Interoceptive Inference and Music: Integrating the neurofunctional 'Quartet Theory of Emotion' with predictive processing in music-related emotional experience

Shannon Proksch — MSc Student University of Edinburgh — shannon.proksch@outlook.com

PROBLEM

How and what emotional content is communicated or evoked by music constitutes a central question for music cognition. Emotion in music has been attributed to the extra-musical associations such as imagery and memory evoked responses, as well as violation and confirmation of musical expectancies based on statistical regularities of musical features from an agent's past experience (Meyer, 1956; Narmour, 1990; Huron, 2006). A number of studies have shown that the brain does, in fact, respond to the violation of expectation in various features of music, including rhythmic violations (Vuust et al.2008; Vuust and Witek, 2014), harmonic violations (Patel, 1998; Koelsch et al, 2005), and melodic-syntactic violations (Carrus, Pearce, & Bhattacharya, 2013). **While "it may indeed be by a flirtatious generation of disorder and its subsequent resolution that music communicates its emotional effect" (Fristion, 2013, 48), it remains unclear how responses to the violations of structural music expectations are related to the emotional content evoked by music.**

INTEROCEPTIVE PREDICTIVE PROCESSING (ACTIVE INTEROCEPTIVE INFERENCE)

In attempt to explain emotional experience, the predictive processing (PP) framework expands upon traditional, bottom-up, two-factor theories of emotion as passive physiological evidence-building and subsequent cognitive appraisal (James, 1884; Lange, 1885; Schachter & Singer, 1962), **by incorporating active inference based on predictive models of the causes of external and internal stimuli** (Seth & Critchley, 2013; Seth & Friston, 2016). In the same vein, thorough descriptions of music-related emotional experience need to appeal to the active processing of expectancies of both interoceptive (internal, bodily) information as well as exteroceptive (external, structural) information, an explanation that is offered by the minimization of prediction error achieved through active inference within predictive processing.

In predictive processing (PP), cascading predictions from top-down generative models are met with bottom-up prediction errors, which serve to either update the predictive models in our brain or to motivate action in the world (Clark, 2013). **PP models of exteroceptive signals enable inference regarding the states of affairs in the external world** which are most likely to cause a set of sensory states through the process of prediction error minimization (PEM). **Exteroceptive signals serve to enable perception and action in our external world.**

Similarly, PP models of interoceptive signals enable inference regarding the states of affairs in the internal world, which are most likely to cause a set of (interoceptive) sensory states through PEM. However, **interoceptive predictions serve control and regulation of physiological states with the goal of maintaining *homeostasis.***

- Model updating (perceptual inference)
- World updating (active inference)

EXPLANATORY UPSHOTS FOR MUSIC-RELATED EMOTIONS

Cognitively 'lower-level', short term physiological homeostasis

- ***Brainstem-centered system & basic arousal states*** – active inference within the brainstem and diencephalon-centered affect systems (model updating), contextualized by motor or attention effector systems associated with moving along to the beat of a song (world updating)
- Music, as an exteroceptive cue for interoceptive activity, already affects physiological arousal at the level of the brainstem (Koelsch, 2014). The location of the reticular formation, a part of the brainstem which is implicated in maintaining arousal and homeostatic functioning, is at a structurally advantageous location to mediate and integrate information from the cochlear and vestibular nerves. **Music, as an auditory event, activates both of these nerves through acoustic signals and affects the movement of fluid within the vestibular system. In fact, this vestibular fluid movement partially explains a drive to move your head along to the beat of a particularly rhythm of bass heavy song (Phillips-Silver, 2009). I posit that this movement along to a beat corresponds to the world-changing, active inference of matching physiological vestibular state by activating motor effector systems to move your body or head to the external, musical stimulus.**
- ***Diencephalon-centered system & nuanced emotion states*** – active inference within the diencephalon-centered system, especially the dopaminergic reward system (model updating), contextualized by peripheral physiological arousal and motor systems associated with 'chills' or 'frisson' at peak musical moments (world updating)
- While the insular cortex of PP accounts of emotion tracks and integrates specifically interoceptive processes, the brainstem-centered and diencephalon-centered systems of QTE perform integration of input from a variety of other affect systems. The reward circuit within the diencephalon-centered system is particularly activated in response to peak emotion in music, associated with a goose-bumps sensation called 'chills' or 'frisson' (Salimpoor et al., 2011). **As part of the peripheral physiological arousal effector system, frisson may occur as a result of the fulfillment, or release, of built up tension within a musical piece. I posit that this occurs because the world (the music) finally matches the expectation of the listener, and the sudden minimization of exteroceptive prediction error, results in the physiological response of goosebumps.**

Cognitively 'higher-level', long-term habitual homeostasis

- ***Hippocampus- and Orbitofrontal-cortex centered system & complex social emotion states***
- The habitual predictions of the hippocampus-centered and OFC-centered systems contributes to higher level habitual predictions which are learned early in life. **These habitual predictions are associated with the socio-cultural norms to which individuals are exposed early in life, resulting in firmly established models of such phenomena as the language and music they hear, and social environments in which they interact. I posit that if higher-level expectations are met *at the same time* as the lower-level physiological homeostatic expectations are met, then prediction error among these quartets of systems are being maximally minimized, leading to a sense of group belonging.**

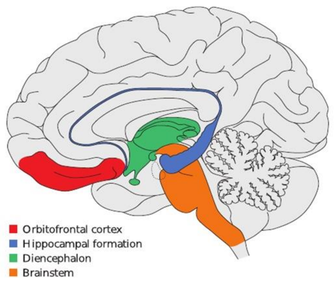
PROPOSAL

An explanation of the physical processes underlying music-related emotion must appeal to more than extra-musical associations and violation of musical features in the external auditory signal. A thorough account of music-related emotion must also take into account the physiological processes involved in the affective experience of the listener/performer.

- The Predictive Processing framework provides an explanation of the privileged role of expectation in action and perception, and active interoceptive inference specifically provides further emphasis on the physiological homeostatic drive underpinning emotional experience. PP importantly lacks a *detailed* description of the neural and physiological mechanisms enabling these homeostatic processes, and needs to go beyond explanations of mere positive or negative affective valence.
- The Quartet Theory of Emotion, as proposed by Koelsch et al. (2015) provides a detailed neural mechanism which takes into account how the brain and associated biological systems actually effect homeostatic maintenance and generate various aspects of in emotional experience, with the potential to address limitations within PP accounts of emotion and to further expand the framework of predictive processing as a whole.

An integration of the computational framework of predictive processing, paired with the neurofunctional mechanism offered by the QTE, provides a more nuanced explanation of emotional experience than either framework/theory offers alone—a claim which I justify through an explanation of how PP and QTE, together, enhance our understanding of music-related emotional experience, shedding light on the answer to the question of how and what emotional content is communicated or evoked by music.

QUARTET THEORY OF EMOTION + PREDICTIVE PROCESSING



Koelsch, S., Jacobs, A. M., Menninghaus, W., Liebal, K., Klann-Delius, G., von Scheve, C., & Gebauer, G. 2015. The quartet theory of human emotions: An integrative and neurofunctional model. *Physics of Life Review*, *13*, 1-27.

Model updating: Neurobiological affect systems

- ***Brainstem-centered***: monitors basic arousal states, structurally and functionally advantageous location to communicate with auditory system and vestibular system (both affected by music, as an acoustic stimulus), associated with fulfilling short term homeostatic needs (e.g. heart rate, breathing rate)
- ***Diencephalon-centered:*** including dopaminergic reward system, associated with peak-experience in music and the sensation of 'chills'
- ***Hippocampus-centered:*** 'higher' levels of cognitive processing such as memory and social behavior, rooted in homeostatic aims associated with fulfilling long term homeostatic needs such as attachment-related affects
- ***Orbitofrontal cortex-centered***: some level of traditional 'cognitive appraisal' through forming concepts and norms (which may not be propositionally available, such as implicit norms of musical structure)

World updating: Neurobiological effector systems – performing the 'active' component of active interoceptive inference

- Motor systems, peripheral physiological arousal systems, attention systems, memory systems

AUDIENCE SPOTLIGHT

A Hypothetical case study:

Let's reflect on the summer that Rylan, a college student in her late teens or early twenties, takes Henry, her grandfather, to a Rammstein concert. Henry wants to know what the young kids are listening to these days, and Rylan promised him that he would love their music. In fact, these industrial metal, hard rock performers make up Rylan's favorite band. Although she doesn't know German, **Rammstein concerts always leave Rylan feeling *energized and exuberant, as if she is right where she belongs*.**

Henry prefers country music oldies: Merle Haggard, Patsy Cline. But he loves his granddaughter, so he had happily tagged along for the experience. **This concert, however, leaves Henry feeling *agitated and uneasy, as if he distinctly doesn't belong.***

Arousal Emotions:

- As Rylan moves along with the beat of *Du Hast*, she is actively responding to the energy of the music and (unknowingly) to the movement of vestibular fluid in her ears via coordinated activity between the brainstem-centered affect system and motor effector systems. **This minimizes prediction error through matching the movement of fluid in her ears with the expected correlated movement of her body, maintaining an *energized* emotional state. Henry, however, standing still in staunch rebellion is remaining in a state of unresolved prediction error, maintaining his *agitation*.**

Nuanced Emotion:

- In the middle of *Du Hast*, all of the back-up instrumentation including the heavy beat of the guitar and drums disappears, leaving only the lead singer, a light synthesizer, and a sudden wealth of prediction error for a period of about fifteen seconds. When the beat 'drops' (to use the terminology of a rave), Rylan and the other Rammstein fans begin again to jump and headbang along to the beat. **Rylan experiences chills in reaction to this sudden minimization of exteroceptive prediction error and resetting of homeostatic arousal states at the peak emotional moment of the song. Henry does not have this reaction, perhaps because when the background instrumentation dropped out, he expected the song to be ending. In this case, the sudden reintroduction of the heavy beat instruments gives Henry an unexpected, and unappreciated surprise, increasing the deviation from his physiological homeostasis.**
- Henry's agitation over time is enhanced through not only an increased amount of prediction error, but also the high rate of prediction error as his body continues to deviate from its levels of homeostatic norms even after peak moments in the music. **This persistent prediction error over time causes him not just to feel a general negatively valenced arousal state, but contributes to Henry's nuanced emotional experience of *unease*. In contrast, Rylan's persistent minimization of prediction error, as well as return to an expected homeostatic range after peak moments, causes her to feel not just a positively valenced arousal state, but contributes to her nuanced emotional experience of *exuberance*.**

Complex Social Emotion:

- Rylan has been listening to music like Rammstein since she was in grade school, and she has well-developed habitual predictions surrounding heavy metal music, metal concerts, and metal fans. In addition, the positive effect of their music on her physiological homeostasis and corresponding emotional experiences has been well established. **Because prediction error is being minimized at both higher-level conceptual processing, as well as lower-level physiological processing, at this Rammstein concert Rylan feels as if she is *right where she belongs*.**
- Henry, however, has not grown up in an environment surrounded by metal music and has developed none of the firmly established models of the social environment surrounding metal music. **Neither his higher-level conceptual expectations, nor his lower-level physiological expectations are being met, causing Henry to feel as if he *distinctly doesn't belong* at this Rammstein concert and around these Rammstein fans.** Even if it is the case that Henry logically expects a certain social environment at this concert, and has developed a higher-level conceptual model of a metal concert environment, **because this higher-level prediction error is not being met at the same time as lower-level homeostatic prediction error, he does not have all of the necessary ingredients for a feeling of belonging.**

Scan QR code to download poster: