The emergence of music from the Theory of Mind

Steven Robert Livingstone

Department of Psychology

Macquarie University

Sydney, NSW, 2109, Australia

Steven.Livingstone@psy.mq.edu.au

William Forde Thompson

Department of Psychology

Macquarie University

Sydney, NSW, 2109, Australia

Bill.Thompson@mq.edu.au

Correspondence to Steven.Livingstone@psy.mq.edu.au

Abstract

It is commonly argued that music originated in human evolution as an adaptation to selective pressures. In this paper we present an alternative account in which music originated from a more general adaptation known as a Theory of Mind (ToM). ToM allows an individual to recognise the mental and emotional state of conspecifics, and is pivotal in the cultural transmission of knowledge. We propose that a specific form of ToM, *Affective Engagement*, provides the foundation for the emergence of music. Underpinned by the mirror neuron system of empathy and imitation, music achieves engagement by drawing from pre-existing functions across multiple modalities. As a multimodal phenomenon, music generates an emotional experience through the broadened activation of channels that are to be empathically matched by the audio-visual mirror neuron system.

Introduction

Adaptationist theories of music origins dominate the intellectual landscape. The resounding tenet is that music is the result of biological evolution - that within each of us are genes that evolved specifically to create, perform and engage with music. Adaptationist accounts operate under the premise that changes in the genotype that promoted musical behaviours conferred reproductive or survival benefits for those who possessed these genes. Among the current adaptationist theories are proposals that musical behaviours were selected because they promote the mother-infant bonds (Dissanayake, 2000a; Falk, 2004; Trehub, 2001), social bonds and group coordination (Brown, 2000a; Dunbar, 1993, 1996; Kogan, 1997; Roederer, 1984), sexual selection (Dean *et al.*, in press; Miller, 2000) and cognitive development or play (Cross, 1999, 2003b; Papousek, 1996).

Two recent reviews of such accounts urged researchers to adopt rigorous criteria for making adaptationist claims (Justus & Hutsler, 2005; McDermott & Hauser, 2005). These critiques did not propose alternatives to adaptationist accounts, but warned against the premature acceptance of music as an evolutionary adaptation when exaptation and culture may be sufficient. Exaptation refers to a structure or attribute whose current use differs from its originally evolved function (Gould & Lewontin, 1979). For example, bird feathers may have originally arisen for thermal regulation, but their use was later exapted for flight (Buss, Haselton, Shackelford, Bleske, & Wakefield, 1998; Norell *et al.*, 2002; Regal, 1975).

Non-adaptationist models have been proposed (Barrow, 1995; Sperber, 1996), but it was Pinker's (1997) non-adaptive pleasure-seeking account that motivated the development of a large number of adaptationist models (Huron, 2001). Pinker proposed that music emerged in the absence of music-specific genes as a purely hedonistic pursuit, drawing from pre-existing domains of emotion, motor control, auditory scene analysis, habitat selection and language. While Pinker's claims have been questioned on a number of grounds (Cross, 2003a; Fitch, 2006; Huron, 2001; Trehub, 2001), there has been a curious lack of attention to the possibility that musical origins can be explained without the need to invoke music-specific genes.

There are a number of difficulties with the view that musical behaviours were naturally selected and encoded in the human genotype (Fitch, 2006; Patel, 2008). Ethnomusicologists have struggled to locate universals beyond a series of basic elements (Nettl, 2000). In contrast, musical behaviours and functions such as dance (Nettl, 2000), its use in rituals and ceremonies (Cross, 2001), and its connection with affect (Dissanayake, 2000a) appear almost universal. It is this broader nature of music that will frame our discussions of music origins. Cross proffers a suitably encompassing definition of music for this task - "[m]usic embodies, entrains and transposably intentionalises time in sound and action" (Cross, 2003a, p. 80). In this view, music is not restricted to sounded events but extends to actions such as dance.

The evidence for a genetic explanation of music is modest. Justus and Hutsler (2005) conclude that only tonal encoding of pitch may be innate (genetic) and specific to the domain of music (McDermott & Hauser, 2005; Peretz, 2006). Deficits in music processing such as acquired and congenital amusia may also be relevant (Ayotte, Peretz, & Hyde, 2002; Peretz, 2001a, 2001b; Peretz *et al.*, 2002; Peretz & Hyde, 2003). Evidence for modularity of musical functions provides little support one way or the other for theories of musical origins. Double dissociations have been demonstrated in which music sight-reading (a recent cultural invention) is impaired while text reading remains at normal performance levels (Brust, 1980; Cappelletti, Waley-Cohen, Butterworth, & Kopelman, 2000; Hébert & Cuddy, 2006). That is, cortical areas may become specialised for specific tasks merely because such regions are the most aptly suited to handle these tasks (Justus & Hutsler, 2005, Patel, 2008). Music also exhibits a close relation to language (Brown, 2000b; Lerdahl & Jackendoff, 1983; Pinker, 1997) where the apparent overlap in brain regions associated with the two domains is consistent with the possibility that music is an instance of a broader function (Koelsch *et al.*, 2002, 2004; Patel, 2003; Peretz & Zatorre, 2003). Finally, it has been suggested that early infant music abilities acts as an indicator of innate predispositions (Trehub, 2001, 2003). However, other evidence suggests that infant music perception relies on domain-general mechanisms (Trehub & Hannon, 2006).

We now explore an alternative hypothesis, that a ToM acted to bring together domain and non-domain specific cognitive faculties in the creation of music, as previously suggested for language (Tomasello, 1999; see also Tomasello, 1995). Selected faculties employed in the creation and experiencing of music are listed in Table 1. In the model that we outline below, these and other faculties were brought together in music to serve the broad purpose of affective engagement. Affective engagement, which is supported by ontogenetic and phylogenetic factors (Heyes, 2003), is proposed as a specific and advanced form of ToM that is strongly associated with human behaviour and cultural practice. Affective engagement rests on the foundation of more primitive forms of ToM, while various musical, artistic and other practices emerge as instances of affective engagement. Such musical and artistic practices are culture-specific, in that affective engagement can be accomplished in any number of ways. Just as affective engagement rests upon ToM, a range of musical and other cultural practices rest upon a human instinct for affective engagement.

Table 1. Cognitive faculties employed by Theory of Mind in music.

Cognitive Faculty	Music-specific Candidates	References
1. Tension and Expectancy	No	Huron, 2006
2. Affect and Empathy	No	Dissanayake, 2000b; Juslin & Sloboda, 2001; Scherer & Zentner, 2001
3. Habitat Selection	No	Pinker, 1997
4. Isochronic pulse	Yes	Arom, 2000; Jackendoff & Lerdahl, 2006; Merker, 2002; Patel, 2006, 2008
5. Language overlaps	No	Fitch, 2006
6. Vocal imitation[i]	No	Kugiumutzakis, 1993, 1998; Studdert-Kennedy, 1983; for imitation/mimicry process separation see Jones, 2007[ii]
7. Emotional Calls	No	Pinker, 1997
8. Motor Control	No	Haueisen & Knoesche, 2001; Huron, 2006
9. Tonal Encoding of Pitch	Yes	McDermott & Hauser, 2005; Jackendoff & Lerdahl, 2006; but see Trehub & Hannon, 2006
10. Auditory Scene Analysis	No	Bregman, 1990

First, we review evidence that music is an effective means of affective communication. Next, we discuss the connection between music, other arts and speech. Finally, we outline the strong connections between music and ToM, which appear to operate at both the lower level of empathy and simulated response, and at a higher cultural level.

Music as Affective Engagement

Underpinning much of the arts is a motivation to communicate and engage affectively with others. We propose that the motivation for this exchange lies in the desire of humans to understand and learn about the emotional states of others and themselves, and the ways in which this is communicated. This motivation for sharing of emotions is a crucial and unique component of human ToM (Tomasello, 2005). For this we introduce the notion of affective engagement, defined as the degree of emotional connectedness between two or more individuals for the exchange of affective state used in the construction of mental models of conspecific emotion. This term is more suitable than *primitive emotional contagion* described by Hatfield *et al.* (1994), which describes a form of emotional mimicry in which emotions are caught and synchronised with the sender through primitive, automatic and unconscious behaviours.

Affective engagement takes place in a variety of forms including music, dance, painting and sculpture, story telling and acting. Each form has the capacity for meaningful affective communication and relies on a variety of cross-domain, multimodal channels of expression (reviewed later in Table 2).

To support the hypothesis that music is used in the construction of mental models of conspecific emotion, we must first demonstrate that music can function as a means of emotional communication.

The Communication of Emotion in Music

Arguably the most enduring feature of music is its capacity to communicate and induce affect, and there are abiding discussions of how and why this occurs (Gabrielsson, 2001; Kivy, 2002; Scherer, 2004). Indeed, the capacity of music to engender affect in individuals in such powerful, meaningful and abstract ways has been a topic of vigorous research. While not all music is created for the specific purpose of eliciting affect, its universal application in rituals and its capacity to mark important events highlights its nuanced capacity to "change an individual's consciousness or […] ambiance of a gathering" (Nettl, 2000, p. 468). Here, music is used "to coordinate the emotions of participants, and thus to promote conjoinment" (Dissanayake, 2000a, p. 390).

Research on emotion and music (Juslin & Sloboda, 2001) has provided a wealth of psychological and physiological evidence on the subject. While much of the literature pertains to Western classical music, it nonetheless illustrates the affective potential of music. Figure 1 highlights the various actors and elements involved in the communication of emotion in Western tonal music. The framework is a case study that applies specifically to Western classical music, and hence illustrates only one possible instantiation of the universal capacity of music to communicate and induce emotional states. Other musical idioms instantiate affective engagement in other ways, with different actors, elements and structural relationships, operating in ways that are largely culture-specific. But all these systems, in spite of the many superficial differences, appear to serve a broader function of affective engagement.

In Western classical music the communication of emotion may be conceived as a multi-stage process that incorporates emotivist and cognitivist elements. Emotivists (Davies, 2003) recognise that emotion can be induced in listeners following exposure to musical stimuli, while cognitivists (Kivy, 1990) emphasise the expression and representation of emotional messages in music. Both associations between music and emotion are represented in the Music-Emotion Framework (adapted from Livingstone, Muhlberger, Brown, & Loch, 2007).

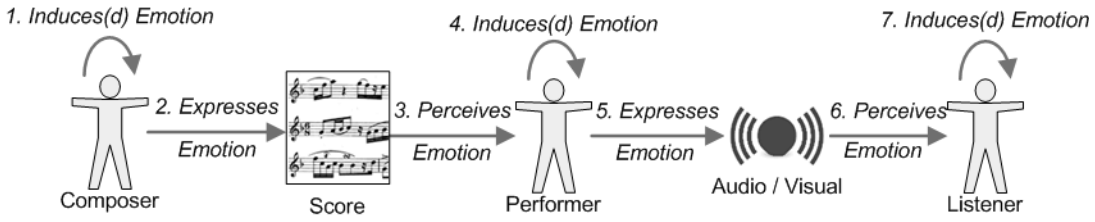

Figure 1. The Music-Emotion Framework illustrates the roles of composer, performer and listener in the communication of emotion that takes place in a Western classical music work. The framework does not mandate the occurrence of all steps, nor does it discuss the effects of listener feedback on the performer, or inter-performer, inter-listener communication. It should be emphasized that actor roles are often conflated, with intertwined emotion forms. Adapted from (Livingstone, Muhlberger, Brown, & Loch, 2007).

The framework highlights a distinction between three forms of emotion: induced, expressed and perceived (Gabrielsson, 2002; Juslin & Laukka, 2004; see also Keltner & Ekman, 2003). Induced emotion is that felt by, or elicited in, an observer in response to a given stimulus. Expressed emotion is the emotion embodied in audio, visual, or other forms of stimulus. Perceived emotion is that sensed/detected as being expressed in the stimulus. Each form of emotion is distinct and the link between each step is not always transitive. For example, a person need not express the same emotion that they are feeling (induced) due to social norms; conversely, an expressed emotion may not result in the same emotion being induced in the observer. Nonetheless, the translation or "code" between expressed and perceived emotion in music is generally robust and accurate (Gabrielsson & Juslin, 2003).

Beginning with step 1, it is a long-held belief that composers embody their current and/or past emotional state in their works (Brown, 2000a; Cook & Dibben, 2001). In step 2, this emotion is expressed in the score by the composer using a set of structural music-emotion rules. These rules govern the relationships between musical features such as tempo, mode, pitch and harmonic complexity, and their emotional association (Gabrielsson & Lindstrom, 2001). For example, a major mode is commonly associated with happy, upbeat works.

Figure 2 illustrates structural rules identified by Livingstone and Brown (2005) in a meta-analysis of over 100 empirical studies (Gabrielsson & Juslin, 2003; Gabrielsson & Lindstrom, 2001; Schubert, 1999). Affect is represented in two dimensions: arousal (activity) and valence (positive/negative). The position of musical features, or cues, on each axis reflects the number of studies showing the association, where features are represented in emotion quadrants if the corresponding association was observed in multiple independent studies. The meta-analysis revealed a 93% level of agreement across studies for reported music feature – emotion association (music-emotion rules).

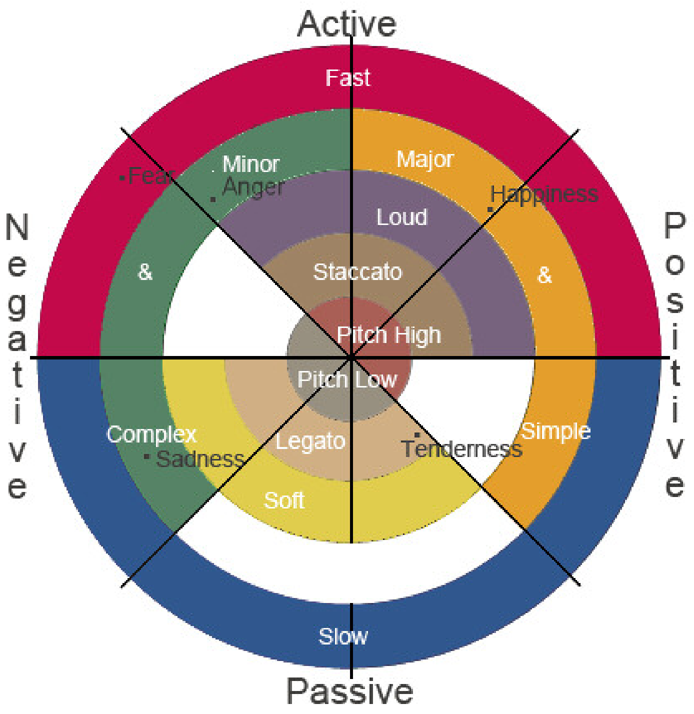

Figure 2. Meta-analysis of music attributes and commonly found emotions. Attributes near the periphery of the circle are based on the results of the greatest number of studies; those near the centre are based on the results of the fewest number of studies. As an example, happy upbeat music tends to have a faster tempo (red, 20 studies), be composed in a major mode (orange, 19 studies), have above average loudness (purple, 10 studies), staccato articulation (light brown, 5 studies) and above average pitch (clay, 4 studies). Adapted from (Livingstone & Thompson, 2006).

The use of these features as a means of affective communication is represented by their reflective symmetry. For example, a positively valenced work typically has a major mode and simple harmonies (orange), while a negatively valenced work has a minor mode and complex harmonies (dark green). The emotional power of these structural rules has been demonstrated with a computational model, where the perceived emotional message of selected music works could be influenced towards any other emotion quadrant through modification of these rules with an accuracy of 63%

(Livingstone, 2008; Livingstone, Muhlberger, Brown, & Loch, 2007; Livingstone, Muhlberger, Brown, & Thompson, in review). For example, changing a musical work from the major mode to the minor mode, slowing the tempo and decreasing the overall dynamic level shifted the perceived emotional message of the music shifted from joy to sadness. The results demonstrated the power and simplicity of the rules used by composers when expressing different emotional intentions (Thompson & Robitaille, 1992).

Moving to step 3, there are many studies that emphasise the importance of interpreting emotions. Sloboda writes that "even inexperienced performers soon become aware that their performance must be expressive" (Sloboda, 1985, p. 82), and Rigg articulates the idea that a musical score is interpretable emotionally; "if the composer leaves no interpretation of his production, it is usually not long before one is invented" (Rigg, 1942) cited in (Juslin & Laukka, 2004, p. 218). A survey of 15 performers revealed that 11 used a form of emotion-memory association when learning or conceptualising how a piece of music should be performed (Persson, 2001). That is, when learning the score, performers actively interpret the embodied emotional message, storing it as an emotional memory.

Step 4 sees a recurrence of induced emotion, however the effect here relates to self-inducement by the performer prior to playing. It has long been suspected that for a performer to efficaciously convey an emotion they must currently be experiencing that or a similar emotion (Juslin & Laukka, 2004). Further work by Persson (2001) revealed that all 15 of the previously discussed respondents reported using a form of self-inducement of the kind of emotion they wished to express prior to performance.

In step 5 the performer expresses their emotional interpretation of the musical work through a set of performance music-emotion rules (Gabrielsson, 1999, 2003; Juslin & Sloboda, 2001); these rules are similar to those used by the composer in step 2. Like composers, performers also use similar sets of rules when communicating differing emotional intentions (Gabrielsson, 1994, 1995). The power of these rules was recently demonstrated with a computational model that utilised both structural and performance rules. The system was highly successful in changing the emotion of selected music works to any of the emotional quadrants with an accuracy of 78%. That is, changing a happy sounding Mozart piano sonata into a sad work, an angry work, or a tender work (Livingstone, 2008; Livingstone, Muhlberger, Brown, & Thompson, in review). While ethnomusicologists have identified only a limited set of potential universals in music, an increasing number of studies have highlighted cross-cultural similarities in the low-level cues (tempo, dynamics, timbre) used by musicians to express particular emotions. These studies have considered the music of India (Balkwill & Thompson, 1999), Japan (Ohgushi & Hattori, 1996), Italy (Baroni & Finarelli, 1994), Germany (Langeheinecke, Schnitzler, Hischer-Buhrmeister, & Behne, 1999), Sweden (Juslin, 1997) and Russia (Kotlyar & Morozov, 1976); and instrumental performances of Japanese, Western and Hindustani music (Balkwill, Thompson, & Matsunaga, 2004). This body of research further suggests an underlying universal code for emotion communication that are intertwined with culture-dependant cues (Becker, 2001; Gabrielsson & Juslin, 2003), as has been observed with speech (Scherer, Johnstone, & Klasmeyer, 2003; Thompson & Balkwill, 2006). In addition to audio, performers leverage a variety of channels in the communication of emotion through music, discussed in the next section.

Step 6 occurs when a listener is exposed to a musical stimulus and perceives the emotion being expressed. The music at this point is a combination of the composer's original score and the performer's interpretation of it. It has been noted that the translation accuracy between the expression and perception of emotion, seen in steps 2-3 and 5-6, is typically accurate and robust. This accuracy suggests that music functions as an emotional communication mechanism (Juslin, 2001; Korhonen, 2004). This function is not unique to music, with high levels of decoding accuracy also demonstrated in Ekman's (1972, 1973) facial affect programs and Levenson *et al.*'s (1992) investigation of autonomic nervous system responses. While not marked in the framework, the act of performing music itself may also induce emotion in the performer, or act as a form of catharsis (Persson, 2001).

The Music-Emotion Framework concludes with step 7 in which emotion is induced in the listener in response to the musical stimuli. Since the time of Aristotle and Plato it has been reported that music elicits emotion in its listeners. Physiological studies support the inducement of emotion by music (Bartlett, 1996; Gabrielsson, 2001; Hodges, 1996; Panksepp, 1995; Rickard, 2004; VanderArk & Ely, 1992, 1993), along with profound effects on levels of relaxation and contentment (Gupta & Gupta, 2005). Mood also has been shown to affect memory of music based on the emotion being expressed by the music (Houston & Haddock, 2007). Musical preferences or social attitudes are also known to affect a listener's emotional response (Livingstone, 2008; Livingstone, Brown, & Muhlberger, 2005).

The framework described in this section illustrates how one system of music - the Western tonal idiom - instantiates in a very specific way the universal function of affective engagement. That is, Western classical music uses culture-specific structures and mechanisms to accomplish a broad human function that transcends cultural boundaries. Furthermore, other culture-specific mechanisms operate to achieve the same affective engagement across the arts. To support these hypotheses, we next review evidence from cross-domain studies of emotion in music and the arts.

Cross-domain studies of affect

As noted by Cross (2003a), Dissanayake (2000b) and Huron (2006), music shares a great deal with other art forms. We propose that music is non-specific in its ability and motivation to communicate emotion, leveraging a subset of multimodal channels employed across the arts and speech. Table 2 outlines a set of cross-cultural artistic forms and the channels they leverage in the expression and elicitation of emotion.

Table 2. Channels used in the expression and elicitation of emotion in the arts. ● represents common usage, • is less common usage, while T represents a transitive relation. For example, a painting may depict a body or facial expression, which itself is the channel. Visual channel refers to visual imagery elements, or objects used for visual effect.

	Form				
Channel	Music	Dance	Painting / Sculpture	Story Telling / Acting	Clothing & adornments
Body	●	●	T	●	
Facial	●	●	T	●	T
Visual	•	•	●	●	●
Tactile	•	●	●		•
Rhythmic	●	●			
Sound	●	●	•	●	•

As we will discuss, there is an emerging body of evidence for the affective significance of a multimodal display of music that incorporates body gestures, facial signals, tactile response and kinaesthetic elements in addition to the acoustic signal. Many art forms share channels with music, most strikingly dance. In many cultures dance is a strongly multimodal display, such as the Māori of New Zealand's Haka dance, which incorporates strong body, facial, rhythmic, sound (chant & music) and visual elements (Youngerman, 1974). Painting and sculpture often depict human subjects that communicate emotion with expressive body, facial and visual cues. In addition, some forms of cave art are also suspected to have employed acoustical effects (Goldhahn, 2002; Waller, 1993). Story telling (oral tradition) and acting often involve the communicator to be gesturally and facially expressive, often incorporating visual props, sound aids and music (Pellowski, 1977). *Wayang Purwa*, the traditional shadow play of the Javanese, thought to have originated from the Neolithic practice of ancestor worship, incorporates all of these elements (Ulbricht, 1970). Finally, clothing and adornments are primarily composed of visual elements and tactile materials that sometimes incorporate adornments displaying facial or gestural expressions. Masks, facial markings and body dress have all been used to elicit an emotional response, often for fear or intimidation.

The overlap observed in the arts and music extends beyond channel usage to include emotion types and methods of expression. Krumhansl and Scheck (1997) exposed three groups of participants to different versions of a choreographed musical work: dance only, music only, dance and music. Participants judged the similarity of the emotion expressed in each condition, with results indicating that the choreography and musical component were capable of communicating the same intended emotions. The capability of observers to correctly decode intended emotions in dance has also been confirmed in experiments that involved exposing observers to four versions of the same work, expressing one of anger, fear, grief, or joy (Camurri, Lagerlöf & Volpe, 2003; Lagerlöf & Djerf, 2002).

A connection between the emotion expressible by music performance gestures and accompanying audio has been similarly implicated. In Vines *et al.* (2004), 30 musically-trained subjects either saw, heard, or both saw and heard a performed classical music work. Results indicated that the visual stream carried much of the same structural information (that related to the music score) as the audio. Further, that performance gestures reflected musical phrasing and increased the experience of musical tension - a core property of musical emotionality (Lerdahl & Krumhansl, 2007; Meyer, 1956). These results later led the authors to state "there may exist an emergent quality when musical performances are both seen and heard" (Vines *et al.*, 2006 p. 80). Additional studies have yielded similar results (Dahl & Friberg, 2004; Davidson, 1993, 1994, 1995; Vines *et al.*, 2005; for further review, see Seitz, 2005).

A series of studies by Thompson *et al.* (2005) found that facial expressions of performers also carry a variety of structural and emotional cues for observers, both in vocal and instrumental music. Listener ratings of dissonance and interval sizes increased when coupled with facial expressions associated with dissonant points in the music or with large sung intervals, than when coupled with neutral or incongruent expressions. Visual imagery was also found to increase affect ratings of melodic intervals when paired with congruent audio samples. The presence of visual imagery was also found to significantly alter the overall valence evaluation of a sung musical work.

The association between music and visual imagery also has a long history. In recent times this link has been employed to great effect in cinema (Cohen, 2001). Music's close integration with a film's storyline was demonstrated by Vitouch (2001). Participants were presented a brief film excerpt with one of two musical soundtracks, original or fake, and asked to predict the plot continuation; the affective qualities of the music significantly influenced the predicted development of later scenes. Time synchronisation of audio-visual components is also important. Lipscomb (2005) has reported that a higher degree of synchronisation between the audio and visual components resulted in increased levels of emotional effectiveness.

Other research illustrates the connection between music and tactile or kinaesthetic responses. One study reported that infants are better able to encode rhythm when it is accompanied by congruent body movement (Phillips-Silver & Trainor, 2005). This finding may implicate an association between rhythm and the vestibular system, a connection also suggested by Truslit in 1938 (cited in Repp, 1993). A kinaesthetic relation with music is also exhibited by performers in the communication of emotion. The performance of rubato and dynamics has long been suggested to mimic physical motion. The growth in computational models of music performance has provided support for this hypothesis, demonstrating a strong mathematical correlation (Friberg & Sundberg, 1999; Kronman & Sundberg, 1987; Repp, 1992; Todd, 1992; Widmer & Tobudic, 2003; for additional music-motion discussion see Jackendoff & Lerdahl, 2006).

Affective mechanisms associated with music also overlap with those of speech prosody. Juslin and Laukka's (2003) meta-analysis of 104 studies on vocal emotion and 41 on music performance emotion revealed a strong similarity between the two in the communication accuracy and the rules (cues) used to express particular emotions. When these rules were coupled with the aforementioned structural music-emotion rule set, a 96% degree of 3-way rule agreement in the rules used to express emotion in speech, music performance and music structure was found (Livingstone, 2008). Similarities between music and speech prosody have also been observed at the level of interval prevalence. Ilie and Thompson's (2006) investigation of music and speech cues supported those of Juslin and Laukka (2003), while stressing the need for discrimination between energetic and tense aspects of arousal.

The connection between music and speech may run deeper than their reliance on a common emotional code (Patel, 2008). Huron (2006) has argued that musical expectations are learned and culture-specific, and that the underlying learning mechanisms also operate in speech and other domains. Saffran *et al.* (1996) found that 8-month old infants use statistical learning through exposure to determine appropriate word boundaries based on transitional probabilities, and that this process also operated for non-linguistic audio streams (3-note sequences) in both adults and infants (Saffran *et al.*, 1999). Both music and language also employ a series of similar devices to guide listener syntactic, semantic (language only) and emotional expectations (Astésano *et al.*, 2004; Kerkhofs *et al.*, 2007; Koelsch *et al.*, 2003; Kotz & Paulmann, 2007; Magne *et al.*, 2005; see also Besson & Schön, 2001). At the syntactic level, language also incorporates fluent ("the") and disfluent ("um", "uh") articles to guide expectations regarding the ongoing discussion of objects, both past and future (Arnold *et al.*, 2004). Finally, Vos and Troost (1989) examined the intonation behaviour of Albanian, Bulgarian, Iberian, Irish, Macedonian, Norwegian and African-American folk sounds and found that small intervals (those of 4 semitones or less) exhibit "tumbling melodies" (Sachs & Kunst, 1962), which as noted by Huron (2006) has long been known to linguists as "declination" (Cohen & t' Hart, 1967).

While story telling, painting and adornments also act in referential/informational capacity that leverage multimodal channels, the strong overlap between music and dance (see music definition) suggest their primary purpose is as affective engagement mechanisms. While we have outlined the affective basis of music and its broader connection with the arts, this alone does not account for why music appeared and remains universal across all human cultures. We next discuss the emergence of Theory of Mind in human evolution, and how it enabled the usage of a wide range of pre-existing cognitive faculties, and its importance for the creation and application of music in the construction of mental models of conspecific emotion.

The Origins of Music and a Theory of Mind

It has often been suggested that music and the arts appeared in Homo sapiens concurrently with a variety of other cultural practices including language (Henshilwood *et al.*, 2002; Mithen, 1996; Morley, 2002), with Huron (2001) and Fitch (2006) conservatively placing the emergence of music between 40,000 to 150,000 years ago. More specifically, it has been suggested that cognitive advancements such as music, language, culture, advanced tool use, symbolic art, social organisations and hierarchical thought appeared as a result of an increasing number of sophisticated innate, domain-specific cognitive faculties. An alternative view, however, is that a single cognitive faculty known as a ToM was responsible for the emergence of all of these activities.

Astington and Baird (2005, p. 3) define ToM as the ability to understand "people as mental beings who have beliefs, desires, emotions, and intentions and whose actions and interactions can be interpreted and explained by taking account of these mental states". This understanding of other beings enables an individual to imagine themselves “in the mental shoes” of the instructor. This cognitive advancement is thought to be responsible for the efficient transmission of knowledge and technology, enabling cumulative cultural evolution, or the *cultural ratchet effect* (Tomasello, 1999).

Two perspectives on ToM - *Theory Theory* and *Simulation Theory* – offer alternative explanations for the emergence of music through the use of pre-existing cognitive faculties. *Theory Theory* suggests that humans construct a model of the mental states of conspecifics in order to predict and explain their actions. *Simulation Theory* suggests that the *mirror neuron system* is engaged in order to simulate a hypothetical state of mind, providing a basis for empathy. Both viewpoints are compatible with the possibility that music is an instance of affective engagement, allowing the construction of mental models of the emotional states of conspecifics. That is, music and other arts act as an affective sandbox for safe emotional exploration, an ontogenetic concept that has its roots in Aristotle’s “paradox of tragedy” in *Poetics* (Halliwell, 1987); famously reformulated by Hume (1987). At the cultural level, music’s association with symbolic rituals and its use in social identity was enabled by higher-order ToM, further contributing to its continued use and dissemination. We will now review evidence to support these proposals.

A Theory of Mind

The appearance of Homo sapiens 200,000 to 250,000 years ago marked a significant shift in human evolution, heralding an explosion in technical, social and cultural phenomena. Their arrival is thought to coincide with the complete maturation of a ToM (Burns, 2004), as evidenced by the appearance of the earliest forms of fictional art and symbolic adornments during the Middle and Upper Palaeolithic periods (Baron-Cohen, 1999; Henshilwood *et al.*, 2002; Kunej & Turk, 2000; Mithen, 1996).

Along with cultural artefacts, Homo sapiens were also characterised by their use of advanced stone tools, symbolic communication and social structures. It has been argued that these advancements were enabled by the evolution of late-appearing mental modules (Chomsky, 1975; Fodor, 1983; Peretz & Morais, 1989) or the integration of such modules (Mithen, 1996). But it is problematic to assume that modularity is a genetically-encoded outcome of evolution.

The six million years that separate humans from the great apes is a very short period in evolutionary terms, with Homo sapiens sharing approximately 99% of their DNA with chimpanzees. As Tomasello argues, "there simply has not been enough time for normal biological evolution involving genetic variation and natural selection to have created, one by one, each of the cognitive skills necessary for modern humans" (1999, p. 2). This issue of time led to the proposition of a cultural ratchet effect enabled by ToM, whereby modern humans progressively made modifications or improvements to tools or practises, faithfully passing on these developments through social transmission. Operating on time scales orders of magnitude faster than genetic evolution, cumulative cultural evolution allows individuals to leverage previously developed skills and knowledge (Tomasello *et al.*, 1993). The ability to understand people as mental beings enables an individual to imagine themselves “in the mental shoes” of another person. As we will discuss, this capacity has important implications for music at the cultural level.

Mirror Neurons and Empathy in Musical Emotion

Two contrasting hypotheses have been proposed in the explanation of ToM. The first of these, known as Theory Theory (Gopnik, 1993; Gopnik & Wellman, 1992), proposes that humans can understand the behaviours of a conspecific through the application of higher-order theories about the mind (folk psychology). From these, we construct the mental state of a conspecific to predict and explain their actions. Theories of other minds can arise in either of two ways. The first route, called the ‘child as scientist’ explanation, is an ontogenetic selective process whereby developing individuals test theories empirically and discard those that are unsuccessful (Gopnik & Meltzoff, 1997; Klahr, 2000). The second route is through the phylogenetic result of pre-existing domain-specific innate modules (Baron-Cohen, 1995; U. Frith & Frith, 2003; Leslie, 1987, 1994).

The rival hypothesis, known as Simulation Theory, posits that there is no need to construct an abstract model of other minds, because the observer uses their own mind as the model of a conspecific - to be in the “mental shoes” of another (Goldman, 1989; Gordon, 1986; Harris, 1992; Nichols, Stich, Leslie, & Klein, 1996). Simulation theory has received tangible support from neuroscience literature relating to the discovery of mirror neurons. Located in the ventral premotor cortex (area F5) (di Pellegrino *et al.*, 1992; Ferrari *et al.*, 2003; Gallese *et al.*, 1996; Rizzolatti *et al.*, 1996) and posterior parietal cortex (Fogassi *et al.*, 2005; Gallese *et al.*, 2002; Rizzolatti *et al.*, 2001) of the macaque monkey, mirror neurons are also suspected to exist in the human brain (for review, see

Gallese, 2006). Mirror neurons were classified as those that were shown to discharge when a monkey both executes or observes an action (Gallese *et al.*, 1996), with a subset discharging when the final part of the action is hidden, indicating goal inference (Umiltà *et al.*, 2001; Fogassi *et al.*, 2005). Some mirror neurons have also been shown to respond equally whether an action is seen, seen and heard, or only heard (Keysers *et al.*, 2003; Kohler *et al.*, 2002). Gallese proposes that this direct-matching property of mirror neurons underpins Simulation Theory (Gallese, 2006, 2007; Gallese & Goldman, 1998), a position we adopt.

Mirror neurons also play a role in the *Perception Action Model* (PAM) of empathy (Preston & de Waal, 2002; Decety & Grèzes, 2006). PAM states than an observer will resonate with the perceived emotional state of a conspecific, activating the observer's corresponding representations and in turn their own somatic and autonomic responses. PAM also provides a framework for understanding the phenomenon of emotional contagion (Hatfield *et al.*, 1994), which has been widely cited in connection with musical affect (Jackendoff & Lerdahl, 2006; Juslin, 2001; Juslin & Laukka, 2003, 2004; Peretz, 2006; Scherer & Zentner, 2001). In PAM, emotional contagion is subsumed into a broader model of empathy.

One view is that both Theory Theory and Simulation Theory are in operation within humans (Meltzoff & Gopnik, 1993; Perner, 1996; Preston & de Waal, 2002). According to this hypothesis, Simulation Theory functions as a lower-level motor - audio (imitation) and empathy learning mechanism, which in turn assists in the formation of complex, higher-order models about the mind (Theory Theory). This notion finds support from recent evidence which found that the response of the human mirror system to previously observed motor events is not wholly innate or fixed, and can be modified through retraining (Catmur *et al.*, 2007; Press *et al.*, 2007). Such a system provides a flexible model of low-level sensorimotor learning. The nature of perspective taking observed in ToM is a higher-order process than the imitation of action and emotion seen in Simulation Theory, and involves both emotional and non-emotional tasks (Preston *et al.*, 2007). There is also growing neuroscientific evidence to support the existence of both mechanisms (Decety & Grèzes, 2006; C. D. Frith, 2007; U. Frith & Frith, 2003; Saxe & Kanwisher, 2004; Saxe & Powell, 2006); though some remain cautious (Apperly *et al.*, 2007; Apperly *et al.*, 2005).

It is this two-stage model of ToM that is of particular interest to our theory of music origins. We propose that both modes operate in the creation and propagation of music, and are intimately connected with two of music's universal tendencies – emotion and its use in cultural rituals. In particular, music appears to leverage empathy in the elicitation of emotion in listeners. The Perception Action Model of empathy underpins this notion and further supports a multimodal view of music. The observation of multiple channels through the audio-visual mirror neuron system results in a heightened emotional experience for the listener, due to the broadened activation of channels that are empathically matched or "mirrored"[iii] (see Table 2). Strong physiological reactions can occur in listening to music due to our "response *with*" mechanism (matching responses sad with sad), but they are tempered by our "response *to*" the performer (Preston & de Waal, 2002, p. 5) - a later developing, learnt prosocial behaviour which allows for the visible inhibition of both the automatic "response *with*" and typical "response *to*" patterns (for example, consolation for fearful or angry conspecifics).

An empathy-only model of music and emotion is overly simplistic however, because it fails to explain why negatively-valenced (*e.g.*, sad) music maintains popularity in musical repertoires. It is here that we can consider how a higher-order ToM might operate in music. The emotional experience generated by music permits a safe learning environment for affective engagement - the exchange of psychological state for the development of higher-order models of emotion regarding conspecifics. In this view, "a ToM is essentially a derivative of the Aristotelian position in which our experience with artworks is a form of affective sandboxing; a means of pursuing affective exploration / hypothesis testing in a safe environment" (Livingstone & Thompson, 2006, p. 91). Thus, while a listener may feel sad as a result of certain types of music, the experience is still a beneficial or positive one.

As Tomasello *et al.* (2005) discusses, the motivation for sharing of emotions is a crucial and unique component of human ToM. Evidence for ToM has been reported for a number of non-human animals, but humans appear to possess an advanced form of ToM that includes and motivates affective engagement. Both primates and corvids are suspected to possess an early form of ToM. Interestingly, however, while ToM is associated with the prefrontal cortex of primates (Call *et al.*, 2004; Povinelli *et al.*, 2003; Tomasello & Rakoczy, 2003), it is believed to lie in the forebrain of corvids (Bugnyar & Heinrich, 2005, 2006). Thus, despite the divergent brain evolution of the two species, ToM evolved across primates and corvids as a result of convergent mental evolution due to similar socioecological challenges (Emery & Clayton, 2004). As humans also faced similar challenges during their evolution, this provides further support for ToM as an adaptation.

In humans, traces of ToM are evident at early developmental stages. While higher-order ToM capabilities do not fully mature in children until the age of 4[iv], the perception-action model of empathy, along with cognitive faculties required for music processing (see Table 1) are functioning from infancy. Further, following the "child-as-scientist" model, the desire for the creation of higher-order models of conspecific emotion is an ongoing process from early-age. Together these elements can account for the appearance of musical behaviours in children below the age of 4, while prior to ToM maturation. In the next section we will discuss how a higher-order ToM also acts at the cultural level in music.

Music and Culture

A suspected universal feature of music is its use in symbolic ceremonies and cultural rituals (Nettl, 2000). Tomasello (1999) argues that a ToM is essential for symbolism, where symbolic representations are thought to arise through collaborative problem solving referred to as *joint-attentional activities* (Tomasello, 2003). In this process individuals engage in a form of *dialogic interaction* to produce a shared or joint understanding. However, music itself is not symbolic - it is not referential through joint understanding in the way that language functions. Nonetheless, music does have the capacity to be referential, and certain musical forms, styles and works can obtain an indexical relation (Turino, 1999). One modern case is Felix Mendelssohn's Wedding March, written for Shakespeare's "A Midsummer Night's Dream." The association between this music and the Western wedding ritual is an *indexical* relation, where the Wedding March indicates a wedding by virtue of their mutual association - their contiguity in time and place (Tolbert, 2001), and their co-occurrence with experience (Turino, 1999). Indexes lack the arbitrary association between sounds and objects that symbols possess. However, music universally accompanies symbolic events (Nettl, 2005), and the symbolic nature of these events is enabled by ToM. Many cultural rituals, such as burial, pair union, war and so on have continued for hundreds of thousands of years. The strong association between music and such rituals continues today and may illustrate how ToM assisted in the persistence, propagation and diversification of music at the cultural level.

As Nettl discusses (2005, pp. 254-58), a significant function of music revolves around the concept of identity. Hargreaves and North (1999, see also, Sloboda & O'Neill, 2001) suggest that music functions as a means of self-expression and identity. Adolescent males, in particular, often use music in the construction of their self-identity (North, Hargreaves, & O'Neill, 2000). The sense of identity, whether at an individual level or at the level of national identity, is a direct consequence of ToM – the ability to recognise a conspecific as having their own thoughts, beliefs and desires. Like the use of music in symbolic rituals, its use in self and group identity is enabled through ToM, contributing to its continuing presence in human culture.

Along with symbolism, another consequence of ToM is a process known as *reflective thought*. Through this capacity an individual can construct a statement or thought, and then make an evaluative statement about it (Lohmann, Tomasello, & Meyer, 2005); Mithen (1996) refers to this capacity as *reflexive consciousness*, while Gilardi *et al.* (2006) believe it is a form of epistemological thought. Reflective thought serves two functions[v]. The first is to facilitate the faithful transmission of cultural knowledge. For example, when a child expresses their view about a particular situation, it is common for an adult to critique this view. In the process, the child is forced re-examine their view by imagining the situation from the perspective of the adult. The continual refinement and progression of music and the arts may have continued over time through this process of reflective thought (Livingstone & Thompson, 2006). In addition, reflective thought allows a performer/composer to reflect back on themselves and consider what they are trying to express, how, and why. This ability allows the performer/composer to consider how to better communicate these thoughts and emotions, with additional contextual considerations.

The second function of reflective thought relates to individual skill development. Tomasello notes that after a person reaches a certain level of *skill mastery* on a particular task (*e.g.*, juggling), the individual starts to reflect on their reasons for success at the task, identifying properties of their approach that led to success. The result of this reflective thinking process may sequestrate features related to how the individual achieves their goal and the type of objects involved in the task. This process of abstracting *properties of success* serves as an excellent mechanism in the cultural transmission of knowledge, or cultural ratcheting. This reflexive and contemplative process permits the creator to make refinements and progressions based on their own personal aesthetic judgements, and those of others, over time. In the case of music, experiential knowledge regarding music performance, either vocally or instrumentally, could be better passed to conspecifics to enable a greater proficiency in music making.

Conclusion

There are a number of persuasive arguments for adaptationist accounts of music origins. However, strong claims of a music-specific genotype are premature (Huron, 2001; Justus & Hutsler, 2005; McDermott & Hauser, 2005). Singling out three universal aspects of music for investigation – its connection with affect, dance and use in cultural rituals - we proposed that a *Theory of Mind* provides a more compelling account for music and other arts.

Examining a broad range of psychophysiological cross-domain evidence, we established that music is non-specific in the arts for its uses, mechanisms and diversity in the expression, perception and elicitation of emotion. The evidence suggested that an underlying, cross-domain, universal method of emotional communication, referred to as affective engagement, was in operation. The broader connection with the arts also accounted for strong parallels between music and dance (Nettl, 2000, 2005).

While innate, domain-specific cognitive faculties have been proposed as an evolutionary explanation of the diversity of human thought and culture, this view suffers from the issue of time. Theory of Mind has been proposed as an alternative, and relies on the more rapid process of cumulative cultural evolution (Tomasello, 1999). Enabling the use of language and symbolism, Theory of Mind accounts for the significance of music at the cultural level, including its use in symbolic rituals (Nettl, 2000), in identity of self and ethnicity (Nettl, 2005), and its continuous growth in complexity and diversity.

Two elements of ToM – Simulation Theory and Theory Theory - may be viewed as complementary stages (Decety & Grèzes, 2006; C. D. Frith, 2007; U. Frith & Frith, 2003; Meltzoff & Gopnik, 1993; Preston & de Waal, 2002; Saxe & Kanwisher, 2004; Saxe & Powell, 2006). The involvement of mirror neurons in the Perception Action Model of empathy (subsumes emotional contagion), provided a strong empirical basis for the low-level emotional response to music and the arts. Broadened activation of channels to be empathically matched accounts for the multimodal nature of music and the arts, where evidence of increased emotional response to multimodal displays was frequently reported. At a higher level, the 'child-as-scientist' view of Theory Theory accounted for why humans engage in such behaviours, when direct adaptive functions of music-specific behaviours are lacking. Instead music is merely one example of a broader biological function of affective engagement.

References

Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W.-L., & Humphreys, G. W. (2007). Testing the domain-specificity of a theory of mindnext term deficit in brain-injured patients: Evidence for consistent performance on non-verbal, "reality-unknown" false belief and false photograph tasks. *Cognition, 103*, 300-21.

Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: evaluating neuropsychological evidence. *Trends in Cognitive Sciences, 9*, 572-77.

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science, 15*, 578-82.

Arom, S. (2000). Prologomena to a biomusicology. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 27-29). Cambridge, Massachusetts: MIT Press.

Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research, 18*, 172-84.

Astington, J. W., & Baird, J. A. (2005). Introduction: Why language matters. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 3-25). Oxford: Oxford University Press.

Ayotte, J., Peretz, I., & Hyde, K. L. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain, 125*, 238-51.

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*, 43-64.

Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research, 46*, 337-49.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S. (1999). The evolution of a theory of mind. In M. C. Corballis & S. E. G. Lea (Eds.), *The Descent of Mind: Psychological Perspectives on Hominid Evolution*. Oxford: Oxford University Press.

Baroni, M., & Finarelli, L. (1994). *Emotions in spoken language and vocal music*. Proceedings of Third International Conference for Music Perception and Cognition, Liège, Belgium.

Barrow, J. D. (1995). *The artful universe*. Oxford: Clarendon Press.

Bartlett, D. (1996). Physiological Responses to Music and Sound Stimuli. In D. A. Hodges (Ed.), *Handbook of music psychology* (2nd ed., pp. 343-85). Lawrence, KS: National Association for Music Therapy.

Becker, J. (2001). Anthropological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 135-60). Oxford: Oxford University Press.

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences, 930*, 232–58.

Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.

Brown, S. (2000a). Evolutionary models of music: From sexual selection to group selection. In N. S. Thompson & F. Tonneau (Eds.), *Perspectives in Ethology* (Vol. 13, pp. 221-81). New York: Plenum.

Brown, S. (2000b). The "musilanguage" model of music. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 271-300). Cambridge, Massachusetts: MIT Press.

Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind—evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews, 30*, 437-55.

Brust, J. C. (1980). Music and language: musical alexia and agraphia. *Brain, 103*, 367-92.

Bugnyar, T., & Heinrich, B. (2005). Ravens, Corvus corax, differentiate between knowledgeable and ignorant competitors. *Proceedings of The Royal Society B: Biological Sciences, 272*, 1641-46.

Bugnyar, T., & Heinrich, B. (2006). Pilfering ravens, Corvus corax, adjust their behaviour to social context and identity of competitors. *Animal Cognition, 9*, 369-76.

Burns, J. K. (2004). An evolutionary theory of schizophrenia: Cortical connectivity, metarepresentation and the social brain. *Behavioral and Brain Sciences, 27*, 831-55.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, Exaptations, and Spandrels. *American Psychologist, 53*, 533-48.

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). Unwilling’ versus ‘unable’: Chimpanzees' understanding of human intentional action. *Developmental Science, 7*, 488-98.

Camurri, A., Lagerlöf, I., & Volpe, G. (2003). Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies, 59*, 213-25.

Cappelletti, M., Waley-Cohen, H., Butterworth, B., & Kopelman, M. (2000). A selective loss of the ability to read and write music. *Neurocase, 6*, 332-41.

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology, 17*, 1527-31.

Chomsky, N. (1975). *Reflections on language*. New York: Pantheon

Cohen, A., & t' Hart, J. (1967). On the anatomy of intonation. *Lingua, 19*, 177-92.

Cohen, A. J. (2001). Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249-74). Oxford: Oxford University Press.

Cook, N., & Dibben, N. (2001). Musicological approaches to emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 45-70). Oxford: Oxford University Press.

Cross, I. (1999). Is music the most important thing we ever did? Music, development and evolution. In S. W. Yi (Ed.), *Music, mind and science* (pp. 10-39). Seoul: Seoul National University Press.

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences, 930*, 28-42.

Cross, I. (2003a). Music and Evolution: Consequences and Causes. *Contemporary Music Review, 22*, 79-89.

Cross, I. (2003b). Music, cognition, culture and evolution. In I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 42-56). Oxford: Oxford University Press.

Dahl, S., & Friberg, A. (2004). Expressiveness of musician's body movements in performances on marimba. In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in Human-Computer Interaction* (Vol. 2915/2004, pp. 479-86): Springer Berlin / Heidelberg.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21*, 103-13.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies, 26*, 279-301.

Davidson, J. W. (1995). What does the visual information contained in music performances offer the observer? Some preliminary thoughts. In R. Steinberg (Ed.), *The Music Machine: Psychophysiology and Psychopathology of the Sense of Music Perception* (pp. 105-13). Berlin: Springer Verlag.

Davies, S. (2003). *Themes in the philosophy of music*. Oxford: Oxford University Press.

Dean, R. T., Byron, T., & Bailes, F. A. (in press). The pulse of symmetry: On the possible co-evolution of rhythm in music and dance. *Musicae Scientiae, Special Issue on Music and Evolution*.

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research, 1079*, 4-14.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*, 176-80.

Dissanayake, E. (2000a). Antecedents of the temporal arts in early mother-infant interaction. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 389-410). Cambridge, Massachusetts: MIT Press.

Dissanayake, E. (2000b). *Art and intimacy: How the arts began*. Seattle: University of Washington Press.

Dunbar, R. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences, 16*, 681-735.

Dunbar, R. (1996). *Grooming, gossip and the evolution of language*. Cambridge, MA: Harvard University Press.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation*, 1971 (Vol. 19, pp. 207–82). Lincoln: University of Nebraska Press.

Ekman, P. (1973). *Darwin and facial expression: A century of research in review*. New York: Academic Press.

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science, 306*, 1903-07.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eurpoean Journal of Neuroscience, 15*, 399-402.

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences, 27*, 491-541.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eurpoean Journal of Neuroscience, 17*, 1703-14.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition, 100*, 173-215.

Fodor, J. (1983). *The Modularity of Mind*. The MIT Press.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 302*, 662-67.

Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion: A model of final ritardandi derived from measurements of stopping runners. *Journal of the Acoustical Society of America, 105*, 1469-84.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 362*, 671-78.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 358*, 459-73.

Gabrielsson, A. (1994). *Intention and emotional expression in music performance*. Paper presented at the Stockholm Music Acoustics Conference 1993, Stockholm.

Gabrielsson, A. (1995). Expressive intention and performance. In R. Steinberg (Ed.), *Music and the mind machine: The psychophysiology and the psychopathology of the sense of music* (pp. 35-47). Berlin: Springer Verlag.

Gabrielsson, A. (1999). The Performance of Music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed.) (pp. 501-602). San Diego: AP press.

Gabrielsson, A. (2001). Emotions in strong experiences with music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 431-49). Oxford: Oxford University Press.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special Issue 2001-2002*, 123-47.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 31*, 221-72.

Gabrielsson, A., & Juslin, P. N. (2003). Emotional Expression in Music. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 503-34). Oxford: Oxford University Press.

Gabrielsson, A., & Lindstrom, E. (2001). The Influence of Musical Structure on Emotional Expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 223-48). Oxford: Oxford University Press.

Gallese, V. (2006). Intentional attunement: A neurophysiological perspective on social cognition and its disruption in autism. *Brain Research, 1079*, 15-24.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 362*, 659-69.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593-609.

Gallese, V., Fogassi, L., Fadiga, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B.Hommel (Eds.), *Attention and performance XIX* (pp. 247-66). Oxford, UK: Oxford University Press.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*, 493-501.

Gilardi, S., Bruno, A., & Pezzotta, C. (2006). Discursive practices and mentalization ability in adults at work. In A. Antonietti, O. Liverta-Sempio & A. Marchetti (Eds.), *Theory of mind and language in developmental context* (pp. 173-92). New York: Springer.

Goldhahn, J. (2002). Roaring rocks: An audio-visual perspective on hunter-gatherer engravings in northern Sweden and Scandinavia. *Norwegian Archaeological Review, 35*, 29-61.

Goldman, A. (1989). *Interpretation psychologized. Mind and Language, 4*, 161-85.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences, 16*, 1-14.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language, 7*, 145-71.

Gordon, R. (1986). Folk psychology as simulation. *Mind and Language, 1*, 158-71.

Gould, S. J., & Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of The Royal Society of London B: Biological Sciences, 205*, 581-98.

Gupta, U., & Gupta, B. S. (2005). Psychophysiological responsivity to Indian instrumental music. *Psychology of Music, 33*, 363-72.

Halliwell, S. (1987). *The poetics of Aristotle: Translation and commentary*. London: Duckworth.

Hargreaves, D. J., & North, A. (1999). The functions of music in everyday life: Redefining the social in music psychology. *Psychology of Music, 27*, 71-83.

Harris, P. L. (1992). From simulation to folk psychology: the case for development. *Mind and Language, 7*, 120-44.

Hatfield, E., Cacioppo, J., & Rapson, R. L. (1994). *Emotional contagion*. Cambridge: Cambridge University Press.

Haueisen, J., & Knoesche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience, 13*, 786-92

Hébert, S., & Cuddy, L. L. (2006). Music-reading deficiencies and the brain. *Advances in Cognitive Psychology, 2*, 199-206.

Henshilwood, C. S., D'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G. A. T., et al. (2002). Emergence of modern human behaviour: Middle Stone Age engravings from South Africa. *Science, 295*, 1278-80.

Heyes, C. (2003). Four routes of cognitive evolution. *Psychological Review, 110*, 713-27.

Hodges, D. A. (1996). Neuromusical Research: A Review of the Literature. In D. A. Hodges (Ed.), *Handbook of Music Psychology* (pp. 197-284). Lawrence, KS: National Association for Music Therapy.

Houston, D., & Haddock, G. (2007). On auditing auditory information: the influence of mood on memory for music. *Psychology of Music, 35*, 201-12.

Hume, D. (1987). Of tragedy. In E. F. Miller (Ed.), *Essays: Moral, political, and literary* (pp. 216-25). Indianapolis: Liberty Classics.

Huron, D. (2001). Is Music an Evolutionary Adaptation? *Annals of the New York Academy of Sciences, 930*, 43-61.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge: The MIT Press.

Ilie, G. & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception, 23*, 319-29.

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

Jones, S. S. (2007). Imitation in infancy: The development of mimicry. *Psychological Science, 18*, 593-99.

Juslin, P. N. (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14*, 383-418.

Juslin, P. N. (2001). Communicating emotion in music performance: a review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249-74). Oxford: Oxford University Press.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770-814.

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research, 33*, 217-38.

Juslin, P. N, & Sloboda, J. A. (Eds.). *Music and Emotion: Theory and Research.* Oxford: Oxford University Press.

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

Keltner, D., & Ekman, P. (2003). Introduction: Expression of Emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences* (pp. 411-14). Oxford: Oxford University Press.

Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. J. (2007). Discourse, syntax, and prosody: The brain reveals an immediate interaction. *Journal of Cognitive Neuroscience, 19*, 1421-34.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research, 153*, 628-36.

Kivy, P. (1990). *Music alone: Philosophical reflections on the purely musical experience*. Ithaca, New York: Cornell University Press.

Kivy, P. (2002). *Introduction to a Philosophy of Music*. USA: Oxford University Press.

Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge: MIT Press.

Koelsch, S., Gunter, T. C., Cramon, D. Y. V., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical ''language-network'' serves the processing of music. *NeuroImage, 17*, 956-66.

Koelsch, S., Gunter, T. C., Schröger, E., & Friederici, A. D. (2003). Processing tonal modulations: An ERP study. *Journal of Cognitive Neuroscience, 15*, 1149-59.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T. C., & Friederici, A. D. (2004). Music, language, and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7*, 511-14.

Kogan, N. (1997). Reflections on aesthetics and evolution. *Critical Review, 11*, 193-210.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*, 846-48.

Korhonen, M. D. (2004). *Modeling continuous emotional appraisals of music using system identification*. Unpublished Masters, University of Waterloo.

Kotlyar, G. M., & Morozov, V. P. (1976). Acoustic correlates of the emotional content of vocalized speech. Soviet Physics. *Acoustics, 22*, 370-76.

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research, 1157*, 107-18.

Kronman, U., & Sundberg, J. (1987). Is the musical retard an allusion to physical motion? In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music* (Vol. 55, pp. 57-68). Stockholm, Sweden: Royal Swedish Academy of Music.

Krumhansl, C. L., & Schenck, D. L. (1997). Can dance reflect the structural and expressive qualities of music? A perceptual experiment on Balanchine's choreography of Mozart's Divertimento No. 15. *Musicae Scientiae, 1*, 63–85.

Kugiumutzakis, G. (1993). Intersubjective vocal imitation in early mother–infant interaction. In J. Nadel & L. Camioni (Eds.), *New perspectives in early communicative development* (pp. 23-47). London: Routledge.

Kugiumutzakis, G. (1998). Neonatal imitation in the intersubjective companion space. In S. Braten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 63–88). Cambridge, UK: Cambridge University Press.

Kunej, D., & Turk, I. (2000). New Perspectives on the Beginnings of Music: Archeological and Musicological Analysis of a Middle Paleolithic Bone "Flute". In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 235-68). Cambridge, Massachusetts: MIT Press.

Lagerlöf, I., & Djerf, M. (2002). *On cue utilization for emotion expression in dance movements*, Manuscript in preparation, Department of Psychology, University of Uppsala.

Langeheinecke, E. J., Schnitzler, H.-U., Hischer-Buhrmeister, M., & Behne, K.-E. (1999). *Emotions in singing voice: Acoustic cues for joy, fear, anger, and sadness*. Poster presented at the Joint Meeting of the Acoustical Society of America and the Acoustical Society of Germany, Berlin, March 1999.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: The MIT Press.

Lerdahl, F., & Krumhansl, C. L. (2007). Modelling Tonal Tension. *Music Perception, 24*, 329-66.

Leslie, A. (1987). Pretence and representation. The origins of 'theory of mind'. *Psychological Review, 94*, 412-26.

Leslie, A. (1994). A theory of ToMM, ToBy, and Agency: core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-48). New York: Cambridge University Press.

Levenson, R. W., Ekman, P., Heider, K., & Friesen, W. V. (1992). Emotion and autonomic nervous system activity in the Minangkabau of West Sumatra. *Journal of Personality and Social Psychology, 62*, 972-88.

Lipscomb, S. (2005). The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology, 12*, 37-67.

Livingstone, S. R. (2008). *Changing musical emotion through score and performance with a computational rule system*. Unpublished Doctoral Thesis, The University of Queensland, Brisbane. Submitted February 4 (under review).

Livingstone, S. R., & Brown, A. R. (2005). Dynamic response: Real-time adaptation for music emotion. *ACM International Conference Proceeding Series, 123*, 105–11.

Livingstone, S. R., Brown, A. R., & Muhlberger, R. (2005). *Influencing the Perceived Emotions of Music with Intent*. Paper presented at the Third International Conference on Generative Systems.

Livingstone, S. R., Muhlberger, R., Brown, A. R., & Loch, A. (2007). Controlling musical emotionality: An affective computational architecture for influencing musical emotions. *Digital Creativity, 18*, 43-54.

Livingstone, S. R., Muhlberger, R., Brown, A. R., & Thompson, W. F. (in review). *Changing musical emotion through score and performance with a computational rule system.* 10th International Conference of Music Perception and Cognition, Sapporo, Japan.

Livingstone, S. R., & Thompson, W. F. (2006). Multi-modal affective interaction: A comment on musical origins. *Music Perception, 24*, 89-94.

Lohmann, H., Tomasello, M., & Meyer, S. (2005). Linguistic communication and social understanding. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 245-65). Oxford: Oxford University Press.

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., & Besson, M. (2005). On-line processing of ''Pop-Out'' words in spoken French dialogues. *Journal of Cognitive Neuroscience, 17*, 740-56.

McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, Uniqueness, and Evolution. *Music Perception, 23*, 29-59.

Meltzoff, A. N., & Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds, perspectives from autism*. Oxford: Oxford University Press.

Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae, 6*, 3-21.

Meyer, L. B. (1956). *Emotion and Meaning in Music*: The University of Chicago Press.

Miller, G. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 329-60). Cambridge, Massachusetts: MIT Press.

Mithen, S. (1996). *The Prehistory of the Mind*. London: Thames and Hudson.

Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: from motion to 'e'motion. *Social Cognitive and Affective Neuroscience, 1*, 235-41.

Morley, I. (2002). Evolution of the Physiological and Neurological Capacities for Music. *Cambridge Archaeological Journal, 12*, 195-216.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 463-72). Cambridge, Massachusetts: MIT Press.

Nettl, B. (2005). *The study of ethnomusicology: Thirty-one issues and concepts* (2nd ed.). Champaign, IL: University of Illinois Press.

Nichols, S., Stich, S., Leslie, A., & Klein, D. (1996). Varieties of off-line simulation. In P. Carruthers & P. K. Smith (Eds.), *Theories of Theories of Mind* (pp. 39-70). Cambridge: Cambridge University Press.

Norell, M., Ji, Q., Gao, K., Yuan, C., Zhao, Y., & Wang, L. (2002). Palaeontology: 'modern' feathers on a non-avian dinosaur. *Nature, 416*, 36-37.

North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology, 70*, 255-72.

Ohgushi, K., & Hattori, M. (1996). *Acoustic correlates of the emotional expression in vocal performance*. Paper presented at the Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, 2-6 December 1996.

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception, 13*, 171-207.

Papousek, H. (1996). Musicality in infancy research: biological and cultural origins of early musicality. In I. Deliège & J. A. Sloboda (Eds.), *Musical Beginnings*. (pp. 88-112). Oxford: Oxford University Press.

Parncutt, R. (In press-a). Prenatal and infant conditioning, the mother schema, and the origins of music and religion. *Musicae Scientiae, Special Issue on Music and Evolution*.

Parncutt, R. (In press-b). Prenatal development. In S. Hallman, I. Cross & M. Thaut (Eds.), *Oxford handbook of music psychology*. Oxford: Oxford University Press.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674-81.

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception, 24*, 99-104.

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

Pellowski, A. (1977). *The world of storytelling*. New York: Bowker.

Peretz, I. (2001a). Brain specialization for music: New evidence from congenital amusia. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (pp. 153-65). New York: The biological foundations of music.

Peretz, I. (2001b). Music perception and recognition. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. 519-40). Philadelphia: Psychology Press.

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

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., et al. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33*, 185-91.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences, 7*, 362-67.

Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary Music Review, 4*, 279-93.

Peretz, I., & Zatorre, R. J. (Eds.). (2003). *The cognitive neuroscience of music*. Oxford: Oxford University Press.

Perner, J. (1996). Simulation as explicitation of prediction-implicit knowledge about the mind: arguments for a simulation-theory mix. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 90-104). Cambridge: Cambridge University Press.

Persson, R. S. (2001). The subjective world of the performer. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 275-90). Oxford: Oxford University Press.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308*, 1430.

Pinker, S. (1999). *How the mind works*. New York: Norton.

Povinelli, D. J., Theall, L. A., Reaux, J. E., & Dunphy-Lelii, S. (2003). Chimpanzees spontaneously alter the location of their gestures to match the attentional orientation of others. *Animal Behaviour, 66*, 71-79.

Pratt, G. (1990). *Aural awareness: Principles and practice*. Milton Keynes: Open University Press.

Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of The Royal Society B: Biological Sciences, 274*, 2509-14.

Preston, S. D., Bechara, A., Damasio, H., Grabowski, T. J., Stansfield, R. B., Mehta, S., et al. (2007). The neural substrates of cognitive empathy. *Social Neuroscience, 2*, 254-75.

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1-72.

Regal, P. J. (1975). The evolutionary origin of feathers. *Quarterly Review of Biology, 50*, 35–66.

Repp, B. H. (1992). Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's "Träumerei". *Journal of the Acoustical Society of America, 92*, 2546-68.

Repp, B. H. (1993). Music as motion: A synopsis of Alexander Truslit's (1938) *Gestaltung und bewegung in der musik*. *Psychology of Music, 21*, 48-72.

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music, 32*, 371-88.

Rigg, M. G. (1942). The expression of meanings and emotions in music. In F. P. Clarke (Ed.), *Philosophical essays in honor of Edgar Arthur Singer* (pp. 279-94). Oxford: University of Pennsylvania Press.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131-41.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Neuroscience Review, 2*, 661-70.

Roederer, J. (1984). The search for a survival value of music. *Music Perception, 1*, 350-56.

Sachs, C., & Kunst, J. (1962). *The wellsprings of music*. In J. Kunst (Ed.), The Hague: Marinus Nijhoff.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-28.

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*, 27-52.

Saxe, R., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 55*, 87-124.

Saxe, R., & Powell, L. J. (2006). It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind. *Psychological Science, 17*, 692-99.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33*, 239-51.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 433-56). Oxford: Oxford University Press.

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 361-92). Oxford: Oxford University Press.

Schubert, E. (1999). *Measurement and Time Series Analysis of Emotion in Music*. University of New South Wales.

Seitz, J. A. (2005). Dalcroze, the body, movement and musicality. *Psychology of Music, 33*, 419-35.

Sloboda, J. A. (1985). *The musical mind: the cognitive psychology of music*. Oxford: Clarendon Press.

Sloboda, J. A., & O'Neill, S. A. (2001). Emotions in Everyday Listening to Music. In *Music and Emotion, theory and research* (pp. 415-29): Oxford Press.

Sperber, D. (1996). *Explaining Culture*. Oxford: Blackwell.

Studdert-Kennedy, M. (1983). On learning to speak. *Human Neurobiology, 2*, 191-95.

Thompson, W. F., & Balkwill, L.-L. (2006). Decoding speech prosody in five languages. *Semiotica, 158*, 407-24.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica, 156*, 203-27.

Thompson, W.F. & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts, 10*, 79-89.

Todd, N. P. M. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America, 91*, 3540-50.

Tolbert, E. (2001). Music and meaning: An evolutionary story. *Psychology of Music, 29*, 84-94.

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development, 10*, 131-56.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press

Tomasello, M. (2003). The key is social cognition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 47-58). London: The MIT Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675-91.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16*, 495-552.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language, 18*, 121-47.

Trehub, S. E. (2001). Musical Predispositions in Infancy. *Annals of the New York Academy of Sciences, 930*, 1-16.

Trehub, S. E. (2003). Toward a developmental psychology of music. *Annals of the New York Academy of Sciences, 999*, 402-13.

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*, 73-99.

Truslit, A. (1938). *Gestaltung und Bewegung in der Musik* [Shaping and motion in music]. Berlin-Lichterfelde: Chr. Friedrich Vieweg.

Turino, T. (1999). Signs of imagination, identity, and experience: A Peircian semiotic theory for music. *Ethnomusicology, 42*, 221-55.

Ulbricht, H. (1970). *Wayang Purwa: Shadows of the past*. Kuala Lumpur: Oxford University Press.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing: a neurophysiological study. *Neuron, 31*, 155-65.

VanderArk, S. D., & Ely, D. (1992). Biochemical and galvanic skin responses to music stimuli by college students in biology and music. *Perceptual and Motor Skills, 74*, 1079-90.

VanderArk, S. D., & Ely, D. (1993). Cortisol, biochemical, and galvanic skin responses to music stimuli of different preference values by college students in biology and music. *Perceptual and Motor Skills, 77*, 227-34.

Vines, B. W., Krumhansl, C., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2005). Dimensions of emotion in expressive musical performance. *Annals of the New York Academy of Sciences, 1060*, 462-66.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition, 101*, 80-113.

Vines, B. W., Wanderley, M. M., Krumhansl, C. L., Nuzzo, R. L., & Levitin, D. J. (2004). Performance gestures of musicians: What structural and emotional information do they convey? In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in Human-Computer Interaction* (Vol. 2915/2004, pp. 468-78). Berlin / Heidelberg: Springer

Vitouch, O. (2001). When your ear sets the stage: Musical context effects in film perception. *Psychology of Music, 29*, 70-83.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception, 6*, 383-96.

Waller, S. J. (1993). Sound reflection as an explanation for the content and context of rock art. *Rock Art Research, 10*, 91-101.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41*, 989-94.

Widmer, G., & Tobudic, A. (2003). Playing Mozart by analogy: Learning multi-level timing and dynamics strategies. *Journal of New Music Research, 32*, (259-68).

Youngerman, S. (1974). Maori dancing since the eighteenth century. *Ethnomusicology, 18*, 75-100.

Acknowledgements

The work reported in this paper has been funded in part by the Cooperative Research Centre Programme through the Australian Government's Department of Education, Science and Training. I am grateful to Robert Accordino and Felecity Baker for their helpful references, the guest editors of this special issue Oliver Vitouch and Olivia Ladinig, and two anonymous reviewers for their helpful and detailed suggestions.

---

[i] The role of vocal imitation in the evolution of language, and consequently singing, is currently under vigorous debate (Fitch, Hauser, & Chomsky, 2006; Hauser, Chomsky, & Fitch, 2002; Pinker & Jackendoff, 2005)

[ii] While the mirror neuron system has been implicated in vocal imitation (Gallese, 2007), evidence to date does not yet support a direct link (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Watkins, Strafella, & Paus, 2003; for review see Gallese, 2007).

[iii] Mirror neurons in music, emotion and movement have previously discussed by Molnar-Szakacs & Overy, 2006; Vines, Krumhansl, Wanderley, & Levitin, 2006; and briefly by Jackendoff & Lerdahl, 2006.

[iv] For a review of the ontogeny of a Theory of Mind, see Brüne and Brüne-Cohrs (2006).

[v] For a different perspective on the importance of reflective thought in the origins of music, see Parncutt (In press-a, In press-b)