Jordan Bailey

Second Notes – Pass

Primary Sources

1. ***Measuring Cardiac and Electrodermal Responses of Emotional States and their Persistence***

    a. Assuming that music can generate happy and sad emotions, it is possible to argue that valenced music could also generate moods. In this perspective, the mood can be defined as a general affective background that persists over time without a specific stimulus (Beedie et al., 2005). To assess the musical influence on mood, Eich and Metcalfe (1989) produced continuous musical induction throughout a 90-min session, while participants performed an encoding and retrieval task. Mood states were measured throughout different moments in the session. Results showed that, in the beginning, positive (or negative) songs made participants feel happy (or sad) as measured by self-report, and with repeated exposure, the mood manipulation lost some of its intensity. (**p.2**)

    b. The literature on the persistence of the emotional effects after musical exposure on HR and SCL signals are scarce, yet critical. If an induced emotion persists after its induction, one may also expect to observe a variation in physiological responses of the subjects induced by music. One of the few studies investigating the persistence of emotions used four different valence–arousal film clips with its original soundtracks to induce positive high-arousal, positive low-arousal, negative high-arousal, and negative low-arousal. These mood inductions lasted 10 min; they were followed by a computer task (online shopping) aiming to observe whether induced emotions would last throughout the cognitive task. The authors reported lower SCL for the negative videos during the accomplishment of the task. In contrast, the HR results for those participants who watched the two positive videos did not present significant differences. Also, after approximately 9 min of the computer task, no self-reported arousal effects were observed (Gomez et al., 2009). (**p.3**)

    c. The length of each music excerpt was precisely 3 min, and each musical excerpt was edited using Audacity R 2.1.2 to avoid startling participants. "Fade out" and "fade in" were applied when necessary and normalized to the same Root Mean Square loudness level. The musical excerpt used for negative induction was Albinoni "Adagio" composed in G-minor with 3/4 time signature. For positive mood induction, we used Bach "Brandenburg concert n◦ 2" composed in F-Major with 2/2 time signature. The neutral excerpt was "Steve Reich-Variations for winds, strings, and keyboard" composed in C-minor/C-flat, and B-major, with varied time signatures (**p.4**)

d. Participants rated their emotional state on the valence dimension selecting one out of nine adjectives (adapted from Plutchik, 2001). There were three adjectives related to positive mood (happy, excited, and euphoric), three to negative mood (sad, melancholic, and distressed), each one corresponding to a low, medium, or high-intensity level, respectively. We also selected three adjectives to rate neutral mood (neutral, indifferent, and unresponsive) . Although the valenced adjectives were rated as low, medium, and high intensities, participants were not informed about these intensities. The arousal dimension was measured immediately after valence, in which participants were requested to rate the arousal of the previously selected adjective on a 7-point analog scale about how they felt at that moment. One corresponded to "I feel very little aroused" and 7 to "I feel very much aroused." Moreover, to assess ratings of musical pleasantness, a 7-point scale was also included, with 1 meaning unpleasant and 7 meaning very pleasant. (**p.4**)

e. This study was the first to examine the persistence of an EIM paradigm. According to our valence–arousal self-report, the valenced adjectives changed quickly to neutral ones after the second minute of the recovery phase, which is congruent with Gomez et al. (2009) and Kuijsters et al. (2016), both using videos. However, over time, SCL remained increased at least 4 min more for positive and negative EIM. One possible explanation for this longer-lasting arousal effect could be that, contrary to the physiological measures, the self-report rates are discrete measurements that cannot continually capture all the variations of the emotions felt (Mauss and Robinson, 2009) or participant's attention to the stimulus decreases. (**p.11**)

2. *Embodied Cognition is not what you think it is*

a. Because perception is assumed to be flawed, it is not considered a central resource for solving tasks. Because we only have access to the environment via perception, the environment also is not considered a central resource. This places the burden entirely on the brain to act as a storehouse for skills and information that can be rapidly accessed, parameterized, and implemented on the basis of the brain's best guess as to what is required, a guess that is made using some optimized combination of sensory input and internally represented knowledge. This job description makes the content of internal cognitive representations the most important determinant of the structure of our behavior. *Cognitive science is, therefore, in the business of identifying this content and how it is accessed and used.* (**p.2**).

b. What is the task to be solved? Embodied cognition solutions solve specific tasks, not general problems, so identifying how an organism produces a given behavior means accurately identifying the task it is trying to solve at the time. Taking things one task at a time opens up the possibility of smart solutions (Runeson,1977). Organisms using smart solutions solve particular problems using heuristics made possible by stable features of the task at hand, rather than general purpose rote devices which apply algorithms to solve the task. For common tasks, smart solutions are typically more efficient, more stable, and more economical than rote solutions (e.g., Zhu and Bingham, 2008, 2010). (**p.3**)

c. What are the resources that the organism has access to in order to solve the task? Embodied cognition implies that there are resources, plural, available to the organism. These resources include the brain but also the body, the environment, and the relations between these things (e.g., the motion of our bodies through the environment). A task analysis should include an exhaustive list of resources available that might contribute, beginning with those available via perception and action and only hypothesizing more complex cognitive resources once the capabilities of these other resources have been exhausted. An exhaustive list is possible if you are able to characterize your task formally; tasks are differentiated from each other in terms of their underlying dynamics (e.g., Bingham, 1995) and thus it is becoming common practice to formalize the task description using the tools of dynamical systems (e.g., Fajen and Warren, 2003; Bingham, 2004a,b; Schöner and Thelen, 2006). (**p.2**).

d. How can these resources be assembled so as to solve the task? Solving a specific task means creating a smart, task-specific device that can do the job (Bingham, 1988). To be more specific, it means assembling the required resources into a dynamical system that solves the task at hand as its behavior unfolds over time. Remember, these resources can be distributed over brain, body, and environment. Since we only have access to information about our bodies and the environment via perception, an embodied analysis must include a detailed account of the perceptual information used to connect the various resources (Golonka and Wilson, 2012). (**p.3**).

e. Does the organism, in fact, assemble, and use these resources? It is always an empirical question whether the dynamical system hypothesized in step 3 is, in fact, an accurate description of the system the organism has assembled to solve the task. The basic experimental tool for establishing the identity of a dynamical system is the perturbation experiment; systems respond to perturbations of resources in a manner that is specific to the role that resource plays in the system, and this allows you to map the composition and organization of the system at hand (e.g., Kay et al., 1987, 1991; Wilson and Bingham, 2008)(**p.3**).

3. Event-related skin conductance responses to musical emotions in humans

a. Measuring electrodermal activity is one technique that provides readily accessible autonomic indices, such as the skin conductance response (SCR). SCR is due to rapid fluctuations in eccrine sweat gland activity, which result from the liberation of acetylcholin by the sympathetic nervous system [2]. This measure has the advantage over other measures of the autonomic nervous system such as heart rate, since SCR is under strict control of the sympathetic branch of the nervous system. (**p. 145**)

b. The study of emotions induced by music has already benefited from the use of electrodermal indices. However, studies have typically monitored tonic levels of electrodermal activity over long-duration periods of music listening (30-s–6-min) [1,7], rather than the phasic, more transient changes occurring in response to brief stimulations (i.e. a few seconds). Tonic changes in skin conductance level, such as those recorded over a few minutes during long-duration stimuli, and phasic changes in the form of SCRs occurring 1–4 s after discrete stimuli, may not reflect the same underlying physiological and psychological processes (**p. 145**)

c. Stimuli consisted of 28 musical clips of 7 s duration, for which the emotional content has been previously assessed and validated [14]. These clips were composed in a film music genre for research means, and so were completely novel for the subjects. The clips were equally distributed across four emotions, and equated in terms of arousal and valence. However, due to a strong habituation effect across the entire set of musical clips, SCRs were analyzed for the first half of excerpts presented. In this smaller set, four excerpts were to evoke a sense of fear (‘peur’), another four happiness ('gaiete´'), three others sadness ('tristesse') and three clips peacefulness ('apaisement') (see musical scores in Appendix A). As in the literature, happy excerpts were characterized by a fast tempo and were played in a major mode whereas sad excerpts had a slow tempo and were played in a minor mode [13]. The structural determinants of the peaceful and fearful excerpts are less well established and were empirically verified in our previous study [14]. (**p. 146**)

d. In the validation of our material, arousal and valence of each stimulus was evaluated on a 10-point scale, from 1 to 10 (with 1 meaning 'very calming' and 10 'very stimulating' on the arousal scale; and with 1 meaning 'very unpleasant and 10 'very pleasant' on the valence scale…A significant effect of the emotion category was observed on the ratings of arousal ($F(3,36) = 139.6$, $P < 0.001$) and valence ($F(3,36) = 58.4$, P<, 0.0001). Both Fearful and happy melodies were rated more stimulating than those peaceful or sad melodies (Tuckey test, $P < 0.05$). In contrast, melodies associated with peacefulness and happiness were judged as more pleasant than melodies associated with sadness and fear (Tuckey test, $P < 0.05$). (**p. 146**)

e. The major finding of the present study is that musical emotions induce SCRs varying according to arousal. The activation of the sympathetic nervous system by musical excerpts is itself under control of many subcortical structures such as the amygdala, the hypothalamus, and cortical structures, especially the orbito-frontal cortex and the temporal lobes [2,4]. Thus, SCR measurement may offer a window to the brain structures involved in emotional responses to music. Indeed, it possible that SCRs sensitivity to arousal in musical emotions is subject to top-down regulations. In further studies, the recording of SCRs, in addition to the use of neuroimaging techniques, would allow the assessment and the localization of neural networks associated with event related autonomic responses to music arousal. (**p. 148**)

4. ***Role of tempo entrainment in psychophysiological differentiation of happy and sad music?***

a. One way to explore the emotional responses to music is to consider both conscious emotional assessments in recognition task (emotion identification), for example, and automatic psychophysiological responses to stimuli that reflect emotional responses likely related to feeling. (**p. 17**)

b. Nevertheless, several experiments have demonstrated clear physiological differentiations between happy and sad music. In general, happy excerpts elicit larger skin conductance responses (Khalfa et al., 2002), faster heart and respiration rates relative to sad excerpts. In a study where participants listened to 3 min of sad, fearful and happy music, happiness and sadness were differentiated by the electrodermal conductance level (reflecting activation of the sympathetic nervous system), and by cardio-respiratory responses (Krumhansl, 1997). The heart rate was slower during sad music than when listening to happy music, the diastolic blood pressure increased more for sad than happy excerpts, and the breathing rate increased less for sad than happy music. (**p. 18**)

c. This happy–sad music distinction may be the result of responses to the acoustical parameters responsible for this distinction. Indeed, on one hand, happy music is typically fast (and written in major mode) while sad music is typically slow (and written in minor mode) (Balkwill and Thompson, 1999; Gabrielsson and Juslin, 1996; Peretz et al., 1998). The happy/sad distinction relies on two major musical features: tempo (i.e. the number of beats per minute) (Gabrielsson and Lindstrom, 2001; Schellenberg et al., 2000), and mode (i.e. the specific subset of pitches used to write a given musical excerpt) (Krumhansl, 2001). The specific manipulation of musical tempo affects the perception of happiness and sadness (Dalla Bella et al., 2001; Khalfa et al., 2005). On the other hand, entrainment to tempo or rhythm which is not specific to music has been described. It is a phenomenon in which two or more independent rhythmic processes synchronize with each other (Clayton et al., 2005; Jones and Boltz, 1989). This is why Etzel et al. (2006) pointed out that the significant respiration rate difference between happy and sad excerpts they observed seems to be attributable to entrainment of respiration to tempo (Haas et al., 1986), and/or to rhythm (i.e. temporal patterning conveyed by tones' perception). (**p. 18**)

d. The happy excerpts were written in a major mode at a fast tempo (average Metronome Marking = 136 beats/min) (range: 110–154). In contrast, the sad excerpts were written in minor mode at an average slow tempo of 52.3 beats/min (Metronome Marking) (range: 40–69). Two other versions of the original excerpts were created with the Encore 3.0 software, by successively removing the pitch variations (rhythmic version), and pitch plus temporal variations (tempo version) in each melody (see sad and happy stimuli respectively in Appendix 1a and 1b). Six musical conditions, each comprising three excerpts, were presented to the participants: 1) 3 happy— original melodies (happy), 2) 3 sad—original melodies (sad), 3) 3 happy—rhythm (happy R), 4) 3 sad—rhythm (sad R), 5) 3 happy—tempo (happy T), 6) 3 sad-tempo (sad T). (**p. 19**)

e. As expected from previous studies on tempo variations (Dalla Bella et al., 2001; Gabrielsson and Lindstrom, 2001; Peretz et al., 1998), happy and sad music were clearly identified with emotion recognition rates superior to 89%. However, the happy/sad distinction ability decreased when the pitch variations were removed, and more importantly, the stimuli became less pleasant. However, the emotion originally intended for the sad Tempo-only stimuli remained well recognized (superior to 90%). This latter finding parallels previous results (Schellenberg et al., 2000), and supports the hypothesis that equating tone durations in sad excerpts give them a special "funeral" quality making them sound very sad. Moreover, since the sad Rhythm, and sad Melody stimuli were found to be more pleasant than the sad Tempo stimuli, it may also explain why sad Tempo stimuli were more easily associated to the negative (sadness) than the positive (happiness) emotion (**p. 23**)

5. ***Emotion elicitation during music listening: Subjective self-reports, facial expression, and autonomic reactivity.***

a. Listening to music activates brain regions involved in emotion and reward, including the ventral striatum, amygdala, orbitofrontal cortex, anterior cingulate cortex, and the insula (Blood & Zatorre, 2001), as well as other areas typically associated with cognitive processes such as the anterior hippocampus and auditory cortex (Koelsch, 2020). Interestingly, brain activation reveals a segregation of subcortical areas that responds to differences in affective dimensions. For example, chills evoked by pleasant music tend to correlate with an increased activation in nucleus accumbens and insula (Blood & Zatorre, 2001). In addition, high pleasurable moments during music listening are related to dopamine release in the ventral region of the striatum (Salimpoor et al., 2011). In contrast, amygdala seems to be specifically activated during unpleasant chill responses (Klepzig et al., 2020), as well as during the fear and tension evoked by music (Koelsch, 2006; Koelsch & Skouras, 2014). (**beginning of introduction**)

b. Regarding the methodological diversity among studies, one important caveat influencing the comparability of results is the duration of music excerpts, which have been ranging between 15 and 30 s (Brushan & Asai, 2018; Gomez & Danuser, 2004) and even more than 60 s (Baumgartner et al., 2006; Bullack et al., 2018; Etzel et al., 2006; Guhn et al., 2007; Lundqvist et al., 2009; Roy et al., 2009; Sammler et al., 2007). Interestingly, only a few studies have used music stimuli with durations shorter than 15 s (Dellacherie et al., 2011; Gringas et al., 2015; Khalfa et al., 2002; Vieillard et al., 2012), despite being the standard procedure in other stimuli modalities such as affective pictures (IAPS: Bradley et al., 2001) or emotional sounds (IADS: Bradley & Lang, 2000). (**middle of introduction**).

c. A total of 42 musical excerpts (14 unpleasant, 14 pleasant, and 14 neutral) were selected from the Film Music Stimulus Set (FMSS; Eerola & Vuoskoski, 2011) based on the Spanish normative values for two affective dimensions (valence and energy arousal) (Fuentes-Sánchez et al., 2020). Particularly, unpleasant and pleasant excerpts were evaluated below 4 and above 6 in hedonic valence, respectively, whereas all stimuli in both categories were rated above 6 in energy arousal. Music excerpts classified as neutral were rated between 4 and 6 in hedonic valence, and below 4 in energy arousal…Furthermore, with the aim of reducing other confounding variables, all excerpts used in this experiment were also rated as certainly unfamiliar in the previous Spanish validation study (Fuentes-Sánchez et al., 2020). Particularly, the normative values in familiarity (3-point scale: 0 unfamiliar; 1 somewhat familiar; and 2 very familiar) were as follows: $M = 0.30$, $SD = 0.17$, 95% CI [0.24–0.35]. (**Stimuli and Design**).

d. Pairwise correlations showed a high correspondence between subjective evaluations for shortened (obtained in this experiment) and longer original excerpts (Spanish normative values), in hedonic valence, energy arousal, and tension arousal ratings, $r(1) = .96$, .95, and .94, all $ps < .0001$, respectively. Similarly, strong correlations were found in all the ratings assigned to discrete emotions for shortened and longer excerpts. Specifically, significant associations were found for happiness, anger, fear, tenderness, and sadness, $r(3) = .98$, .97, .98, .89, and .93, all $ps < .0001$, respectively. (**Results**)

e. Current findings revealed a strong positive correlation between subjective evaluations of the shortened and original excerpts, both when analyzing the ratings for the three affective dimensions and specific emotions, suggesting that the selected film music excerpts conveyed similar emotions independently of their duration. Additionally, our findings demonstrate that shortening the film music excerpts did not alter the emotional perception contained therein. This result goes in line with a previous study that certainly reported consistency of the emotional responses between long (30-s) and short (1-s) classical music excerpts (Bigand et al., 2005). Accordingly, it becomes necessary to open a methodological debate about how much time is needed to elicit steady and consistent emotions in laboratory controlled conditions, especially when using dynamic and complex stimuli such as music excerpts. (**discussion**)

6. ***The Role of Peripheral Feedback in Emotional Experience With Music***

   a. Evidence for the role of peripheral feedback in emotional experience comes from a variety of sources. Schachter and Singer (1962) theorized that physiological arousal combined with cognitive labeling resulted in an emotion. More specifically, they argued that when individuals have no immediate explanation for their arousal, they will label their emotion in terms of the cognitions available to them; when they have an explanation at hand, they are unlikely to use other information to "label" their feeling state; and they will describe their feelings as emotions according to the extent of their physiological arousal. (**p. 81**)

   b. What, then, is the process by which body state can influence emotional experience? Philippot et al. (2002, pp. 606–607) provide a useful conceptualization of emotion and body states in terms of three strands of thinking. *The first conception* of emotion and body state, the "undifferentiated arousal model," argues that arousal feedback can intensify emotional states and that the arousal-emotion relationship is mediated by causal attributions regarding the source of the arousal (Reisenzein, 1983, p. 258; Schachter, 1964) but does not differentiate between different emotions. *The second conceptualization*, the "cognitive appraisal model," suggests that body changes are the result of cognitive appraisal or action readiness (Frijda, 1986). This suggests that the body changes that accompany emotion feeling states are differentiated according to the kind of appraisal or the kind of response readiness required to deal with the environment (Frijda, 1986, p. 165). *The third conception*, which Philippot et al. term the "central network model," argues that the different components that make up emotion are connected by a neural or cognitive network. This group of theories suggests that body changes are differentiated according to the emotion experienced. (**p. 83-84**)

   c. The research reported in this article investigates the influence of arousal on emotional experience directly for the first time, by focusing on the effect of induced arousal on emotional experience of music and on how this is mediated by the valence and energy (arousal) characteristics of the music. By inducing arousal through exercise, rather than through a film, which has its own valence characteristics, the experiments carried out here aim to investigate the effects of undifferentiated arousal on emotional experience of music. (**p. 84**)

d. It was found that two groups of participants with different levels of induced arousal reported different emotional experiences: an exercise group, with elevated pulse rates, reported feeling a greater intensity of emotions congruent with the valence of the music than did a relaxation group. This result suggests that physiological arousal can influence listeners' experience of emotion in relation to music, as hypothesized. The association of elevated arousal levels with increased ratings of emotions congruent with the valence of the piece suggests that physiological arousal intensifies the dominant valence of the response and confirms the pattern found in nonmusical studies. These findings are consistent with undifferentiated arousal models, which suggest that arousal facilitates performance of the dominant response. Failure to find any influence of physiological arousal on the energy (arousal) dimension of emotional response suggests that listeners do not use their body state as information about the emotional energy of the music. (**p. 94**)

e. When we experience emotional feeling states we also experience body sensations associated with those states. But, as this research shows, this phenomenon can also work in the opposite direction: body sensations from one source can be misattributed/transferred to another. In this case, the body sensations of arousal due to exercise are attributed or transferred to the music, and those sensations can intensify the emotion thought to be expressed by the music and/or the emotion felt in relation to the music. (**p. 113**).

7. ***Cardiovascular and respiratory responses during musical mood induction***

a. It is generally accepted that large and reliable changes in physiological states are associated with emotional responses, regardless of the manner in which the emotional response was induced. There is consensus that such physiological changes are a reliable correlate of certain psychiatric disorders, including anxiety and panic disorders and depression (Berntson and Cacioppo, 2004; Berntson et al., 1998; Grossman, 1983; Wientjes, 1992). However, whether specific physiological patterns for each unique normal emotional state exist is controversial (e.g. Collet et al., 1997; Hagemann et al., 2003; Levenson and Ekman, 2002). (**p. 58**).

b. The present study expands the literature of psychophysiological measurements of musically induced emotions by examining individual changes in physiological activity during the stimuli and coordination of respiration with the music. The goal of the present study was to determine whether consistent cardiovascular and respiratory changes occur while subjects experience emotions induced by music. We chose our music stimuli in a pilot study based on its ability to reliably induce reports of strong happiness, sadness, and fear in the listeners. We hypothesized that (a) the induction of emotion would be associated with reliable changes in heart rate and respiration, and that (b) these changes in heart rate and respiration would differ systematically between the different induced moods. It was expected that changes would be consistent with those reported in previous studies: decreased respiration and heart rate during sadness compared to fear or happiness inductions, with the measures highest on the happiness inductions and intermediate during fear. (**p. 59**).

c. Music stimuli were selected from a large pool of potential stimuli using a pilot study. The chosen stimuli produced the most intense and specific reported experience of each target emotion: happiness, sadness, and fear (details are presented in Johnsen, 2004). The stimuli consisted of 12 music clips; four different clips were chosen to induce each target mood (fear, sadness, or happiness). Details of each stimulus appear in Table 1. The stimuli were short classical music selections taken from movie soundtracks ranging in length from 74 to 189 s (M = 136 s). Stimuli of various lengths were used so that each clip could form a musically complete unit. The stimuli are labeled by a letter indicating the targeted mood (H = happiness, F = fear, S = sadness) and a number indicating its place in the presentation order (the presentation order was the same for all subjects). The music was selected based on how well it induced each specific mood; no effort was made to match tempo, mode, or pitch. The stimuli were presented via headphones at a loud, but comfortable, volume. (**p. 59**).

d. Subjects reported experiencing the targeted mood at a stronger intensity than the other emotions following each induction. The mean rating given by the subjects on each clip to each question appears in Fig. 1, with bars indicating standard error of the mean. The ratings were lower on the sad clips than the fear or happy ones, representing a more mixed reaction to the sad and fear clips than the happy ones. Nevertheless, the questionnaire results indicate that the music clips were effective at eliciting the targeted mood. Analysis of the dial and questionnaire data is presented elsewhere (Johnsen, 2004); this study did not attempt to relate the dial and physiological responses. Few subjects reported familiarity with the music and no subject was able to correctly identify the source of any clip. (**p. 61-62**).

e. The subjects' median emotional intensity ratings indicate distinct emotional experiences: the median rating was 6 for the fear and happiness inductions and 4 for the sadness induction on a scale of 0 (''none'') to 9 (''very much''). No differences were found on any of the traditional cardiovascular measures examined, but evidence was found that the heart rate decelerated during the sadness induction and accelerated during the fear inductions. Differences in total breath length and total expiration length were found in the expected direction (slower respiration during the sadness than the fear or happiness inductions) on the time-domain measures but not when changes during the clips were examined. (**p. 65**).

8. *The Emotional Power of Music: Multidisciplinary Perspectives on musical arousal, expression, and Social Control*

a. Accordingly, studies using electroencephalography (EEG) showed that when listening to isochronous metronome beats, the periodicity of beats creates expectations (Brochard et al. 2003) and can be tracked in the brain waves (Will and Berg 2007; Nozaradan et al. 2011). However, the precise neural basis of meter perception is still an issue of research. In fact, it is a feature of expressive music to also play with tempo variations, adding to expectations and violations thereof that can also be induced by melodic features. Meyer (1956) as well as Huron (2006), state that music is a powerful tool for generating melodic, harmonic, and, through the metrical structure, also temporal expectations. Both authors claim that either meeting this expectancy or disappointing it can produce an emotion. However, disappointment of expectation in this sense means only slight deviations from the metronomic tempo, for example small accelerations and decelerations, or anticipations of a note. Thus, these kinds of deviations, which are essential to the expressiveness of music, do not disturb the global metrical structure of the music but reconfirm the regularities of the composition and help to reinforce entrainment processes. (**p.312**)

b. In fact it was shown that several motor circuits become activated when listening to metrical music (Grahn and Brett 2007). This might imply an internal preparation of actions and explain the spontaneous initiation of movements when music is heard. Several studies also showed that the basal ganglia, a crucial subcortical relay in motor pathways, are consistently implicated in rhythm processing (Grahn 2009; Kotz et al.

2009; Schwartze et al. 2011). In Parkinson patients, who have problems with initiating movements (including with gait) due to basal ganglia dysfunction, astonishing improvements have been documented in response to music (Pacchetti et al. 2000). Listening to marching music, with strong and pronounced beats, some patients could be helped to walk very smoothly and without hesitation. This effect is probably due to a stimulation of the basal ganglia via the rhythm of the music, which in consequence facilitates the execution of gait movements. The basal ganglia are also strongly implicated in the processing of pleasant emotion. For example, activations in the striatum are associated with different forms of reward, like food, sex, and monetary gain (Berridge and Robinson 1998; Small et al. 2003) but also with pleasant music (Blood and Zatorre 2001; Brown et al. 2004; Menon and Levitin 2005). The basal ganglia therefore seem to represent an ideal candidate for an anatomical and functional basis of interactions of rhythm and emotion processing, as they are involved in both the dopaminergic reward system and the motor circuits. (**p. 313**)

c. In a recent study, Salimpoor et al. (2009) asked participants to bring their favorite music to the lab and counterbalanced the stimuli in such a way that each participant was listening to his preferred music and to the favored music of another participant as a control condition. The physiological reactions between preferred music and control condition differed enormously. As the pieces were different for each participant, no conclusion can be drawn regarding the tempo entrainment. However, this study shows clearly that subjective pleasantness and musical preferences can drastically change the physiological reactions to music. This result might be explained by the fact that recognizing one's preferred music will trigger a state of high arousal and interest. This effect could also be described by an induction mechanism that Juslin et al. (2010) called "evaluative conditioning," simply representing a conditioned physiological response triggered by a familiar stimulus. (**p. 314**)

9. ***Evolutionary musicology meets embodied cognition: Biocultural coevolution and the enactive origins of human musicality.***

a. The phenomenon of entrainment may be observed in many ways and over various timescales in both biological and non-biological contexts (de Landa, 1992; Clayton et al., 2005; Becker, 2011; Knight et al., 2017). Most fundamentally, it is understood in terms of the tendency for oscillating systems to synchronize with each other2 . Accordingly, biological and social systems can be conceived of as dynamically interconnected systems of oscillating components (from metabolic cycles to life cycles, from single neuron firing to regional patterns of activity in the brain, from individual organisms to social groups and the broader biological and cognitive ecology; McGrath and Kelly, 1986; Oyama et al., 2001; Varela et al., 2001; Ward, 2003; Chemero, 2009). Importantly, the components of such systems influence each other in a non-linear or recursive way. As such, organism and environment are not separate domains, but rather aspects of "one non-decomposable system" that evolves over time (Chemero, 2009, p. 26). Moreover, the development of coupled systems is guided by local and global constraints that allow the system to maintain stability—to be resistant to perturbations, or to regain stability once a perturbation has occurred. This is, of course, crucial for living systems, which must maintain metabolic functioning within certain parameters if they are to survive. (**p.5**)

b. Here it should be noted that the biocultural approach also develops a theory about the origins of vocal musicality, albeit one that is deeply connected to the rhythmic factors just described. This entails the development of a repertoire of "gesture calls" similar to those found in modern primates and many mammalian species (grunts, pants-hoots, growls, howls, barking, and so on; see Tomlinson, 2015, p. 89–123). These do not involve the abstract, symbolic-representational, and combinatorial properties employed by modern languages. Rather, they are tightly coupled with the same mimetic, emotional, and embodied forms of communication that characterize pre-human toolmaking. It is suggested that the vocal expressions associated with these gesture-calls reflected the sonic aspects (rhythmic and timbral) of these environments, the motor-patterns of production, as well as the gestural and social rhythms (e.g., turn taking, social entrainment) that developed within the cultural ecology. (**p.7**)

c. With regard to this point, ethnomusicological and sociological research has revealed musical activity around the world to be central for human well-being—it is inextricable from work, play, social life, religion, ritual, politics, healing, and more (Blacking, 1973, 1995; Nettl, 1983, 2000; DeNora, 2000). Moreover, in many cultural environments music is highly improvisational in character, and the acquisition of musical skills begins in infancy and develops rapidly, often without the need for formal instruction (Blacking, 1973; Cross, 2003; Solis and Nettl, 2009). It has also been suggested that because certain physical and cognitive deficits need not hinder survival and well-being in modern Western society, certain "musical" impairments may go almost completely unnoticed (van der Schyff, 2013a). Likewise, music's relevance for human survival across evolutionary time has been considered in terms of its importance for bonding between infants and primary caregivers, and between members of social groups (Benzon, 2001; Tolbert, 2001; Dissanayake, 2010; Dunbar, 2012). (**p.3**)

10. ***Electrodermal activity during dissonant music in musicians and non-musicians***

   a. The EDA research presented in this manuscript describes a general increase in SCR among those who experience unpleasant music or sounds, and some research has shown a greater reaction in the presence of musical experience. The data of EMG responses to unpleasant stimuli are similarly conclusive, particularly in the case of the corrugator muscle reaction to unpleasant stimuli (Biehl, 2015). This study seeks to utilize the remaining data from research done by Bumgarner in 2015 to explore hypotheses informed by previous research findings. This study proposes that EDA responses in all subjects will be more pronounced in the presence of dissonant music as they were rated as more unpleasant. Additionally, the EDA response to dissonant music will be greater in musicians than in non-musicians despite the non-significant effect in pleasantness ratings among High Music Experience participants and Low Music Experience participants. (**p. 21**)

   b. There was a significantly different response to dissonance between the HE and LE participants. This difference showed a more pronounced elevation of SCR to dissonant music among HE participants. This finding is similar to those observed by Bumgarner (2015) and Biehl (2015) who also found higher EMG responses tied to negative emotional valence observations collected during the study. A more pronounced response to dissonance among the HE participants was a finding in the study from which this one is based (Dellacherie et al., 2010) and provides more evidence of the effects of musical training on music perception. (**p.39**)

11. ***Effects of Sad and Happy Music on Mind-Wandering and the Default Mode Network***

   a. Notably, our results reveal that participants' mental activity while listening to sad vs. happy music was self-referential, in line with (i) individuals reporting to mind-wander about personally significant matters7 and (ii) evidence of a putative role of the DMN's midline core in self-referential processing19. Interestingly, a recent meta-analysis40 showed that activity within regions of the DMN, such as mPFC, PCC and pIPL (the same areas that exhibited increased centrality in response to sad vs. happy music in the present study), was associated with both personal goal processing and mind-wandering. Therefore, future research could stratify the self-referential component of mind-wandering evoked in response to sad music by specifically testing whether sad music, compared with happy music, increases reflections on personal goals. (**p.5**)

b. Sad music is slow-paced music associated with low levels of arousal, while happy music is fast-paced music associated with high levels of arousal. Because arousal is an intrinsic component of music-evoked emotions, it is challenging to disentangle its contribution to the observed relationship between sad music and increased mind-wandering (i.e., to determine to what extent spontaneous thoughts are triggered by the evoked low arousal independent of the evoked sadness). Although we controlled for the tempo of the music stimuli in Experiment 1B, happy music evoked higher arousal than sad music. This underscores that while tempo clearly influences arousal, arousal does not simply vary as a function of tempo. In spite of the difficulty to empirically separate emotion (sadness/happiness) from arousal, further research may try to pin down the mind-wandering effects on one of these two factors by using, for example, music evoking different emotions with similar arousal (e.g., sad and peaceful music, or happy and fearful music). Nevertheless, our data suggest that mind-wandering is modulated not only by arousal levels but also by the quality of the evoked emotions (sad, happy), because fast sad music, compared with slow happy music, tended to elicit stronger mind-wandering, despite evoking significantly lower arousal (Fig. 3A). Moreover, note that meditation and relaxation practices aimed at facilitating mindfulness (and at the same time avoiding mind-wandering) usually make use of music evoking low arousal emotions with peaceful and relaxed (but not sad) emotional tone. Therefore, it is unlikely that arousal is the only factor driving the changes in mind-wandering. (**p. 6**)

c. We demonstrate that music modulates self-generated thought: During sad (vs. happy) music, listeners direct their attention inwards, engaging in spontaneous thoughts, which are related to the self and emotional aspects of life; during happy (vs. sad) music, listeners are more focused on the music itself and exhibit reduced mind-wandering levels. Tus, our findings highlight the capability of music to trigger specific mental processes as a function of its emotional tone, opening a novel line of future research elucidating the impact of music on internally-oriented cognition. This has crucial implications for the application of music in a variety of domains including education and psychotherapy. (**p.6**)

12. ***Emotion response and regulation to "happy" and "sad" music stimuli: Partial synchronization of subjective and physiological responses***

a. The majority of participants reported feeling an emotional response to both the “sad” and “happy” music samples, and the majority were also reported being able to “regulate” their response to at least two of the “sad” pieces and one of the “happy” pieces. In support of the second hypothesis, self-reported emotional responses were associated with significant changes in autonomic nervous system activity (heart rate and skin conductance). Finally, our data reveal the first evidence to our knowledge to show partially synchronized changes in a self-reported and physiological measure of emotion response as a result of regulation of emotions experienced in response to music. Participants’ self-reported ratings of their emotional response were concordant with that predicted from the normative perceived emotion data of the musical excerpt source (Eerola & Vuoskoski, 2011) with regard to the direction of the emotional change (i.e., increased sadness to “sad” music and increased happiness to “happy” music). In the current study, the “sad” music pieces induced these congruent self-reported emotional responses in a greater proportion of the sample than did the “happy” music. However, the “sad” music also induced a more mixed response than reported by Eerola and Vuoskoski, with moderate happiness induced also. Whether this difference in responsiveness to “sad” and “happy” music reflects a real difference in human capacity to be more responsive to emotionally sad stimuli than to happy stimuli warrants further investigation in a larger sample using a greater variety of stimuli of positive emotions. Furthermore, such individual variation in happiness responsiveness may be explained by individual factors, such as differences in trait resilience (for a review, see Tugade & Fredrickson, 2007). (**p.10**)

b. The self-reported emotional responses to both “happy” and “sad” music in this study were accompanied by significant changes in averaged skin conductance and heart rate. The direction of physiological changes for the “sad” music was consistent with previous research, with decreases in skin conductance and heart rate (Khalfa et al., 2008; Krumhansl, 1997; Lundqvist et al., 2009; Nykliček, Thayer, & Doornen, 1997). This represents a synchronized emotional response to music stimuli in line with current theories of emotion (Levenson, 1994; Scherer, 2009). The decrease in heart rate and skin conductance in response to “happy” music was, however, inconsistent with previous research, which has demonstrated increased autonomic nervous system activity to arousing and pleasant music (Hodges, 2010; Khalfa et al., 2008; Lundqvist et al., 2009). This inconsistent physiological response may reflect error in measurement or it may be a result of mixed emotional responses to the pieces of music used in this condition, which when averaged over the 15-s period, appear contradictory. For instance, while participants reported liking the “happy” music, it is possible that some participants may have associated this music with sad or relaxing visual imagery or episodic memories (see Juslin & Västjfäll, 2008), resulting in a decline in autonomic arousal associated with an emotional experience mixed with peace, nostalgia or feeling moved. Furthermore, the experimental demand (in particular, cognitive load) involved here may have had its own additive influence on the resulting autonomic response. Given the dynamic nature of music, future research may attain greater sensitivity if a time series analytical approach is taken instead, to explore whether changes in responsiveness across measures are correlated more closely at this more temporally defined level. Nevertheless, the current findings are consistent with recent demonstration that skin conductance appears to be more closely synchronized with a self-reported negative than positive emotional response (Balconi, Grippa, & Vanutelli, 2015). (**p.11**)

Secondary Sources

1. *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*

a. By imitation I mean not only the overt behavior of "monkey see, monkey do" but also covert imitation that occurs only in imagination. These forms of imitation occur whenever we attend to the behavior of others, whether in the performing arts or athletics, or in learning a particular skill from someone else's demonstration, or in merely taking an interest in what others are doing. When we imitate overtly or covertly, in effect we are responding to two implicit questions: What's it like to do that? and its twin question, what's it like to be that? We answer these questions in part by overtly and covertly imitating the behavior of others. (**p.11**).

b. I will refer to overt mimetic behavior as mimetic motor action (MMA), and for the relevant muscle-related brain processes that do not manifest in overt actions I will use the term mimetic motor imagery (MMI): mimetic for imitative, motor for muscle related, and imagery for "thought," "imagination," or "mental representation." I intend imagery to include not only voluntary and conscious forms, but especially those forms that occur automatically and with or without our awareness. The involuntary and nonconscious forms of MMI are in some respects the most significant in the construction of musical meaning. (**p.12**).

c. In plain terms, participants were better at comprehending heard actions (the sound of musical performance) that shared an exertion schema with actions that they had performed previously. These findings are also consistent with the proposition that rhythm perception involves real time MMI, or mimetic enactment of the rhythms heard, based on previous overt imitation (MMA). (**p.19**).

d. Bodily representation refers to the bodily states and changes of state that occur in response to an external stimulus. These representations involve both nonmimetic and mimetic processes. Via MMI and MMA we represent an observed action in the musculoskeletal system and correlated portions of the brain, as well as in systems that control respiration, heart rate, and blood chemistry. To understand the role of MMI and MMA, we have to relate these representations to established meanings of five terms: perception, comprehension, conceptualization, recognition, and inhibition. (**p.40**).

e. Perceptions of external stimuli are nonmimetic representations, including the auditory perception of musical sounds. However, when we give our attention to musical sounds, they are also represented in the form of MMI and often in the form of MMA…Conceptualizations can be understood in part as enduring representations based on ephemeral mimetic and nonmimetic representations. For example, the musical concept of a step is grounded in nonmimetic perceptions of sounds and mimetic representations of sounds as imitable actions…Recognition of external stimuli is based in part on information gained via the external senses—the appearance, sound, feel (texture, temperature, weight, etc.), smell, and/or taste of external stimuli. But recognition is also based on other responses and representations in the perceiver, including mimetic representations…[and] Inhibition in the present context refers to neurological inhibition of action. (**p.40**).

2. ***Cross-Cultural Comparisons of Affect and Electrodermal Measures While Listening to Music***

a. In 1997, Krumhansl [7] measured EDL among other signals from 38 college students with musical backgrounds, while they listened to six classical music excerpts, divided into three groups based on their intended emotion: fear, sadness and happiness. Results showed a decrease in EDL for all emotions, with sadness being the lowest, followed by fear and happiness. Khalfa et al. [8] measured the magnitude of EDRs from 34 participants, with no specified background, who were presented with 28 musical clips lasting seven seconds that were unknown to participants. These were classified in four groups: fear, happiness, sadness, and peacefulness. EDRs showed higher magnitudes for both fear and happiness, and lower magnitudes for sadness and peacefulness. No significant differences in EDR magnitudes were found between happiness and fear or between sadness and peacefulness. (**p.2**).

b. In 2007, Grewe et al. [11] utilized EDA from 38 participants (the majority of whom were musicians), to study pieces that could evoke chills. EDRs were compared against self-report measures (participants pressed a button when experiencing a chill), and segments in which both were not observed were discarded. Pieces were selected by the researchers and participants, with no emotion classifications. They found that chills coupled with EDRs are relatively rare events, supporting previous findings from [12], [13], and that the musical event that triggered most chills was the entrance of a singing voice. (**p.3**).

c. Due to the conflicting results and the non-uniformity of the methods, it is difficult to extract meaningful relationships between physiological responses, music, and affect from these as well as other studies in the literature [18]. Based on this, EiM was designed with two specific goals. First, to address the concern of the generalizability of results when physiological responses are measured in a laboratory environment [19], [20], data is collected outside of a laboratory in a more ecological environment. Second, to address issues that may arise in an environment that is not completely controlled, as well as the many potential differences in physiological responses with respect to age [21], [22], sex [23], [24], race and ethnicity [21], [25], personality [26], [27], and even the time of day [28], [29], among others, a large sample size is imperative. It is also important that the study draw from demographically diverse populations [30]. Finally, it is important to highlight that this approach allows for the exploration of a number of different research questions. (**p.3**).

d. With regards to further investigation, our results suggest that there may have been differences between groups that did and did not present the HE response for familiarity with the stimulus, musical preference, age, and felt arousal. Specifically, we suggest that studies that incorporate other stimuli available in the EiM database and target these relationships specifically. In this and our other published work, musical preference has repeatedly demonstrated an effect on both self-reported and psychophysiological response. Studies that delineate the characteristics of this relationship would have important implications for the field of affective computing. (**p.7**).

e. The limitations of these results highlight the importance of databases like EiM, as they serve as a starting point for the finer-grained studies required to thoroughly explore the relationships between affect and physiology. They provide the means for exploratory analyses that are not only strong enough in power to uncover legitimate links between affect and physiology, but also the data necessary to outline the design of the more focused studies required to reveal the details of these relationships. (**p.7**).

3. *Origin of Music and Embodied Cognition*

a. At lower levels KI acts automatically: sensory-motor experiences are directly embodied. But at higher levels abstract knowledge is called abstract exactly because it does not exist pre-formed in the world, it is created through the interaction of the world and the mind. But cognitive dissonance (CD), a mechanism opposite to KI, might interfere at higher levels. CD is a discomfort caused by holding conflicting cognitions (Festinger, 1957; Cooper, 2007; Harmon-Jones et al., 2009). This discomfort is usually resolved by devaluing or discarding the conflicting cognition. This discarding often occurs below the level of consciousness; it is fast and momentary (Jarcho et al., 2011). It is also known that the majority of new knowledge originates through the differentiation of previous knowledge, which is the mechanism for several broad empirical laws: Zipf's law, the power law, Pareto law emerge when new entities (or usage) evolve from pre-existing ones (Simonton, 2000; Newman, 2005; Novak, 2010). Therefore, almost all knowledge contradicts other knowledge to some extent. According to CD theory, any knowledge should be discarded before its usefulness becomes established (Perlovsky, 2013a). (**p.2**)

b. Does music help in overcoming CD? Masataka and Perlovsky (2012a,b) have reproduced the above experiment with music in the background and observed that the toy is not devalued. Another experiment demonstrated that academic test performance may improve while listening to music. Perlovsky et al. (2013) demonstrated (1) that students allocate less time to more difficult and stressful tests (as expected from CD theory), and (2) with music in the background students can tolerate stress, allocate more time to stressful tests, and improve grades. These experiments confirmed the hypothesis that music helps in overcoming CD. It is likely that music emerged and evolved for a fundamental cognitive function: music makes the accumulation of knowledge and human evolution possible. (**p.2-3**).

c. How can multiplicity of emotions be explained and justified from a cognitive and evolutionary standpoint, and why has this ability emerged? The proposed hypothesis relating music to CD suggests the following explanation. CD produces a variety of emotional discomforts, different emotions for every combination of knowledge—in other words, a huge number of emotions. Most of these emotions are barely noticed because they lie below the level of consciousness, and in these unconscious states they produce disincentives for knowledge. Music helps to overcome these emotional discomforts by developing a huge number of conscious musical emotions. The mind being conscious of the multiplicity of emotions can bring into consciousness emotions of CD, and thus be prepared to tolerate them. We enjoy even sad and difficult musical emotions for their positive effect of overcoming difficult CD. Possibly this explains the mysterious enjoyment of sad music: it helps us to overcome CD of life's unavoidable disappointments, including the ultimate one, the knowledge of our finiteness in the material world. (**p.3**)

d. Masataka and Perlovsky (2013) demonstrated that consonant music helps "everyday" decision-making in the presence of cognitive interfering evidence, whereas dissonant music increases interference effects. Is music limited to a few emotions, or does every musical phrase evoke a different shade of emotion? Researchers take opposite sides of this issue (Scherer, 2004; Cross and Morley, 2008; Juslin and Västfjäll, 2008; Zentner et al., 2008; Koelsch, 2011; Juslin, 2013). As reviewed in Perlovsky (2012a,b), this does not affect the main argument of this paper for embodiment of abstract concepts through music. (**p.2**)

e. According to a theory of drives and emotions developed by Grossberg and Levine (1987), instinctual drives are neural mechanisms similar to internal sensors in the mind-body. They measure vital bodily parameters and indicate to the organism their deviations from safe ranges. Emotional neural signals connect these instinctual indications to decision-making brain-mind regions (Grossberg and Levine, 1987; Grossberg, 1988). The emotions felt are associated with these neural signals. For example, a bodily sensor-like mechanism measures sugar level in the blood and indicates when it is below a safe range. This generates emotional neural signals, which are felt as hunger and drive decision-making mechanisms to look for food. Instincts and emotions are the mechanisms of embodiment. (**p.1**)

4. *Musical Interaction reveals music as embodied language*

a. This proposal heeds Ian Cross' call for an investigation of music as an "interactive communicative process" rather than "a manifestation of patterns in sound" (Cross, 2014), with an emphasis on its embodied and predictive (coding) aspects (Clark, 2016; Leman, 2016; Koelsch et al., 2019). In the present paper our goal is: (i) to propose a framework of music as embodied language based on a review of the major concepts that define joint musical action, with a particular emphasis on embodied music cognition and predictive processing, along with some relevant neural underpinnings. (**p.1**).

b. The two components are usually called musicality and music: "Musicality in all its complexity can be defined as a natural, spontaneously developing set of traits based on and constrained by our cognitive and biological system. Music in all its variety can be defined as a social and cultural construct based on that very musicality" (Honing et al., 2015, p. 2, see also Huron, 2001). Life and social sciences often focus on the social nature of music (and language alike). In biology, for example, the three main evolutionary hypotheses about music, i.e., sexual selection (Miller, 2000; Fitch, 2006), parent-infant bond (Dissanayake, 2008; Malloch and Trevarthen, 2009) and group cohesion (Freeman, 2000; Dunbar, 2012), stress its intrinsically social character. Neurobiology thereby stresses the neuronal and hormonal underpinning of musicality. In line with these approaches, the present paper aims to suggest that the proper way to capture the social interactive nature of music (and, before it, musicality), is to conceive of it as an embodied language, rooted in culturally adapted brain structures. (**p.2**).

c. Firstly, consider entrainment, the phenomenon that brings a body rhythm to synchronize to a music rhythm (Clayton, 2012; Phillips-Silver and Keller, 2012; Moens and Leman, 2015). The sensorimotor prediction and adaptation mechanisms are supported by neuronal circuits in the posterior parietal lobe, premotor cortex, cerebro-cerebellum and basal ganglia, giving rise to the phenomenon of "groove" (Janata et al., 2012), suggesting that the same processes that cause bodily motion are involved in musical rhythm perception. (**p.5**).

d. Secondly, movement can also disambiguate a metric structure. In a couple of experiments Phillips-Silver and Trainor (2005) let infants be passively bounced or adults bend their knees to an ambiguous rhythmic pattern. These subjects' oscillations were set to stress either the second or the third beat, thus rendering either a binary or a ternary meter, as was manifest by their answers afterward, when asked to recognize which of two different patterns they moved on (while the adults answered verbally, the infants were observed attending to their preferred pattern between those two). (**p.5**).

e. Thirdly, timing is often not a matter of counting but rather a matter of moving, using outsourcing strategies by which limbs are moved, or choreographies are maintained in loops that don't require cognitive attention. Su and Pöppel (2012) showed that non-musicians rely more than musicians on their own movement in order to feel the pulse of a rhythmic sequence, missing it when such movements are not allowed. However, musicians can also rely on their internal clock to understand the sequence even without moving, thus demonstrating the importance of body movement, in particular where expertise is absent. In addition, it is worthwhile to remind that mirror neurons have been shown to depend also on such a sensorimotor expertise. For example, inferior-frontal and parietal areas typically involved in mirror activation, have been found to be more active (in a fMRI scan) in pianists, compared to naïve subjects, while observing piano-playing, compared to non-piano-playing, finger movements (Haslinger et al., 2005, see also Herholz and Zatorre, 2012). (**p.5**).

5. ***The Routledge Handbook of Embodied Cognition (pp. 81 – 89)***

a. Research on music perception has been guided by a paradigm that focuses on the anticipation of perceived structural components in music (Meyer, 1956; Huron, 2006; Honing, 2011). Anticipation is based on the ability to discern patterns (e.g. melodies, rhythms, timbres) that emerge from music through our senses. These patterns are compared with previously stored knowledge and used to generate expectations about music. The degree of match between the expected and the newly perceived pattern may then generate an outcome that is further processed by the emotional or motor system. (**p.81-82**).

b. In this case the sequence D7-D7- Dmin7 culminating to G7 creates tension, while the G7 going to Cmaj7 creates a relaxation. In short, perceptive processes evoke anticipations that lead to the emergence of tensions and relaxations, which ultimately lock into practices that give rise to signification, such as mood regulation, bonding and aesthetic experience. (**p.82**).

c. Moreover, music perception as a cognitive process can be described in terms of an abstract pattern-processing model that is independent from the physical carrier (body, hardware) of the processing. The processing consists of (1) a pattern-learning module that takes structures from music and organizes them into memory, (2) a pattern-prediction module that uses musical input and previously learned structures to predict the next input, (3) a pattern-association model that relates the consequences of the prediction to cognitive, emotive and motor outcomes. Levels (1) and (2) address syntactic processes, while level (3) addresses semantic processes. (**p.83**).

d. To sum up, the above examples show that some of the core cognitive phenomena associated with music perception require a foundation based on environmental constraints and particularities of the human body. Music perception, at levels that involve cognition, is not detached from the body or from the environment with which it interacts. At this point, it is important to state that there is little discussion about the fact that music perception involves the anticipation of emerging patterns. The question is more about the origins and the resources of anticipation, and to what degree that should be understood in relation to body and environment. Further arguments in favor of embodiment are based on recent findings that show dependencies of music perception on movement and emotions. (**p.83-84**).

e. More direct evidence of embodied cognition is provided by studies showing how changes in the motor system correlate with changes in perception of structural and expressive features of music. Thereby, two categories of changes in the motor system are generally addressed: one category relates to impairments of the motor system (i.e. motor disorders), the other to the development of the sensorimotor system (i.e. sensorimotor learning) (**p.84**).

6. ***Sweet Anticipation: Music and the psychology of expectation***

a. These results support two conclusions. First, naive listeners are sensitive to the frequency of occurrence of various tones and rate the most frequent tones as best fitting. Second, listeners who are enculturated to an appropriate pitch schema experience pitch sequences as evoking some preexisting schema and judge the various tones on the basis of their frequency of occurrence in the totality of their past listening exposure rather than the frequency of occurrence in a given tone sequence. (**p.170**).

b. When an event happens at an expected moment in time, the prediction response is positively valenced. This provides a positive reward for the heuristic used in the prediction and reinforces the use of such a heuristic in making future predictions. As noted in chapter 8, however, the positive emotion evoked by the accurate prediction is typically misattributed to the stimulus itself (the prediction effect). Since the downbeat represents one of the most predictable of event-moments, events that fall on the downbeat tend to evoke positive feelings. Colloquially, we say that the downbeat sounds nice. (**p.184**)

7. ***Sensorimotor grounding of musical embodiment and the role of prediction: A Review***

a. We advocated for more rigorous evidence to substantiate the "radical" embodiment thesis in the domain of music perception and performance (Mahon and Caramazza, 2008; Chemero, 2011; Wilson and Golonka, 2013; Kiverstein and Miller, 2015). Therefore, we provided a focused review presenting empirical evidence and computer models demonstrating that music perception and performance may be directly determined by the acoustics of sound and by the natural disposition and dynamics of the human sensory and motor system. On top of that, we have emphasized the role of long term processes involving learning and prediction in how humans interact with music. At the present, the exact nature of these processes is still a matter of ongoing debate—boldly between inferential and ecological accounts (Orlandi, 2012, 2014a)— yet to be fully determined. However, the collected findings suggest to consider short-term modality-specific processes serving perceptual or sensorimotor functions, and long-term learning and prediction processes as reciprocally determined and interacting; sensorimotor experience may lead to predictions, and predictions may shape sensorimotor engagement with our environment. (**p.6**).

b. It is of interest to link musical expression and the experience of affect to neurodynamical processes (Seth, 2013; Flaig and Large, 2014). In addition to musical expression, the study of the sensorimotor prediction loop can provide deeper insights into aspects of motivation and reward in music. Previous research demonstrated that dopaminergic activity, which relate to feelings of reward, encodes learning prediction errors (Waelti et al., 2001; Schultz, 2007; Hazy et al., 2010). This link of prediction processes to actual physiological responses, in this case dopamine responses, may contribute to our understanding of feelings of pleasure and reward that arise in music-based interactions (Chanda and Levitin, 2013; Zatorre and Salimpoor, 2013). Finally, all of the processes that occur within an individual in interaction with its sensory environment, may be dynamically linked to its social environment, leading to phenomena such as interpersonal coordination and synchronization (Repp and Su, 2013; Moran, 2014; D'Ausilio et al., 2015). (**p.7**)

8. What is entrainment? Definition and applications in musical research

a. In studying entrainment in music too we can find many different phenomena sharing some common features. In fact, I argue that it is important to distinguish between different manifestations of the phenomenon (see Clayton, in press), and that it is convenient to do so at three different levels, viz: 1. Intra-individual entrainment takes place within a particular human being. An important phenomenon at this level is the entrainment of networks of neuronal oscillators, which appear to be responsible for metrical perception (Large & Kolen, 1994; Large, 2000, 2010; London, 2004). Another aspect of intra-individual entrainment, as noted above, is the coordination between individual body parts (e.g. the limbs of a drummer). 2. Inter-individual/ intra-group entrainment concerns co-ordination between the actions of individuals in a group, which is essential for ensemble playing in any musical tradition. This is largely facilitated by the entrainment of attentional rhythms to auditory information, although other sense modalities – vision in particular – often also play a part. 3. Inter-group entrainment concerns the coordination between different groups. This is less widely recognised, being a rare phenomenon in Western art music, but in fact it is a widespread phenomenon (see Lucas et al., 2011). (**p.51**)

b. In summary, musical entrainment is recursive: individuals perceive and generate hierarchical temporal structures; they coordinate their actions within groups; and groups of people coordinate to form larger groups. It is also diverse: it can involve matching periods as well as hierarchical and polyrhythmic relationships, it is out of phase as often as it is in phase; and it can fall almost anywhere on the symmetrical-asymmetrical continuum. Musical entrainment is observed with periods in a range of roughly 100-2000 msec (corresponding to frequencies of 0.5-10 Hz), from the fastest beat to a typical measure (metrical and hypermetrical structures can however be considerably longer than 2000 msec, see e.g. Clayton, 2000, p. 87). (**p.52**)