Recent functional magnetic resonance imaging (fMRI) studies indicate that there are specific parts of the brain, including the superior temporal sulcus, that respond selectively to human voices (Belin et al. 2000). Moreover, electroencephalogram (EEG) research reveals that such processing is faster and more extensive for vocal than for non-vocal sounds (Schirmer et al. 2007). These two findings suggest that our brains can differentiate between emotional voice sounds, and emotional musical sounds, very early in processing.

Studies of patients with brain injuries provide further evidence of the distinction between the processing of voices and non-voices. Griffiths et al. (2004) describe a patient who underwent a stroke that resulted in damage to his left insula and left amygdala, and consequently possessed a deficit in "musical emotion" processing, while retaining normal music perception and voice prosody perception. If the same mechanism is responsible for both processes, how can one be impaired and the other intact? This indicates that separate neural networks may underlie voice prosody and musical emotion perception.

A dissociation was also found in patient I.R., who suffered damage to her left superior temporal gyrus and left insula, among other regions, resulting in the loss of music recognition, while leaving musical emotion perception and speech prosody intact (Patel et al. 1998; Peretz et al. 1998). Although patient I.R. correctly discriminated the emotional nature of musical excerpts, she claimed that she was only guessing, and was surprised to learn that she was doing well. It therefore appears that her discrimination ability was outside of her awareness or intention (Peretz et al. 1998). This is strikingly similar to the phenomenon of blindsight, in which patients are able to respond to visual stimuli, but are not consciously aware of perceiving them (Weiskrantz et al. 1974). Interestingly, I.R.'s speech prosody judgments were normal (she did not have to "guess"). The fact that emotion perception in music was outside of awareness, while emotion perception in the voice was within awareness, is further evidence of distinct systems for voice prosodic processing and musical emotion processing.

Another problem with the emotional contagion mechanism is that it fails to explain emotional reactions to music that are not congruent with the emotional content of the music. For example, when watching a frightening movie and you hear an angry voice, or a "threatening" musical excerpt, this theory predicts that you should also feel angry. However, this is not the case; instead, we feel fearful. From an evolutionary perspective, it would not be adaptive to possess a simulation mechanism that causes you to become angry upon hearing an angry voice or other threatening sound. Therefore, this theory fails to take into account all types of emotional expressions.

J&V present a thoughtful list of mechanisms responsible for musical emotions. Such a framework is much needed, and undoubtedly will be informative for guiding future studies. However, there are weaknesses in one of the proposed mechanisms – emotional contagion – that need to be addressed. One possibility is that the same underlying mechanism is responsible for both musical and vocal emotional expressions, as *super-expressive voice* theory postulates, but that the former is more sensitive to damage than the latter. Further studies must directly compare "musical emotions" and vocal emotions to elucidate these differences.

The role of signal detection and amplification in the induction of emotion by music

doi:10.1017/S0140525X08005529

William Forde Thompson and Max Coltheart

Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia.

Bill.Thompson@mq.edu.au **max@maccs.mq.edu.au**

http://www.psy.mq.edu.au/staff/bthompson/

Abstract: We propose that the six mechanisms identified by Juslin & Västfjäll (J&V) fall into two categories: *signal detection* and *amplification*. Signal detection mechanisms are unmediated and induce emotion by directly detecting emotive signals in music. Amplifiers act in conjunction with signal detection mechanisms. We also draw attention to theoretical and empirical challenges associated with the proposed mechanisms.

We consider Juslin & Västfjäll's (J&V's) article from the perspective of a distinction we propose between two classes of mechanisms: *signal detection* and *amplification*. Signal detection mechanisms are unmediated sources of emotion, including brain stem responses, expectancy, and evaluative conditioning. They are unmediated because they induce emotion by directly detecting emotive signals in music. Amplifiers act in conjunction with signal detection mechanisms. They include episodic memory, visual imagery, and possibly emotional contagion.

Signal detection mechanisms. J&V distinguish *brain stem responses* from the other mechanisms proposed. This neuroanatomical classification presents a source of confusion, however, because the brain stem has multiple functions and may be implicated in the other five mechanisms. An alternative conception is the *psychophysical signal detector*, which encompasses brain stem responses and evaluative conditioning. Balkwill and Thompson (1999) defined psychophysical signals as sound attributes having consistent emotional connotations across domains (e.g., music and speech prosody) and cultures. The signals may be learned or congenital. Learned signals arise through evaluative conditioning, acting on attributes correlated with emotion. Congenital signals trigger hard-wired affective responses including, but not restricted to, brain stem activity.

J&V restrict discussion of *expectancy* to syntax, but syntactic structure represents only one attribute relevant to expectancy. Expectancy implicates multiple mechanisms at several processing levels. Huron's (2006) expectancy model includes imagination, tension, prediction, reaction, and appraisal, underscoring the challenge of defining expectancy as a unified mechanism operating solely on syntactic structure. For example, the *tension response* is a physiological preparation for any imminent event and involves changes in arousal that likely arise from brain stem activity and are adjusted according to the degree of uncertainty about the outcome. *Prediction responses* are transient states of reward or punishment arising in response to accuracy of expectations. Accurate expectations lead to positive states. Inaccurate expectations lead to negative states.

The mechanism of *evaluative conditioning* proposed by J&V conflates a process of *learning* following long-term exposure to environmental regularities with an *emotional-induction mechanism* that detects signals and induces emotion. However, feedback mechanisms that establish learned associations are usefully distinguished from *signal detection mechanisms* that decode emotions during listening. Learning mechanisms act both on musical pieces and on psychophysical attributes of sound. The sadness of Shakespeare's monologue "Tomorrow and tomorrow . . ." nurtures associations between emotions communicated by verbal information and statistical parameters of the acoustic signal, such as slow delivery and little pitch variation. Such psychophysical signals are correlated with emotional states and connote them even when embedded in nonverbal stimuli such as music.

Amplification mechanisms. J&V posit *visual imagery* as an independent cause of emotional experience. But imagery primarily accompanies or amplifies emotional experience; emotional states induced by music are conducive to imaginative processes that elaborate and amplify that experience. Moreover, imaginative processes are not restricted to visual images. Some music has a conversational quality that stimulates an *auditory* image of talking; other music can stimulate a *kinesthetic* image such as floating. Music can even generate *conceptual* imagination, such as the idea of death. Imagery during music listening may have less to do with music than with the absence of visual stimulation to which a listener must attend.

Like visual imagery, *episodic memories* are rarely an independent cause of musically induced emotion but primarily amplify emotional experience. Episodic memory is powerful precisely because there is typically congruence in the emotional connotations of the music and episode. More generally, because self-report studies are susceptible to demand characteristics, the prevalence of episodic memory and imagery is probably overestimated. Most music listening accompanies activities such as driving, reading, and socializing, with little opportunity for imagery and episodic memory. Emotional effects of music are subtle but they occur continuously. In contrast, tangible visual images (a meadow) or episodic memories (a day at the beach) – because they are extraordinary – are over-reported.

According to the authors, *emotional contagion* is triggered by voice-like qualities of music, including intensity, rate, and pitch contour ("super-expressive voice"). However, such music-speech associations must be established in the first place through conditioning, and then decoded by psychophysical signal detectors. Once signals are decoded, emotional contagion converts perceived into felt emotion through a process of mimicry, amplifying the output of perceptual mechanisms. It is feasible that emotional contagion is directly activated by acoustic signals with no mediating process, but it should be engaged not only by voice-like attributes, but any emotional signal.

Conclusions. J&V characterize the literature as confused. A more optimistic interpretation is that the field is developing, and the target article is a valuable stimulus for this progress. Researchers have carefully controlled musical attributes, and cross-cultural studies have elucidated the capacity of people to interpret emotional connotations of music or speech from foreign cultures by relying on psychophysical signals that are culture-transcendent (Balkwill & Thompson 1999; Thompson & Balkwill 2006) and that have similar connotations in music and speech (Ilie & Thompson 2006).

Emotional responses to music seem to arise from three broad sources: psychophysical signal detection, expectancies, and emotional amplifiers. Many issues remain unresolved. The difference between perceived and felt emotion – not explored here – has implications for theories of music and emotion (Schubert 2007). Moreover, research suggests that emotional responses to music implicate multisensory processes not acknowledged in the target article (Thompson et al. 2005; in press). Finally, it is important to define modularity explicitly (Peretz & Coltheart 2003) since this is a much-misunderstood concept (Coltheart 1999).

Music as a dishonest signal

doi:10.1017/S0140525X08005530

Sandra E. Trehub

Department of Psychology, University of Toronto – Mississauga, Mississauga, ON, Canada L5L 1C6.

sandra.trehub@utoronto.ca

http://www.utm.utoronto.ca/7041.0.html

Abstract: Instead of the discrete emotions approach adopted by Juslin & Västfjäll (J&V), the present perspective considers musical signals as functioning primarily to influence listeners in ways that are favorable to the signaler. Viewing music through the lens of social-emotional regulation fits with typical uses of music in everyday contexts and with the cross-cultural use of music for infant affect regulation.

Given widespread disagreement on the existence and nature of musical emotions, their links to non-musical emotions, and the conditions under which they are elicited, Juslin & Västfjäll's (J&V's) attempt to specify induction mechanisms for such disputed emotions may be unwarranted or premature. Although the notion of discrete emotions has faced increasing challenges in the general emotion literature (Barrett 2006; Russell 2003; Scherer 2000a), it continues to prevail in the literature on music cognition. Moreover, it remains central to J&V's conceptualization. It is odd, indeed, that the basic emotions (e.g., happiness, sadness, anger, fear), which concern automatic, innate responses to highly significant stimuli (Izard 2007; Tomkins 1962), have been co-opted for emotions expressed in music and for feelings experienced while listening to music. Typical studies provide listeners with preselected musical materials (based on expert agreement regarding the intended emotion) and highly constrained response choices. High levels of agreement in such contexts are viewed as confirmation of similarities in the perception of emotion in music, both within and across cultures (Balkwill & Thompson 1999; Juslin & Laukka 2003). Open-ended response formats would undoubtedly generate unruly individual differences. For appraisal theorists (e.g., Ellsworth & Scherer 2003), however, who regard specific emotions as arising from subjective appraisals of events, individual differences in affective responses are to be expected.

Bachorowski and Owren's (2003) selfish-gene (Dawkins 1989) account of vocal (nonverbal) emotion offers an alternative to conventional discrete emotions or appraisal approaches. These authors dispute the notion of emotive vocalizations as honest signals reflecting the emotional state of the signaler, even its valence. In their view, the primary function of emotive signals is to influence listeners' affect, attitudes, and behavior in ways that are favorable to the signaler. It follows that the signaling process was shaped over evolutionary time by such consequences. These consequences can be direct, arising from signal acoustics (e.g., amplitude, variability), or indirect, arising from familiarity with the signal or prior affective experiences.

Similarly, musical signals are unlikely to reflect the affective state of the composer or performer, and they may not *express* emotion in any simple sense. Presumably, composers and performers strive to influence the affective state of listeners – to *move* them or connect with them in one way or another. It is possible, then, that emotional responses to music can be approached more productively within the broad context of communication. Diversity in listeners' responses would not be surprising in view of the diversity of personal and musical histories, as well as the variable network of associations with music in general or with specific musical pieces. Indeed, when listeners are given the option of selecting from dozens of empirically derived descriptors of feelings in response to various pieces of music, *relaxed*, *happy*, *dreamy*, *transcendent*, *enchanted*, *nostalgic*, and *touched* are among the most frequently endorsed terms, whereas *sad*, *angry*, and *fearful* are among the least common (Zentner et al., in press). Nevertheless, the same listeners use such negative terms to characterize emotions expressed by the music.

Viewing music through the lens of social and emotional regulation (including self-regulation) fits with typical uses of music in everyday contexts, whether as background or foreground (DeNora 2000; Sloboda & O'Neill 2001). It also fits with the cross-cultural use of maternal vocalizations for regulating infant affect and promoting attachment. Mothers use a distinctly musical style when they speak to their preverbal infants (Fernald 1991), a style that includes individually distinctive, or signature, tunes (Bergeson & Trehub 2007). The efficacy of maternal vocal signals is evident in infants' enchantment with this speech style (Fernald 1991; Werker & McLeod 1989) – its positive affect, in particular (Singh et al. 2002). Mothers across cultures also sing to their infants, doing so by means of a distinctive musical genre (lullabies and play songs) and a distinctive singing style (Trehub & Trainor 1998). Divergent social-regulatory goals are reflected in divergent song choices, with lullabies prevailing in cultures that value calm, contented infants, and play songs prevailing in cultures that value infant vitality and expressiveness (Tsai 2007). Infants prefer such infant-directed singing to typical informal singing (Trainor 1996). They also exhibit greater engagement and more prolonged attention to maternal singing than to maternal speech (Nakata & Trehub 2004), perhaps because the former is especially effective in modulating arousal (Shenfield et al. 2003).