21

Music and Emotion: Psychological Considerations

William Forde Thompson and Lena Quinto

Music is experienced in many different contexts and its significance for human behavior is not always obvious. Music plays an important role in many social contexts such as weddings, funerals, and parties, but its appeal cannot be fully explained by such functions. Music affects us in ways that are personal and require psychological explanation: music energizes, surprises, soothes, delights, and otherwise shapes our emotional states. Research in cognition and neuroscience supports the idea that pleasure and emotions are key motivations for listening to music. Not only does music activate "pleasure centers" in the brain (Blood and Zatorre 2001), it can communicate and induce a range of powerful emotions (Juslin and Sloboda 2001).

This latter capacity—to communicate and induce emotional states—has been the subject of intense scientific investigation. Emotional interpretations and experiences of music are extremely common and play a significant role in our appraisals of music. Indeed, whenever emotions are evoked by stimuli, they are combined with appraisals of those stimuli occurring "on multiple levels of processing ranging from automatic and implicit to conscious conceptual or propositional evaluations" (Scherer 2005: 701). This pairing of emotion and evaluation occurs because, from a biological standpoint, stimuli that elicit emotional responses are (or were, at an earlier point in evolutionary history) relevant to the major concerns of the organism (Huron 2005).

The aim of this chapter is to provide a psychological explanation for the link between musical activity and emotional states. To the extent that aesthetic evaluations are motivated by and intertwined with emotional systems, our discussion contributes to a psychological account of musical aesthetics. However, we take for granted that aesthetic evaluations are based on more than gut feelings. The complex decision-making processes that lead people to assign *value* to phenomena reflect multiple considerations that extend beyond their emotional attributes and the emotional states they induce. For example, humans may be genetically predisposed to value images or sounds depicting or derived from adaptive contexts, such as healthy mates, safe environments, and food sources (Davies 2009). Artistic phenomena may also acquire value by virtue of "making special" the social patterns, conventions, and experiences that allow life to prosper (Dissanayake 2000). Nonetheless, behavioral and neuroscientific evidence indicates that emotional systems are always implicated in such preferences and appraisals of value (Damasio 1994; Lehrer 2009). As such, understanding the capacity of music to communicate and induce emotional states is an essential step in developing a psychological model of musical aesthetics.

A number of theories have been proposed to account for links between music and emotion but fundamental questions remain. Are there general principles that might account for the connection between music and emotion across cultures and historic periods or is musical significance unique to every time and place? Does music influence emotions directly or do cognitive and motor processes mediate this link? Can broad principles of human interaction account for connections between music and emotion?

We first discuss general theories of emotion and how they view the relation between emotion and cognition. We next discuss empirical evidence demonstrating that specific attributes of music are individually associated with distinct emotional interpretations, and may be manipulated by performers and composers to convey complex and dynamic emotional messages. We also describe empirical investigations that attempt to disentangle the contributions of cross-cultural and culture-specific associations between music and emotion. Finally, we introduce a recent body of theory and data concerning the cognitive-motor implications of music, and review an emerging framework for conceiving the link between music and emotion. This conception focuses on the capacity of music to resonate with psychological processes that function in human *synchronization*, and to elicit emotional effects related to these processes. Such effects may be particularly powerful because music accommodates synchronization on multiple levels, including movement (clapping, tapping), attention, and imagination (Livingstone and Thompson 2009; Overy and Molnar-Szakacs 2009). Effects arising from the motor system, in turn, may influence aesthetic judgments (Topolinsky 2010).

1. Emotion and cognition

The relation between emotion and cognition has been the subject of considerable debate, and views on the matter have informed psychological theories of music and emotion. Are emotions entwined with cognitive processes, such as reasoning, planning, and remembering, or is there a distinct emotional system that operates independently from "cold" cognition?

One view is that emotional and cognitive processes are independent of one another (Zajonc 1980). Goleman's (1995) concept of "emotional intelligence" implies two distinct types of intelligence: one rational and the other emotional. These two types of intelligence are presumed to operate independently and need not be consistent with one another. Research on implicit memory supports the idea that preferences do not depend on cognitive processing. For example, mere exposure to stimulus patterns (e.g., melodies, random dot patterns) can lead to increased preference for such patterns, even when there is no explicit recall of them (Zajonc 1980, 1984). Presumably, these unconscious preferences arise because previously encountered stimuli, even if they are not explicitly represented in memory, can still evoke a positively valenced emotion (for a psychological explanation of the mere exposure effect, see Huron 2006).

A second view is that emotional responses are the outcome of a sequence of cognitive processes in which features are coded, classified, and finally appraised. According to this *post-cognitive* conception, emotion is the end state in a causal chain of information processing in which an event occurs, followed by sensory registration, perceptual processing, and finally an appraisal. Emotions arise at the appraisal stage (Lyons 1980; Zajonc 1980).

A third view is that emotions are intertwined with cognitive processing. According to Damasio (1994, 1999), all images are infused with affective tags or "somatic markers" that link those images with emotional associations. Through experience, images and events become tagged with particular emotions. Somatic markers function to bias cognitive processing in a way that is maximally adaptive. For example, a person may decide against a course of action if the images associated with that course have a negative connotation. Somatic markers increase the efficiency of decision-making.

A connection between emotion and cognition is also assumed in a theory of emotion proposed by Carver and his associates (Carver and Scheier 2009). The model begins with the assumption that much of behavior, and hence cognition, is goal-directed and feedback-controlled. Feedback control describes the process by which psychological or physical acts are monitored, compared to a desired state (a goal), and adjusted to decrease the discrepancy between current and goal states. Emotional experiences are thought to arise as part of this feedback control process.

Whenever there is a gap between goals and outcomes, feedback mechanisms register the discrepancy and adjust plans accordingly. The registration of a discrepancy in itself does not generate an emotional experience; it is a neutral error signal that triggers behavioral adjustment. However, the unfolding pattern of adjustments that bring the system towards or away from goal alignment is experienced phenomenally as a positive or negative experience. From an adaptive perspective, such emotional consequences function to reinforce behavioral and psychological adjustments that bring an organism toward greater alignment with goals, and to extinguish adjustments that are ineffective or counter-productive to goal attainment. We will return to this model later, suggesting that its basic components and assumptions can account for some important emotional effects that arise when listening to music.

2. Empirical studies of music and emotion

Listeners are highly sensitive to the emotional aspects of music, and they have broadly convergent emotional interpretations of music. Results from a wide range of investigations over the past century suggest that the various attributes of music, such as intensity (loudness), tempo, dissonance, and pitch height, are strongly associated with emotional expressions. In particular, changes in any of these attributes are correlated with changes in emotional interpretation (Ilie and Thompson 2006) and affective experience (Husain, Thompson, and Schellenberg, 2002; Ilie and Thompson 2011; Thompson *et al.* 2001). Such attributes contribute to an emotional *code* that may be employed by composers and performers to communicate emotions in music, or by speakers when they communicate emotions in their tone of voice (Juslin and Laukka 2003).

One important cue is *tempo*. Melodies that are played at a slow tempo tend to evoke emotions with low energy such as sadness, whereas melodies that are played at a fast tempo tend to evoke emotions with high energy, such as anger or joy. To scientifically investigate the emotional significance of tempo, Hevner (1935) presented listeners with several pieces of classical music performed at slow (63–80 bpm) and fast (102–152) tempi. Listeners heard the performances and selected from a list of adjectives the terms that best described the character of each piece. Although the two versions of each piece were identical in all respects except tempo, the emotions implied by the two versions were strikingly different. The slow tempo performances were described using terms such as serene, calm, sad, tender, and dreamy, whereas those same pieces performed at a fast tempo were described using terms such as joyous, happy, exciting, and restless.

The emotional connotations of tempo might have been learned through passive exposure to the conventions of Western tonal music, but it is also possible that they reflect natural correlations that exist between pace and emotional states. Indeed, there is strong scientific evidence that the emotional consequences of manipulating acoustic attributes such as intensity and pitch height are not restricted to Western listeners or Western music but appear to tap into universal links between the auditory system and emotional responses. This evidence has emerged from two lines of research.

First, the emotional effects of manipulating such attributes in music overlap the effects of manipulating those same attributes in speech. That is, many of the attributes that comprise an emotional code in music are equally effective at communicating emotion by tone of voice, also referred to as the *supralinguistic* dimension of speech or *prosody*. Such findings implicate an emotional communication system that functions effectively across auditory domains (Juslin and Laukka 2003; Thompson and Balkwill 2010).

In support of this idea, Ilie and Thompson (2006) presented listeners with excerpts of instrumental music and spoken passages and asked them to evaluate the emotional connotations of each excerpt along three affective dimensions: valence (pleasant–unpleasant), energy arousal (awake–tired), and tension arousal (tense–relaxed). In both domains, excerpts were manipulated in intensity (loud and soft versions), rate (fast and slow versions), and pitch height (high-pitched and low-pitched versions). Two manipulations had strikingly similar emotional effects in the two domains: for both music and speech, increases in intensity led to reliable increases in both energetic and tension arousal, and increases in rate led to reliable increases in energetic arousal. Manipulations of pitch height had different emotional effects in music and speech, however, illustrating the importance of domain-specific emotional cues.

More recently, Ilie and Thompson (2011) extended this paradigm with longer excerpts of music and speech (7 minutes). In this case, the authors administered the *Profile Of Mood States* (POMS) to evaluate emotional *experiences* induced by the music and speech, and they also evaluated two types of cognitive skill: speed of processing and creative problem-solving. Once again, manipulations of intensity, pitch height, and tempo had overlapping emotional effects following exposure to music and speech. For example, for both music and speech, participants were more energetic after listening to fast stimuli than after listening to slow stimuli. Moreover, they were able to detect and respond to visual patterns more rapidly after listening to fast music or speech than after listening to slow music or speech. Thus, not only do attributes of music and speech communicate emotional messages; they induce emotional states and alter cognitive function.

Second, the emotional effects of manipulating acoustic attributes in Western music overlap with the effects of manipulating the same attributes in non-Western music. Although the significance of genre- and culture-specific emotional cues cannot be overstated, certain acoustic attributes may tap into deep-seated interactions between auditory and emotional neural areas, providing a source of cross-cultural emotional communication within the domains of music and speech prosody (Thompson and Balkwill 2010).

Balkwill and Thompson (1999) asked Western listeners to judge the emotional content of field recordings of Hindustani ragas, and to rate structural attributes in the music. Hindustani ragas were performed with the explicit intention to evoke specific emotions. Although listeners were unfamiliar with Hindustani music, they were able to decode emotional intentions. Ragas intended to convey joy/*hasya* were assigned high ratings of joy; ragas intended to convey sadness/*karuna* were assigned high ratings of sadness; and ragas intended to convey anger/*raudra* were assigned high ratings of anger. Judgments of emotion correlated with perceptions of musical attributes. For example, joy was associated with perceptions of fast tempo and sadness was associated with perceptions of slow tempo.

This ability to decode emotional intentions in unfamiliar music is not restricted to Western listeners. In another study, Balkwill *et al.* (2004) examined judgments by *Japanese* listeners of Japanese, Western, and Hindustani music. Again, listeners were sensitive to the intended emotion in Japanese, Western, and Hindustani music, and judgments were correlated with perceptions of musical attributes. As for Western listeners, joy was associated with perceptions of fast tempo and sadness was associated with perceptions of slow tempo.

These and other studies implicate the existence of an emotional code that is instantiated in music and speech and operates across cultures. Such evidence parallels the landmark studies by Ekman and his colleagues illustrating pan-cultural facial expressions of emotion (e.g., Ekman *et al.* 1969). Positive emotions are associated with smiling whether in Britain, China, Brazil, or the Congo, just as energetic emotional states are associated with rapid and high intensity speech or music.

3. Music as a multimodal emotional signal

When speaking or singing emotionally, emotional facial expressions and gestures are often combined with auditory signals of emotion, creating a powerful multimodal affective experience (De Gelder and Vroomen 2000; Thompson *et al.* 2008). Indeed, researchers are increasingly recognizing the important role of movement in the communication of emotion in music. Singers frequently use facial and body movements to reinforce or supplement emotional messages conveyed in music (Thompson *et al.* 2005). Visual cues associated with body movements may also nurture a sense of emotional connection between musicians and audience members (Kurosawa and Davidson 2005).

Body movements provide an important source of expressivity even in the absence of sounded music.

Davidson (1993) asked musicians to perform excerpts of music in a deadpan, projected and exaggerated manner while their performances were video recorded. Silent videos of these performances were then presented to experimental participants, who provided judgments of expressivity based on the visual information alone. Marked differences in body movements were observed between the deadpan and exaggerated performance conditions, and judgments confirmed that such movements provide reliable signals of expressivity.

Facial expressions also provide viewers and listeners with emotional information. Livingstone *et al.* (2009) recorded singers with motion capture or electromyography (EMG) as they imitated phrases of emotional singing. All singers were shown audiovisual recordings of sung phrases performed with happy, sad, or neutral emotional expressions and were asked to imitate these recordings. Singers made clear and reliable facial movements during their imitations, supporting the emotional message conveyed by the sonic dimension of vocal production. Empirical studies have confirmed that such facial movements significantly influence emotional interpretations of music (Thompson *et al.* 2005, 2008). Facial expressions provide a powerful supplement to vocal signals of emotion because they can occur not only during vocal production but also before and after it. Pre-production facial expressions and gestures function to *prime* the intended emotion for listeners, facilitating accurate interpretation and encoding. Post-production facial expressions and gestures reinforce a stable representation of the emotion that was conveyed during vocal production.

4. Theories of music and emotion

What features of music are capable of inducing emotional responses, and what is the basis for this capacity? Can specific properties of music trigger emotional responses?

If so, what are they? To date there is no one theory of music and emotion with which everyone agrees. A number of theories have been proposed to address such questions and include influential discussions by Aristotle, Charles Darwin, Suzanne Langer, Leonard Meyer, Peter Kivy, and many others.

One view considers music to contain a large number of "cues" that have referential properties, such that it is possible to identify various features that composers can use to communicate fairly specific emotional connotations (for a review, see Juslin and Laukka 2003). Cooke (1959) proposed that music consists of various melodic features and patterns that have recognizable emotional significance. He argued that composers draw on these features and patterns in order to capture the nuanced and dynamic emotions that they wish to express. In effect, music is viewed as a *language* of emotion, with melodic features signifying distinct emotions.

According to Cooke, melodic intervals—the pitch distance between two consecutive notes—provide a particularly important cue. An ascending major third interval (i.e., consecutive notes separated by four semitones) represents joy and triumph; an ascending major sixth (nine semitones) implies a longing for pleasure; the minor sixth (eight semitones) suggests anguish, and the augmented fourth (six semitones) connotes hostility and disruption. Cooke supported his arguments by examining the lyrics that accompany music, observing remarkable consistency in the adjectives that occur in conjunction with particular intervals. The interval of an ascending major third is typically accompanied by words describing positive emotions, whereas the interval of an ascending minor third is more often accompanied by words implying negative emotions.

According to this *language of emotion* perspective, associations between melodic intervals and emotions can be observed across cultures, so are not merely a quirk of Western tonal music. Although Cooke's theory represents a landmark in the study of music and emotion, there is little evidence that the melodic features he identified are consistently associated with specific emotional connotations across or within genres and historic periods. The main limitation of Cooke's theory appears to be that the emotional associations he proposed were too specific. His essential argument—that music can be broken down into a collection of auditory attributes that have emotional connotations—has considerable empirical support (for a review, see Juslin and Laukka 2003).

A second view focuses on the role of expectation in music (Huron 2006; Mandler 1984; Meyer 1956). Expectancy theories are powerful because they do not rely on a referential system for generating meaning. Anything acquires meaning if it is associated with something beyond itself. With designative meaning, symbols and referents are different in kind. Language has this feature in that words are different in kind than the objects and events to which they refer. With embodied meaning, symbols and referents are the same in kind. According to Meyer (1956: 35), "one musical event . . . has a meaning because it points to and makes us expect another musical event." Thus, the emotional power of music lies in the expectations that it creates in the listener. Music has the capacity to generate complex and nuanced emotions because it continuously deviates from our expectations.

According to Mandler (1984), such responses to music instantiate a more general biological response that occurs for all unexpected events. The ability to anticipate events is essential for human survival, and all behavior is guided by anticipatory responses. Failure to predict an event can be life-threatening, and hence leads to heightened arousal and increased attentional resources. Arousal alone does not fully define the emotional experience, however. It is a bodily reaction that includes increases in heart rate, breathing, and blood pressure to put an organism into a state of heightened alert or readiness. Following this bodily response is a process of appraisal that clarifies the precise nature of the emotion experienced. That is, emotional experiences of music reflect a process of appraising bodily reactions to fulfillments and violations of expected musical events. This idea draws from William James's view that "*the bodily changes follow directly the perception of the exciting fact*," and "*our feeling of the same changes as they occur IS the emotion*" (James 1983 [1890]: 449). However, Mandler's theory focuses on physiological changes that occur in response to expectancy violations, and how music capitalizes on such expectancy effects.

In his ITPRA theory of musical expectancy, Huron (2006) extended the ideas proposed by Meyer and Mandler. He identified several ways in which expectations associated with music generate complex and nuanced emotions. Pre-outcome responses (feelings prior to an expected/unexpected event) include imagination and tension responses; post-outcome responses include prediction, reaction, and appraisal responses. *Imagination* entails contemplating future states and acting in a way that makes those states more likely if they are positive, and less likely if they are negative. *Tension* is an immediate physiological preparation for an imminent event and involves changes in arousal. *Prediction* is a transient state of reward or punishment that arises in response to the accuracy of expectation. *Reaction* and *appraisal* are emotional states that arise from assessments of the event itself independent of whether that event was anticipated. *Reaction* is a rapid "knee-jerk" process that occurs automatically and pre-attentively and activates bodily actions and/or visceral responses. *Appraisal* is more considered and need not be compatible with the reaction response.

A third view is that there are multiple mechanisms underlying the connection between music and emotion. For Juslin and Västfjäll (2008), music is capable of inducing emotion through expectancies, by directly stimulating the brain stem (a part of the brain that controls arousal and other basic functions), or by association with other emotional stimuli. The latter process of inducing emotion may be further broken down into associations with emotions themselves (classical or *Pavlovian* conditioning) and associations with events or stimuli which themselves have emotional connotations, including (a) past events (episodic memory), (b) visual imagery, and (c) the human voice.

The *multiple mechanisms* framework is an important theoretical advance because it identifies and differentiates possible mechanisms and provides suggestions for how each might be examined systematically. The framework highlights the importance of conditioned triggers of emotion in establishing links between music and emotion. Although associations can account for many emotional responses to music, they can occur for any type of stimulus and therefore have limited value as an explanation of the *unique* power of music to communicate or induce emotions. Of the remaining two mechanisms—expectancies and brain stem responses—we emphasize the former mechanism. Expectancies provide a rich and pervasive source of musical emotion, and may even account for some of the emotional effects associated with brain stem responses (Huron 2006; Krumhansl and Agnes 2008; Meyer 1956).

One of the most significant proposals by Juslin and Västfjäll (2008) is that *emotional contagion* underlies the link between music and emotion (see also, Davies 1980, forthcoming). Juslin and Västfjäll (2008) focused their discussion of this mechanism on the *voice-like* quality of certain melodies and instruments. In their view, emotional contagion is activated because at certain levels of processing, melodic patterns are registered as a "super-expressive" human voice (Juslin and Västfjäll 2008: 566). Thus, while excited individuals tend to speak rapidly and in a high-pitched voice, an exciting violin melody typically unfolds with much greater extremes in rate and pitch.

Other authors surmise that effects related to emotional contagion and mirroring are even more ubiquitous and, as such, may be relevant to many of the attributes of music, not just the voice-like qualities implied by melodies and instruments (Livingstone and Thompson 2009; Overy and Molnar-Szakacs 2009). An important instance of contagion effects is observed in behaviors that involve synchronizing with music. Synchronization is linked to the expectancies that are generated in music (Huron 2006). Physical responses, action tendencies, and mental events associated with music listening are manifestations of synchronization.

The construct of synchronization subsumes many of the issues surrounding music and emotion into a broader domain of inquiry. Synchronization in its most general sense is pervasive and essential to all human interaction, and depends on continuous and dynamic monitoring or "feedback" processes for its maintenance. Synchronization often entails physical movement (e.g., dancing, clapping, tapping, head nodding, ensemble performing), but it can also occur implicitly, that is, without overt movement. Implicit forms of synchronization include entrainment of attention with the dynamic properties of the music (e.g., rhythm, melodic accents), and alignment of mental representations with the unfolding musical structure (Jones 2009; Large and Jones 1999). According to Overy and Molnar-Szakacs (2009), the precision with which synchronization occurs is guided by activity in the *mirror neuron* system (a system of motor neurons that are activated when performing or observing an action).

Music can be construed as a technology that builds on the capacity of humans to synchronize with external sources. Its structure is optimally tailored to recruit and sustain mechanisms of synchronization. As argued by Carver and Scheier (2009), it is reasonable to surmise that emotional responses arise as experiential manifestations of feedback processes for synchronization. In this view, the error signals and appraisal processes arising from feedback loops for effective synchronization—which are engaged and manipulated so effectively by music—are manifested phenomenologically as feelings of emotion. However, just as the relation between expectancy and affect is complex and must be theoretically unpacked (Huron 2006), so too are the affective consequences of feedback processes in synchronization, and their intricate relation to musical structure and performance. We now elaborate on some of the assumptions of this model.

5. Synchronization as a pervasive construct

Although we celebrate the many differences between people, from a biological standpoint such differences are a source of tension and, presumably, are responsible for much of human conflict. Counteracting this tension is a fundamental instinct to assimilate with others and our environment. This process of assimilation takes many forms, including instinctive mimicry of other people. Synchronization of behaviors is an extreme form of assimilation and has received relatively little attention by researchers. In contrast, mimicry has been vigorously examined in social psychology and is thought to nurture a sense of similarity with other people (Chartrand and Dalton 2009). More generally, mimicry leads to improved social interactions, rapport, empathy, and bonding. It extends from simple actions such as yawning to complex behaviors that can be analyzed on psychological, social, political, and economic levels.

Mimicry refers to a set of behaviors whereby interacting individuals adopt similar speech patterns, bodily positions, gestures, and mannerisms. For example, individuals who are having a conversation tend to use the same words and clauses as their interaction partners, even when other words and clauses could just as easily be used (Levelt and Kelter 1982). They also tend to adjust their bodies to increase postural congruence, and they adopt similar mannerisms to their interacting partners. The logic and flow of human interaction and communication prohibit full synchronization of actions and vocal patterns, but the instinct to assimilate pushes in that direction. Many psychological, social, economic, and political phenomena can be viewed as a compromise between the instinct for synchronization and the constraints of goal achievement in human–environmental interactions.

The discovery of *mirror neurons* suggests a possible neural basis for this instinct (Iacoboni *et al.* 1999; Koski *et al.* 2002). Brain-imaging techniques have revealed that certain cortical regions are automatically activated when people empathize with others. Nummenmaa *et al.* (2008) showed threatening and neutral images to participants, who rated their emotional reactions to those images. Threatening images portrayed an attacker and a victim. Neutral images portrayed a benign social exchange between two individuals. As expected, participants reported experiencing negative emotional reactions to threatening scenes. Participants were also asked to imagine what it felt like to be one of the individuals depicted in the images. During this empathy phase of the experiment, changes in blood flow to regions of the brain were recorded with functional magnetic resonance imaging (fMRI). When participants empathized with others in emotionally charged situations, activation was observed in brain areas associated with facial and gestural processing, as well as the premotor mirror-neuron area. Activation of such areas may help individuals to coordinate their actions in emotionally charged contexts.

What are the benefits of a tendency to synchronize actions with others? Early theories posited that mimicry functions to nurture understanding and togetherness which lead to social benefits. However, mimicry and synchronization can also be observed in individuals who are alone, for example as they are watching a video or listening to music. When individuals hear pleasant music in isolation, they automatically exhibit activity in the zygomatic muscle—a muscle associated with the expression of a smile. Conversely, listening to unpleasant music leads to greater activation of the corrugator muscle—a muscle associated with a frown (Lundqvist *et al.* 2009). If mimicry and synchronization serve social goals, why would mimicking behavior occur in lone individuals? One interpretation is that the tendency to mimic others occurs automatically as part of a hard-wired perception–action link.

Originally proposed by James (1890), *ideomotor theory* suggests that the act of thinking about an action increases the likelihood of that action. Extending this idea, Prinz and his colleagues proposed *common-coding* theory (for a review, see Prinz *et al.* 2009). According to this view, perceiving events and planning actions are accomplished through shared neural resources, resulting in a common representational domain for perception and action (Prinz 1990, 1997). More specifically, action is represented in terms of its perceivable effects, and these representations generate observable behavior once a certain threshold of neural activation is reached.

Research on mirror neurons implies that common-coding theory is not merely a functional model but is instantiated in individual neurons in the brain. It has also been speculated that such a system has adaptive value (Chartrand and Dalton 2009). If members of a group of hominids suddenly start running for the hills, an individual with a built-in perception–action link who instinctively synchronizes her behavior with the group activity will likely survive a charging predator. Survival is less likely for an individual who evaluates the situation, scans the environment for possible sources of sprinting behavior, and rationally decides on the most sensible course of action. In short, a synchronization instinct, undergirded by rapid and automatic perception–action links, may have deep evolutionary roots because of its survival value. Although such an instinct is no longer needed to avoid predators, its genetic encoding means that effects related to synchronization persist, permeating all levels of modern society. As we will argue below, they are especially relevant to musical behaviors and the emotional effects they proffer.

Musical synchronization can occur in obvious ways, such as choral singing, ensemble performance, clapping or tapping in response to music, and dancing to music. In line with common-coding theory, however, synchronization also occurs at the level of attention and imagination. Common-coding theory implies that the perception of any action-based event will also trigger mental representations of those actions. Because actions are required to produce all musical sounds, it stands to reason that the perception of music will result in representations of possible actions for producing such musical sound. For musically trained listeners, these may include representations of specific actions associated with the details of music performance. For other listeners, performance actions are less specified but still function to predict events along the multiple dimensions of time and pitch. For example, hearing an electric guitar may increase the likelihood of simulating the act of playing guitar ("air guitar"). The details of such action-representations will depend, however, on one's experience and knowledge of the electric guitar. In any case, the result is synchronization with music.

Rhythmic expectancies—predictions about temporal location—are associated with a tendency to synchronize movements with stress points in time. The accuracy and complexity of rhythmic synchronization vary with music training and experience. It is also possible to synchronize with non-temporal features of music, however, by forming mental representations of those features to coincide with their occurrence. When engaged in musical behaviors, including music listening, synchronization can occur in several ways that afford a degree of tight behavioral and perceptual coupling that is not possible with most if not all other stimuli. That is, music is unique in its capacity to permit precise and sustained synchronization. Synchronization is permitted because of the strong use of rhythm, but can also occur for other dimensions of music.

Overy and Molnar-Szakacs (2009) proposed a framework for explaining emotional responses to music called *Shared Affective Motion Experience*, or SAME. They emphasize that music is perceived not only as sound but also in terms of the intentional, hierarchically organized sequence of the motor acts that are required to produce the signal. Auditory features are processed in the superior temporal gyrus and are combined with movement information in the mirror neuron system. The anterior insula then connects the mirror neuron system with the limbic system, ultimately allowing a neural mapping of acoustic input onto an emotional state (see also, Carr *et al.* 2003). Depending on the listener's musical experience, they will extract differing degrees of motor information, ranging from a broad intentional level for untrained listeners to the muscular level for highly trained performers.

What is the process by which mirror neuron responses are mapped onto affective experience? Why is this mapping particularly potent for music? Overy and Molnar-Szakacs surmise that music is powerful because of its capacity for *minimized prediction error*, where prediction is a prerequisite for synchronization. They note that, "familiar, predictable music can be enjoyed to the fullest" whereas "the violation of expectancies can be more emotionally dramatic" (Overy and Molnar-Szakacs 2009: 494). In addition, they argue that movements used to produce music often reflect emotional states, and these states can be induced in the listener by a process of emotional contagion. We share many of these ideas, and propose that feedback control may provide a bridge between the tendency to synchronize our movements, attention, and imagination with music and the tendency to experience music as a powerful emotional stimulus.

6. Synchronization guided by dynamic feedback control

If music engages the instinct for synchronization, how is synchronization sustained? How are errors in synchronization corrected? Although the underlying mechanisms of synchronization remain to be fully understood, the changing nature of sensory input in general and music in particular means that synchronization requires feedback control mechanisms that continuously register discrepancies between actions and events and transmit signals to initiate a dynamic process of correction.

According to Carver and Scheier's (2009) model of emotion, all signals arising from feedback control mechanisms are experienced as feelings of emotion. Feedback about goal-directed behavior can account for a wide range of affective experiences, from joy or relief after attaining a goal to frustration, sadness, or anger when attempts to achieve goals are thwarted. Emotional experiences arise whenever the attainment of a desired condition or the avoidance of an undesired condition is facilitated or impeded. In the case of active music listening, an implicit goal is synchronization with the music (and/or other participants in the musical activity) in movement, attention, and imagination. What is unique about music is that it instantiates goal-directed behavior in the form of an especially tight coupling of action–attention–imagination with perceptual input: synchronization. This potential for detailed and accurate synchronization is carefully controlled and manipulated by composers and performers and requires continuous and rapid feedback control for its maintenance.

A feedback control mechanism has four components: an input, a reference value, a comparison, and an output (Carver and Scheier 2009; Mackay 1966). The input is the perception of target points of synchronization whereas the reference value is the goal of aligning behavior to these target points. The comparison is a measure of the degree of alignment, also called the "error signal." The output is an adjustment in response to the detected error. If the comparison process detects no error, current behavior is unchanged. Synchronization feedback is always controlled by discrepancy-reducing loops, so if a discrepancy is detected, behavior is adjusted to diminish the discrepancy.

Musical synchronization can occur on many levels, from activities that mirror the surface or tone-to-tone details of music to social forms of synchronization that account for musical movements. Some levels of synchronization are defined by simple action sequences such as tapping or clapping to the beat. Other levels are defined by brief action sequences such as singing along with a familiar phrase of music. Still other levels are characterized by synchronized social and economic behaviors such as attending concerts and purchasing musical products. Ultimately, all synchronization is related to action, but it can also occur with no observable bodily movement.

Synchronization also occurs at the level of attention (Jones 2004, 2009; Jones and Boltz 1989; Large and Jones 1999). Dynamic attending theory (DAT) defines *entrainment* as a biological process that leads to synchronization between mechanisms that control our attention and unfolding environmental events. These mechanisms are driven by oscillations of neural activation. Neural oscillations readily entrain to temporally regular patterns such as music that has a steady beat, but they may also guide the timing of attention for all biologically significant environmental phenomena that are potentially predictable. When listening to music, time spans at a metric level elicit a neural oscillation that has an internal periodicity that aligns with that metric level. In this way, neural oscillations "tune into" temporally predictable events by adjusting their phase in response to entrainment feedback (Jones 2009). Although attentional synchronization need not entail observable movement, it is a prerequisite for action synchronization. When attentional synchronization occurs without overt action, it may be viewed as an *action tendency* aimed at synchronization (see also, Frijda 1986).

7. Feedback and affect

Affective states are generated in response to goals and whether or not they are attained. When the achievement of a significant goal is thwarted, we may feel sad or frustrated. When an important goal is met we may feel a sense of relief or joy. Such feelings reflect the operation of feedback processes for goal-directed behavior. Carver and Scheier (2009) identify two processes that operate to assist with goal-directed behavior. The first is a behavior-guiding feedback process that registers error signals and acts to correct that error. The second is a feedback loop that monitors discrepancy reduction over time, essentially supervising the first process. Mathematically, the output from the second feedback loop is equivalent to the derivative of the output from the first feedback process. The simultaneous operation of both feedback systems, one controlling position and the other controlling velocity, allows for rapid and effective control of goal-directed behavior.

Figure 21.1 illustrates the model. The figure shows that acoustic input and perceptual analysis lead to processes by which perceivers synchronize their mental and physical state with the input. Synchronization occurs on motor, attentional, and imaginative levels and is guided by the behavior-guiding and monitoring feedback-control mechanisms described above. Experiences associated with the behavior-guiding process are equivalent to the tension and prediction responses identified by Huron (2006) in his ITPRA theory of expectancy. In the tension response, an arousal response is elicited as a target point of synchronization is approached, and the intensity of that arousal varies with the imminence, rate of approach, and significance of the point of synchronization. Changes in basic parameters of the acoustic input such as intensity (loudness) and tempo influence the tension response by altering the perceived significance of points of synchronization and the rate at which they are approached. In the prediction response, positive or negative feedback arises depending on whether

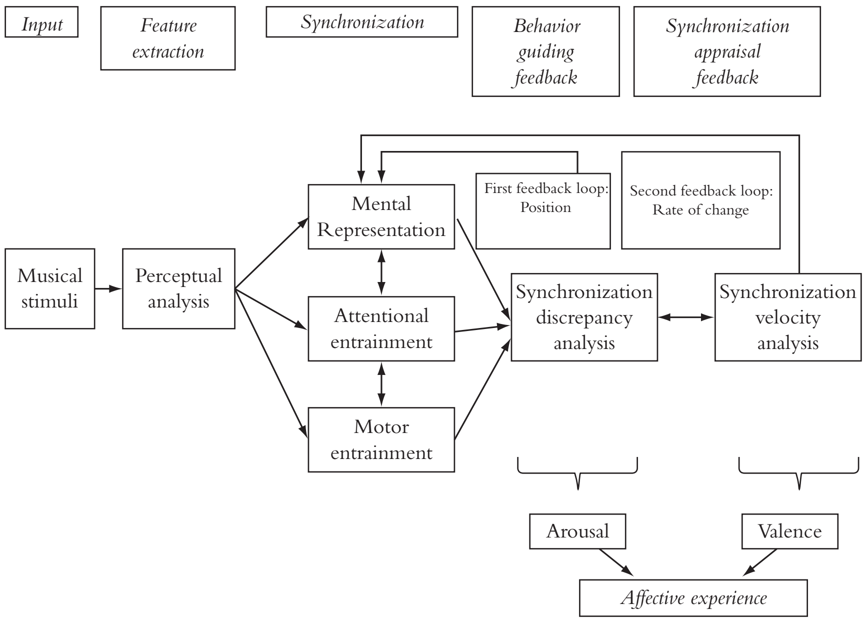

Figure 21.1 An illustration of the synchronization feedback model

synchronization with the target event is correctly aligned. Positive feedback rewards and reinforces alignment with the target; negative feedback motivates increased effort in synchronization.

The second *monitoring* feedback process is also manifested by affective experience. Specifically, if the system determines that there is an increase in the accuracy of synchronization over time, positive feedback results. If the system determines that there is a decrease in the accuracy of synchronization over time, then negative feedback results. The intensity of this second feedback signal is influenced by the significance and rate of increase or decrease in synchronization, which are determined by basic parameters of the acoustic signal. Thus, there are two sources of affective experience generated by feedback loops in synchronization. Moment-to-moment *arousal* and *reward* generated by the (first) behavior-guiding feedback process are combined with positive or negative experiences (*valence*) generated by the (second) monitoring feedback process.

On first glance, such a proposal might appear to be restricted in scope. After all, how often do we synchronize our actions with music, and can feedback mechanisms account for the subtleties of emotional responses to music? Although this is a valid concern when only one level of synchronization is considered (e.g., clapping to music), synchronization takes place along several dimensions and multiple levels of analysis. Synchronization at the level of the tactus—the pulse of the music—may generate only subtle effects when considered in isolation. However, the effects of metric synchronization must be considered in conjunction with higher-order rhythmic levels on which synchronization occurs, and with synchronization effects for other dimensions of the music, including melodic, harmonic, tonal, and phrasing structure. These immediate synchronization effects of music then interact with social synchronization effects that permeate and shape the political economy of music.

Because synchronization relies on prediction and expectation, it follows that any type of emotional response that can be explained by *expectancy mechanisms* can also be explained by synchronization feedback. As demonstrated by Meyer (1956), Huron (2006) and others, the unfolding patterns of violations and fulfillments of expectations that occur while listening to music can account for powerful and complex emotional responses, especially when multiple levels of expectancy are considered simultaneously. As an example, the experience of "awe" may be evoked when low-level violations of expectancy, which generate arousal responses, combine with high-level fulfillments of expectancy, which generate feelings of reassurance. As an analogy, standing at the edge of the Grand Canyon may generate a visceral elevation of arousal because of low-level links between visual perception and brain stem responses, mixed with high-level feelings of reassurance generated by an appraisal of the circumstances as non-life threatening. The combination of feedback from multiple levels of processing generates the emotion of awe.

It is important to note that feedback processes for human synchronization cannot account for all emotional responses to music. The model ignores responses to music that arise from learned associations between music and emotion, or other *mediated* responses. A composition may induce an emotional or aesthetic response because it reminds us of a death or a birth, or lends itself to social and political analyses, or because of its sheer artfulness. These responses are not produced directly by the music but are mediated by associations or cognitive appraisals. In other words, music is not the *object* of these emotional reactions. Such emotional experiences may be powerful but they are not generated directly from the music, and are not our focus.

Instead, our account combines arguments and evidence derived from psychological discussions and investigations of emotions, including Juslin and Västfjäll (2008), Jones (2009), Huron (2006), Carver (Carver and Scheier 2009), Overy and Molnar-Szakacs (2009) and others. We suggest that the unique power of music to elicit emotion lies in its capacity to engage participants in tightly controlled synchronization at multiple levels of abstraction. Music optimally recruits processes of synchronization that are ubiquitous in human behavior and that greatly influence our emotional lives.

References

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

——and R. Matsunaga (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research*, 46/4: 337–49.

Blood, A. J. and R. J. Zatorre (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*, 98/20: 11818–23.

Carr, L., M. Iacoboni, M. C. Dubeau, J. C. Mazziotta, and G. L. Lenzi (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, 100/9: 5497–5502.

Carver, C. S. and M. F. Scheier (2009). Action, affect, and two-mode models of functioning. In E. Morsella, J. A. Bargh, and P. M. Gollwitzer (eds.), *Oxford Handbook of Human Action*. Oxford: Oxford University Press, pp. 298–327.

Chartrand, T. L. and A. N. Dalton (2009). Mimicry: It's ubiquity, importance, and functionality. In E. Morsella, J. A. Bargh, and P. M. Gollwitzer (eds.), *Oxford Handbook of Human Action*. Oxford: Oxford University Press, pp. 458–86.

Cooke, D. (1959). *The Language of Music*. London: Oxford University Press.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.

——(1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.

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

Davies, S. (1980). The expression of emotion in music. *Mind*, 89: 67–86.

——(2009). Life is Passacaglia. *Philosophy & Literature*, 33: 315–28.

——(forthcoming). Music-to-listener emotional contagion. In Tom Cochrane, Bernadino Fantini, and Klaus Scherer (eds.), *The Emotional Power of Music*. Oxford: Oxford University Press.

de Gelder, B. and J. Vroomen (2000). Perceiving emotions by ear and by eye. *Cognition and Emotion*, 14: 289–311.

Dissanayake, E. (2000). *Art and Intimacy: How the Arts Began*. Seattle: University of Washington Press.

Ekman, P., E. R. Sorenson, and W. V. Friesen (1969). Pan-cultural elements in facial displays of emotions. *Science*, 164: 86–8.

Frijda, N. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.

Hevner, K. (1935). The affective character of the major and minor modes in music. *American Journal of Psychology*, 47: 103–18.

Huron, D. (2005). The plural pleasures of music. In William Brunson and Johan Sundberg (eds.), *Proceedings of the 2004 Music and Music Science Conference*. Stockholm: Kungliga Musikhögskolan Förlaget, pp. 65–78.

——(2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

Husain, G., W. F. Thompson, and E. G. Schellenberg (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities: Re-examination of the "Mozart effect." *Music Perception*, 20: 151–71.

Iacoboni, M., R. P. Woods, M. Brass, H. Bekkering, J. C. Mazziotta, and G. Rizzolatti (1999). Cortical mechanisms of human imitation. *Science*, 286: 2526–8.

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

———— (2011). Experiential and cognitive changes following seven minutes exposure to music and speech. *Music Perception*, 28: 247–64.

James, W. (1983 [1890]). *The Principles of Psychology*, with introduction by George A. Miller. Cambridge, MA: Harvard University Press.

Jones, M. R. (2004). Attention and timing. In J. G. Neuhoff (ed.), *Ecological Psychoacoustics*. San Diego, CA: Academic Press, pp. 45–89.

—— (2009). Musical time. In S. Hallan, I. Cross, and M. Thaut (eds.) *Oxford Handbook of Music Psychology*. Oxford: Oxford University Press, pp. 81–92.

—— and M. Boltz (1989). Dynamic attending and responses to time. *Psychological Review*, 96: 459–91.

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

—— and J. A. Sloboda (eds.) (2001). *Music and Emotion: Theory and Research*. Oxford: Oxford University Press.

—— and D. Västfjäll (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31: 559–75.

Koski, L., M. Iacoboni, M. C. Dubeau, R. P. Woods, and J. C. Massiotta (2002). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, 89: 460–71.

Krumhansl, C. L. and K. R. Agnes (2008). Musical expectancy: The influence of musical structure on emotional response. *Behavioral and Brain Sciences*, 31: 584–5.

Kurosawa, K. and J. W. Davidson (2005). Non-verbal interaction in popular performance: The Corrs. *Musicae Scientiae*, 19/1: 111–36.

Large, E. W. and M. R. Jones (1999). The dynamics of attending: How we track time-varying events. *Psychological Review*, 106: 119–59.

Lehrer, J. (2009). *The Decisive Moment: How the Brain Makes Up Its Mind*. Melbourne, Australia: The Text Publishing Company.

Levelt, W. J. M. and S. Kelter (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14: 78–106.

Livingstone, S. R. and W. F. Thompson (2009). The emergence of music from the Theory of Mind. *Musicae Scientiae*: 83–115.

Livingstone, S., W. F. Thompson, and F. A. Russo (2009). Facial expressions and emotional singing: A study of perception and production with motion capture and electromyography. *Music Perception*, 26: 475–88.

Lundqvist, L.-O., F. Carlsson, P. Hilmersson, and P. N. Juslin (2009). Emotional responses to music: experience, expression, and physiology. *Psychology of Music*, 3/1: 61–90.

Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.

Mackay, D. M. (1966). Cerebral organization and the conscious control of action. In J. C. Eccles (ed.), *Brain and Conscious Experience*. Berlin: Springer-Verlag, pp. 422–45.

Mandler, G. (1984). *Mind and Body: Psychology of Emotion and Stress*. New York: Norton.

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

Nummenmaa, L., J. Hirvonen, R. Parkkola, and J. K. Hietanen (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *NeuroImage*, 43: 571–80.

Overy, K. and I. Molnar-Szakacs (2009). Being together in time: Music experience and the mirror neuron system. *Music Perception*, 26: 489–504.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann and W. Prinz (eds.), *Relationships Between Perception and Action: Current Approaches*. Berlin: Springer, pp. 167–201.

——(1997). Perception and action planning. *European Journal of Cognitive Psychology*, 4: 1–20.

——G. Ashersleben, and I. Koch (2009). Cognition and action. In E. Morsella, J. A. Bargh, and P. M. Gollwitzer (eds.), *Oxford Handbook of Human Action*. Oxford: Oxford University Press, pp. 35–71.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44/4: 693–727.

Thompson, W. F. and L.-L. Balkwill (2010). Cross-cultural similarities and differences. In Patrik Juslin and John Sloboda (eds.), *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford: Oxford University Press, pp. 755–88.

——P. Graham, and F. A. Russo (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, 156/1–4: 203–27.

——F. A. Russo, and L. Quinto (2008). Audio-visual integration of emotional cues in song. *Cognition & Emotion*, 22/8: 1457–70.

——E. G. Schellenberg, and G. Husain (2001). Mood, arousal, and the Mozart effect. *Psychological Science*, 12/3: 248–51.

Topolinsky, S. (2010). Moving the eye of the beholder: Motor components in vision determine aesthetic preference. *Psychological Science*, 21/9: 1220–4.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35: 151–75.

——(1984). On the primacy of affect. *American Psychologist*, 39: 117–23.