagery and active perception and fits well with a theoretical framework of predictive processing as a unifying mechanism in perception, action and cognition. In this theoretical framework, predictions are based on multimodal integration processes that match sensory-perceptive inputs and long-term representations creating top-down expectations or predictions, and are of great adaptive value because successes or failures are associated with significant psychological and physiological consequences (Clark, 2013; Schultz, Dayan & Montague, 1997; Steinbeis, Koelsch & Sloboda, 2006).

We suggest here that it could be useful to discuss these definitions also in terms of the working memory processes which are relevant to each type of imagery, considering that imagery and memory are inextricably associated processes. For instance, definition of sensory image by Schaefer as a mental (re-) creation of a sensory experience has as its core meaning the involvement of memory. Schaefer's definition is therefore consistent with Kosslyn's definition of mental imagery as perceptual information being accessed from memory (Kosslyn, Ganis, & Thompson, 2001) and to Hubbard's notion of experimenting sensory qualities, including those drawn from long-term memory, in the absence of corresponding stimuli (Hubbard, 2010, p. 302; Hubbard, 2013, p.52). Moreover, by claiming that constructive imagery is driven by implicit internal representations based on person-specific knowledge and experience, Schaefer clearly assumes the crucial role of long-term memory representations in this type of imagery.

Therefore we suggest that it would be worth investigating the notion that constructive imagery can be seen as involving two stages on the basis of the memory processes relevant to each stage: a first stage characterized by the sensory imagery in which perceptual information is held in the short-term memory, and a second-stage characterized by the constructive imagery which is driven by implicit, long-term internal representations based on person-specific knowledge and experience which are instrumental in interpreting what is perceived. The view of sensory and constructive imageries as differing mainly on the basis of memory processes relevant to each type of imagery fits well with the claim by Cebrian and Janata (2010) that mental auditory representations can arise from either a "bottom-up" sensory processing of an external sound source, thus involving mainly short-term memory process, or a "top-down" process in which an auditory sound is anticipated or imagined on the basis of matching perceptual information to stored representations in long-term memory, which also require working memory processes.

For instance, Halpern & Zatorre (1999) note that imagining both familiar and nonfamiliar melodies after presentation of the first few notes involve working memory processes because melodies have to be remembered before being rehearsed (imagery). The difference between imagery tasks for familiar and nonfamiliar melodies is that in the first subjects are required to retrieve the melody from semantic memory after hearing the first few notes, whereas for nonfamiliar melodies subjects first hear the novel melody entirely before the first few notes are presented. Halpern and Zatorre (1999, p. 698) argued that subtracting brain activations for nonfamiliar melodies from brain activations for familiar melodies "should remove the effects of hearing a note sequence and thereby isolate imagery and working memory processes". In other words, this rationale illustrates how we can distinguish long-term from short-term memory processes in imagery. Interestingly, in perceptual tasks that do not explicitly require imagery, for instance, when subjects are asked to tell if the second of two unfamiliar melodies presents a change or not, working-memory processes support these discrimination tasks even when long-term melodic representations are impaired by lesions involving right anterior fronto-temporal areas, whereas in other patients with posterior brain lesions these representations can be shown to be preserved in singing despite the fact that recognition of familiar melodies is impaired (see Peretz, 2006). Indeed, Halpern & Zatorre (1999) found that semantic retrieval, in addition to auditory and premotor-parietal activations during imagining nonfamiliar melodies, induced significant activations in the right inferior frontal and bilateral middle frontal areas (more significant on the right side).

Neuroimaging studies on healthy non-musicians using discrimination tasks with non-familiar (Zatorre, Evans & Meyer, 1994; Gaab, Gaser, Zaehle, Jancke, & Schlaug, 2003) and either familiar or non-familiar (Platel *et al.*, 1997) musical stimuli have systematically shown that working memory processes are intimately linked to imagery. In addition to activations in the inferior frontal and inferior parietal areas known to be involved in auditory working memory, these melodic tasks also activate a premotor-posterior parietal network, including superior and medial parietal lobes (precuneus) which is crucial not only to motor control, but also to cognitive processes such as mental object-construction task (Mellet *et al.*, 1996) and other visuo-spatial tasks involving spatial working memory either implicitly (Haxby *et al.*, 1994) or explicitly (Smith, Jonides & Koeppel, 1996). These activations have been particularly strong in non-imagery pitch tasks specifically designed to demand working memory processes such as comparing the first

and last notes of eight-note melodies (Zatorre *et al.*, 1994) and either the last or the second to last tone to the first tone of a 6 to 7 sine-wave tone sequences (Gaab *et al.*, 2003). Activation of this premotor-parietal network has been justified by the argument that even in non-musicians these analytic pitch tasks demand visual imagery in terms of “high” and “low”, likely forming a notational mental base line or a mental staff/score (Zatorre *et al.*, 1994; Platel *et al.*, 1997; Gaab *et al.*, 2003), as declared even by naïve listeners after debriefing (Platel *et al.*, 1997).

### IMAGERY, PREDICTION AND MUSIC

Schaefer emphasizes that constructive imagery is crucial to predictive processes which, in turn, “are a driving force in perception, cognition and action”, a notion which is in line with many relevant theoretical frameworks according to which the brain has evolved to anticipate or predict forthcoming events. This efficacy has significant adaptive consequences at both psychological and physiological levels (Wiggins & Bhattacharya, 2014; see Hohwy, 2014 for an excellent non-technical introduction). In this perspective music comes to be of great importance, since the essence of musical composition, musical meaning and appreciation, as well as aesthetic experiences with music and musical emotions are closely connected to confirmation and violation of expectations (Huron, 2006).

Expectation is a basic component of music perception, operating on a variety of levels, including melodic, harmonic, metrical, and rhythmic, and addresses the questions “what” and “when”, that is, what tones or chords are expected to occur and when in a given musical sequence. It is not only presumed to play an important role in how listeners group the sounded events into coherent patterns, but also to appreciate patterns of tension and relaxation contributing to music's emotional effects. Both cognitive and emotional responses to music depend on whether, when, and how the expectations are fulfilled.

Moreover, despite evidence that some cognitive mechanisms crucial to music processing appear to be domain-specific (Peretz & Coltheart, 2003; Peretz, 2006), it is also largely recognized that musical processing is to a great extent transmodal, involving the coupling of perception, cognition and action (Janata & Grafton, 2003). By drawing this brief link we believe that we arrive at the gist of Schaefer's argument in favor of investigating imagery processes in music. Since imagery can be seen as inextricably associated to prediction (a cognitive ability of crucial adaptive value) and since prediction and imagery are the essence of music processing (Hargreaves, 2012), investigating constructive imagery processes in music would be promising for providing invaluable information about how mental models of music are instantiated in the brain.

### COUPLING OF PERCEPTION, COGNITION AND ACTION IN MUSIC

Coupling of perception, cognition and action in music has been emphasized by Schaefer (2014) on the basis of cognitive and neurocognitive evidence (see also Zatorre, Chen, & Penhune, 2007). However, we consider that it may also be important to highlight compelling evidence coming from anthropological and developmental approaches. Music simply moves us. When we are playing, tapping, dancing, or singing along with music, the sensory experience of musical patterns is intimately coupled with action, suggesting musical cognition is necessarily embodied (Maes, Leman, Palmer, & Wanderley, 2013).

Cross-culturally, music involves not only patterned sound, but also overt and/or covert action, and even “passive” listening to music can involve activation of brain regions concerned with movement (Janata & Grafton, 2003; Sievers, Polansky, Casey, & Wheatley, 2013). Developmental precursors of music in infancy, and through early childhood, occur in the form of universal proto-musical behaviors (Trevorthen, 2000) which are exploratory and kinesthetically embedded and closely bound to vocal play and whole body movement. They are also universal. This leads some cultures to employ terms to define music that are far more inclusive than the Western notion of music, like the word *nkwa* that for the Igbo people of Nigeria denotes “singing, playing instruments and dancing” (Cross, 2001, p. 29). Music functions in many different ways across cultures, from a medium for communication (the Kaluli of Papua New Guinea used music to communicate with dead members), for restructuring social relations (the Venda tribes use music for the *domba* initiation), to constitute a path to healing and establishing sexual relationships (Tribes in northwest China play the “flower songs”, *hua'er*) (Cross, 2003). Thus, music is a property not only of individual cognitions and behaviors but also of inter-individual interactions. Developmental studies give further support for the notion that music is embodied.

A common observation in the social-developmental literature is that parent-infant games are often reciprocally imitative in nature. Human infants interact with their caregivers producing and responding to patterns of sound and action. These temporally-controlled interactions involve synchrony and turn-taking, and are employed in the modulation and regulation of affective states (Dissanayake, 2000) and in the achievement and control of joint attention. This rhythmicity is also a manifestation of a fundamental musical competence, and "...musicality is part of a natural drive in human sociocultural learning which begins in infancy" (Trevarthen, 1999, p. 194) and also allows infants to follow and respond in kind to temporal regularities in vocalization, movement, and time, to initiate temporally regular sets of vocalizations and movements (Trevarthen, 1999). When parents shake a rattle or vocalize, infants are likely to shake or vocalize back. A turn-taking aspect of these games is the "rhythmic dance" between mother and child. Adults across cultures play reciprocal imitative games with their children that embody the temporal turn-taking (Trevarthen, 1999). We know that imitation games with music and dance are universal, and the tribal dances itself can be seen as one of the most frequent forms of imitation game, used to develop the in-group sense, the feeling of both "being like the other" and "the other is like me", and thus pertaining to a group. Perhaps not coincidentally, the developmental precursors of music in infancy and through early childhood occur in the form of proto-musical behaviors which are exploratory and kinesthetically embedded, being closely bound to vocal play and to whole body movement. Thus, it is not surprising to observe a significant overlap between neural substrates underlying motor and visuo-spatial tasks and music perception (Zatorre, Chen, & Penhune, 2007). For example, occipital and frontoparietal cortical areas, traditionally involved in visuo-spatial tasks, including the precuneus in medial parietal lobes called the "mind's eye" for being crucially involved in generation of visuo-spatial imagery (Mellet *et al.*, 2002; Mellet *et al.*, 1996), are among the most frequently and significantly activated regions in perception of music either in naïve listeners or in musicians (Nakamura *et al.*, 1999), even in participants performing a task of musical imagery (Halpern & Zatorre, 1999; Meister *et al.*, 2004). Bilateral frontal and inferior frontal activations, mainly premotor frontal areas BA6, dorsolateral prefrontal areas (BA8/9), as well as inferior frontal areas as Broca (BA44/45), insula, and more anterior middle and inferior frontal cortices (46/47), are frequently observed in non-musicians (Platel *et al.*, 1997; Zatorre *et al.*, 1994) and musicians (Ohnishi *et al.*, 2001; Zatorre *et al.*, 1998).

## DISCUSSION

Literature on music processing is consistent with the notion that music can be thought of as a sequence of events that are patterned in time and a "feature space" that is multidimensional and consists of both motor and sensory information (Janata & Grafton, 2003, p. 682) and, hence, supports Schaefer's claim that research on music, with its multi-level structure and the crucial importance of predictive mechanisms and statistical learning for its processing, is exceptionally suitable and informative to investigate concerning interactive mechanisms and mental models.

We also embrace the notion presented by Schaefer (2014) that "attentive music listening is an active process in which we track incoming information on multiple structural levels" and we add that it necessarily requires short-term processes. We argue that short-term processes are inextricably associated with attentive music listening. We further suggest that even perceptual tasks, i.e. discriminating or comparing pairs of melodies, involve imagery at some extent and that explicitly imagery tasks are even more demanding in short-term memory (Halpern & Zatorre, 1999; Hubbard, 2010). Therefore, we emphasize the relevance of investigating memory processes involved in imagery.

On the other hand, although we agree with Schaefer's statement that "there is no assurance that two people will feel the same way about a piece of music, or even necessarily have the same experience at repeated listening" (p.161) we are more cautious in assuming the author's claim that music "is perceived and interpreted through processes driven by the unique listening biography of the listener" (p.161), because several aspects of music, such as the categorization of the octave into a set of tones, the special status of the octave and perfect fifth, pitch processing relative to scales and contours, and basic principles of grouping and meter, are universal among cultures and its perception emerges early in development (Justus & Hutsler, 2005). There is also evidence of common features, across different music styles, in the principles underlying melody and in response to features such as consonant/dissonant intervals, pitch in musical scales and meter. Many of these features are already present in infants when responding to melodic as well as to harmonic consonant patterns, and to complex metric rhythms (Andrade & Bhattacharya, 2003; Justus & Hutsler, 2005; McDermott & Hauser, 2005).

As Schaefer has noted, “interpreting incoming information based on previous experience is also consistent with a statistical learning-based account of music perception” (p.163). However, it is also noteworthy that cross-culturally shared sensitivity to the statistical properties of music influences listeners’ expectations (Oram & Cuddy, 1995; Krumhansl *et al.*, 2000), and is apparently a universal process, exhibited by even 8-month-old infants (Saffran, Johnson, Aslin & Newport, 1999). The literature suggests that in addition to cultural cues listeners possess innate general psychological principles in auditory perception such as sensitivity to consonance, interval proximity, and, finally, the statistical properties, which have been extensively shown to influence listeners’ expectations (Oram & Cuddy, 1995; Krumhansl *et al.*, 2000). For instance, when a melody is presented to the listeners many times, but occasionally altered by a single probe-tone with varying degree of fitness, some universalities are observed in the subject responses (Krumhansl *et al.*, 2000). Cross-cultural studies comparing the fitness ratings given by Indian and Western listeners to North Indian ragas (Castellano *et al.*, 1984), and by Western and Balinese listeners to Balinese music (Kessler, Hansen & Shepard, 1984), as well as native Chinese and American listeners’ responses to Chinese and British folk songs (Krumhansl, 1995) found strong agreement between listeners from these different musical cultures. South Asians (Indians) and Western listeners (Castellano *et al.*, 1984) showed a remarkable consistency in giving higher ratings to the tonic and the fifth degree of the scale, tones that were also sounded most frequently and for the longest total duration and which, theoretically, are predicted to be important structural tones in Indian ragas. Interestingly, Western listeners’ responses did not correspond to tonal hierarchies of major and minor keys, according to the Theory of Harmony in Western music, but rather to theoretical predictions for Indian music. Listeners ranging from experts to completely unfamiliar, namely Sami yoikers (experts), Finnish music students (semi-experts) and Central European music students (Krumhansl *et al.*, 2000) as well as traditional healers from South Africa (non-experts) (Eerola, 2004) rated with strong agreement the fitness of probe-tones as continuations of North Sami yoik excerpts, a musical style from indigenous people (Sami people sometimes referred to as Lapps) of the Scandinavian Peninsula that is quite distinct from Western tonal music (Krumhansl *et al.*, 2000; Eerola, 2004).

It is undeniable that enculturation to a musical tonal system leads to development of culture-specific brain structures, in a way analogous to the learning of a first language (Hannon & Trainor, 2007). This enculturation, such as the implicit knowledge of harmony specific to Western cultures, is also important for emotional experience and emotion perception in music (Juslin & Laukka, 2004). However, there is also evidence that certain basic emotions expressed in music can be understood across cultures, despite dramatic cultural differences. By asking Mafa, a culturally isolated African people from Cameroon and naïve to Western music, and Western listeners to judge happiness, sadness and fearfulness expressed in Western music, Fritz *et al.* (2009) found that both Western and Mafa listeners recognized these three emotions above chance level (but do note that happy emotion was recognized substantially better than the other two). Finally, the authors also found that both Mafa people and Western listeners preferred original versions of both Mafa and Western music over their dissonant versions.

More recently, Perani *et al.* (2010) found that 1- to 3-day-old newborns show differential patterns of brain activity when listening to excerpts of classical music pieces, which mainly evoked right-hemispheric activations in primary and higher order auditory cortex, and when listening to their altered versions in tonal key (in some bars all voices were shifted one semitone upward or downward, thus infrequently shifting the tonal center) or their dissonant versions (when the upper voice of the melody was permanently shifted one semitone upward, rendering the excerpts permanently dissonant), which significantly reduced right auditory cortex activations and elicited significant activations in the left inferior frontal cortex and limbic structures. Since both altered versions contained a higher degree of sensory dissonance than the original music and since significant activation differences occurred when directly comparing original vs. altered music, the authors concluded that the right-lateralized auditory cortex activation during original music could not be considered simply an unspecific response that could have been elicited by any sound in general; instead it was attributable to mainly consonant and structured original music.

In conclusion, we welcome the timely contribution of Schaefer (2014) and hope that it will stimulate researchers to explore further the role of predictive processes in music cognition (Vuust, Ostergaard, Pallesen, Bailey & Roepstorff, 2009; Pearce, Ruiz, Kapasi, Wiggins & Bhattacharya, 2010).

## ACKNOWLEDGMENTS

Joydeep Bhattacharya was supported by EPSRC Research Grant EP/H10294X, “Information and neural dynamics in the perception of musical structure”. We dedicate this article to Andre Michel Andrade who died unexpectedly at the time of writing this commentary.

## NOTES

[1] Correspondence can be addressed to: Joydeep Bhattacharya, Department of Psychology, Goldsmiths, University of London, UK. Email: j.bhattacharya@gold.ac.uk

## REFERENCES

- Andrade, P. E., & Bhattacharya, J. (2003). Brain tuned to music. *Journal of the Royal Society of Medicine*, 96(6), 284-287.
- Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of north India. *Journal of Experimental Psychology: General*, 113(3), 394.
- Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, 3(1), 1-15.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03), 181-204.
- Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences*, 930(1), 28-42.
- Cross, I. (2003). Music as a biocultural phenomenon. *Annual New York Academy of Science*, 999, 106-111.
- Dissanayake, E. (2000). Antecedents of the temporal arts in early mother–infant interactions. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 388–410). Cambridge, MA: MIT Press.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron*, 57(1), 11-23.
- Eerola, T. (2004). Data-driven influences on melodic expectancy: Continuations in North Sami Yoiks rated by South African traditional healers. In *Proceedings of the eighth international conference of music perception and cognition* (pp. 83-87).
- Fodor, J. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A., & Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, 19(7), 573-576.
- Gaab, N., Gaser, C., Zaehle, T., Jancke, L., & Schlaug, G. (2003). Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *Neuroimage*, 19(4), 1417-1426.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9(7), 697-704.
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in cognitive sciences*, 11(11), 466-472.
- Hargreaves, D. J. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of Music*, 40(5), 539–557.

- Haxby, J. V., Horowitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience*, *14*(11), 6336-6353.
- Hohwy, J. (2014). *The predictive mind*. New York: Oxford University Press.
- Huang, Y., & Rao, R. (2011). The predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 580-593.
- Hubbard, T. L. (2013). Auditory aspects of auditory imagery. In *Multisensory Imagery* (pp. 51-76). New York: Springer.
- Hubbard, T.L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*(2), 302-329.
- Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- Janata, P., & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, *6*, 682-687.
- Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*(3), 217-238.
- Justus, T., & Hutsler, J. J. (2005). Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain specificity. *Music Perception*, *23*(1), 1-27.
- 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.
- Köhler, W. (1969). *The task of Gestalt psychology*. Princeton, NJ: Princeton University Press.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635-642.
- Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*, *17*(1), 53-80.
- Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, *76*(1), 13-58.
- Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.
- Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.
- Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. M. (2013) Action-based effects on music perception. *Frontiers in Psychology*, *4*, 1008.
- McClelland, J. L. (1996). Integration of information: Reflections on the theme of attention and performance XVI. In T. Inui & J. L. McClelland (Eds.), *Attention & Performance XVI: Information Integration in Perception and Communication*, (633-656). Cambridge, MA: MIT Press.
- McClelland, J. L. (2001). Cognitive neuroscience. In N. J. Smelser & Paul B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*, (pp. 2133-2139). Oxford: Pergamon.



- McDermott, J., & Hauser, M.D. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23*, 29-59.
- Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R., & Thron, A. (2004). Playing piano in the mind – an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research, 19*(3), 219-228.
- Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M. & Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. *Cerebral Cortex, 12*(12), 1322-1330.
- Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., & Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *The Journal of Neuroscience, 16*(20) 6504-6512.
- Mesulam, M. M. (1998). From sensation to cognition. *Brain, 121*(6), 1013-1052.
- Moore, M. E. (2010). *Imagination and the mind's ear*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.
- Nakamura, S., Sadato, N., Ohashi, T., Nishina, E., Fuwamoto, Y., & Yonekura, Y. (1999). Analysis of music-brain interaction with simultaneous measurement of regional blood flow and electroencephalogram beta rhythm in human subjects. *Neuroscience Letters, 275*(3), 222-226.
- Navarro Cebrian, A., & Janata, P. (2010). Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. *Brain Research, 1342*, 39–54.
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., & Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex, 11*(8), 754-760.
- Oram, N., Cuddy, L. L., & Oram, N. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research, 57*(2), 103-118.
- Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A. & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioral, and neural manifestations of musical expectation. *NeuroImage, 50*, 302–313.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences, 107*(10), 4758-4763.
- Peretz, I. (2006). The nature of music from a biological perspective. *Cognition, 100*(1), 1-32.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688-691.
- Piaget, J. (1971). *Genetic epistemology*. New York: Columbia University Press.
- Pinker, S. (1994) *The language instinct: How the mind creates language*. New York: Harper Perennial.
- Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., ... & Eustache, F. (1997). The structural components of music perception. A functional anatomical study. *Brain, 120*(2), 229-243.
- Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1995). *Parallel distributed processing* (Vol. 1, pp. 184). Cambridge, Mass.: MIT press.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27-52.

- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599.
- Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: Implications for multisensory processing. *Trends in neurosciences*, 31(8), 401-409.
- Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences*, 110(1), 70-75.
- Skinner, B. F. (2011). *About behaviorism*. New York: Vintage.
- Smith, E. E., Jonides, J., & Koeppel, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, 6(1), 11-20.
- Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18(8), 1380-1393.
- Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae*, 3(1), 155–215 (special issue 1999–2000).
- Trevarthen, C. (2000). Autism as a neurodevelopmental disorder affecting communication and learning in early childhood: prenatal origins, post natal course and effective educational support. *Prostaglandins Leukotrienes and Essential Fatty Acids*, 63(1-2), 41-46.
- Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music – brain responses to rhythmic incongruity. *Cortex*, 45(1), 80-92.
- Wiggins, G. A., & Bhattacharya, J. (2014). Mind the gap: An attempt to bridge computational and neuroscientific approaches to study creativity. *Frontiers in Human Neuroscience*, 8, 540.
- Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *The Journal of Neuroscience*, 14(4), 1908-1919.
- Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., & Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences*, 95(6), 3172-3177.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), 547-558.