Faculty of Humanities
Institute of Communication and Cognitive Sciences

Linking musical metaphors, emotions and feelings of entrainment evoked by electronic music

From the physical properties of sounds to the experience of listening to music

Dissertation
Master of Arts in Cognitive Sciences
by

Stefano Politi

Supervised by Prof. Fabrice Clément and Prof. Dider Grandjean

January 2024

Contents

1 Acronyms and abbreviations

Ax designates the audio stimulus index: “A” stands for “audio” and “x” is replaced with a number (between 1 and 16), referring to a specific stimulus

dnb: Drum and Bass

FMA: Free Music Archive

GEMMES: Geneva Musical Metaphors Scale

GEMS: Geneva Emotional Music Scale

Hxy designates a hypothesis: “H” stands for “hypothesis”, “x” is replaced with a number (between 1 and 3), referring to a corresponding section of analysis, and “y” is replaced with a letter (“a”, “b” or “c”), referring to a specific hypothesis. Details are defined in the analytical framework.

ME: Motor Entrainment

MEQ: Musical Entrainment Questionnaire

MIR: Music Information Retrieval

MMA: Mimetic Motor Action

MMI: Mimetic Motor Imagery

PC: Principal Component

PCA: Principal Component Analysis

VE: Visceral Entrainment

2 Abstract

This paper explores the experience of listening to music in terms of relationships between metaphors and emotions. These relationships are considered from the theoretical prism of embodied music cognition and enactivism, where the body is considered central in the emergence of musical meaning and action tendencies are considered as essential components of perception. Through an online survey, 172 participants evaluated two types of electronic music (dance, ambient) in terms of metaphors, emotions and feelings of entrainment. Principal component analyses illustrated two main clusters of correlations: the first associated with elements which refer to overt motion (i.e. body movements), representative of dance music, and the second to covert motion (i.e. internal body sensations), representative of ambient music. These results were compared with those from a similar study on classical music, by Schaerlaeken et al. (2022), which showed consistent correlation patterns, thus suggesting similarities in how meaning is attributed to music through embodied experiences of motion. Furthermore, linear mixed-effects models were computed in order to explore the interaction between the physical properties of music and the experience of listening to music, as evaluated through the survey. The models showed that acoustic features associated with the perception of timbre, rhythm and structure influence the extent to which ambient and dance music are represented by each of the two clusters referring to overt/covert motion. The conclusions of this paper can be relevant for various applications within the fields of music classification, composition, performance and education.

Keywords: metaphors, emotions, entrainment, meaning, enactivism, music perception, embodied music cognition, electronic music, GEMS, GEMMES

3 Introduction

This paper aims to contribute to an embodied music cognition framework, from both a conceptual and an empirical standpoint. In doing so, it also establishes a bridge between physical properties of sounds and the experience of the listener.

The conceptual framework offers a perspective on how meaning emerges from the act of listening to music, where the body is considered central for the integration of auditory musical information. Firstly, this framework provides an operational definition of music, by establishing a connection between the physical properties of sounds and the psychological phenomenology of the perception of music. Secondly, it proposes an enactivist perspective of music perception, inspired by ecological theory and the conceptualization of action tendencies and perception as inseparable. Moreover, music cognition is viewed as fundamentally emerging from sensorimotor patterns mirroring the temporal dimensions of sounds. This is illustrated by the concepts of *rhythmic entrainment* and *mimetic behavior*. Finally, embodied musical meaning is conceptualized in terms of conceptual metaphors and emotions, both grounded in sensorimotor representations intrinsic to music perception and cognition.

The empirical study was composed of three phases. The first phase consisted in the extraction and analyses of a set of acoustic features from a small selection of electronic music excerpts. Dance music was associated with high timbre (in terms of frequency spectrum and energy), simple rhythms and regular structure, while the opposite was true for ambient music. The second phase consisted in the evaluation of these excerpts, via an online survey, according to metaphorical and emotional labels, as well as questions relative to bodily sensations (as expressed by subjective feelings of entrainment). Dance electronic music was associated with elements which refer to overt body movements, whereas ambient electronic music associated with elements which refer to covert bodily sensations. This supports an understanding of embodied music cognition, where the differences between dance and ambient electronic music mirror the differences between stimulative and sedative influences of music on the body of the listener. Furthermore, the relationships between metaphors, emotions and feelings of entrainment were consistent with those found by Schaerlaeken et al. (2022) with classical music, suggesting similarities in how music is understood through it's interaction with the body of the listener. The third phase consisted in creating models to test the interaction between the results from the previous two. Certain physical properties of music seem to influence the experience of the listener similarly, independently of genre or style, whereas others seem to have a specific influence with respects to the distinction between ambient and dance music. The correlations between metaphorical and emotional labels reflect the conceptualization of motion in music, where the sense of meaning emerges from the potential for bodily motion as a consequence of it's interaction with a particular musical motion.

4 Conceptual framework

4.1 What is music ?

In order to address the question of how we (humans) understand and are affected by the experience of listening to music, let us first consider what music is in terms of sound. The physical basis of sound is the vibration of particles in a specific medium (e.g. air), detectable by an individual's auditory system (Pinel, 2011). In response to these vibrations, the brain creates a "mental image" (Levitin, 2006), where objective physical attributes of sound waves (frequency, amplitude and complexity) correspond to the subjective perceptual attributes (pitch, loudness and timbre, respectively) (Pinel, 2011). Hearing sounds also implies perceiving changes in auditory information over time. As such, the perception of rhythm and melodic contour emerges from a sequence of sounds in time. In the context of music, the attributes of combinations of sounds are perceived by our cognitive system into higher-order concepts as meter, key, melody and harmony (Levitin, 2006).

Although there is no universal consensus on the definition of music, the notion of organized sounds seems to be fundamental (Kania, 2017). For the purpose of this work, I refer to the definition proposed by Levinson, where music refers to "sounds humanly made or arranged for the purpose of enriching experience via active engagement (such as through performing, listening, dancing), with the sounds regarded primarily as sounds" (Levinson, 1998). This allows to account for several aspects of music. First and foremost, music is considered as inevitably conceived in terms of sounds, and implicitly by their perceptual attributes. Secondly, music is conceived beyond the intrinsic virtue of the organization of sounds, and the related higher-order concepts, by including a phenomenological dimension: it requires that sounds be "regarded primarily as sounds", and thus that the organization of these sounds be heard as music. This allows to account for a range of environmental, artificial or linguistic sounds to be considered musical, as can be the case for example in *musique concrète*, computer music or spoken poetry. Furthermore, it also allows to account for inter-individual, context- and cultural-dependent differences in the evaluation of what is perceived as musical or not. Thirdly, the notion of active engagement highlight's the role of the individual in interaction with the organized sounds, where the action of listening is also considered as fully constitutive of the musical experience.

Finally, the composer's intention is also accounted for, whether the "purpose" of the music is to express or elicit emotions, to accompany context- or culture-specific activities, or for aesthetic appreciation. As such, even a soundless composition as John Cage's "4'33" can be considered as a musical experience (Davies, 1997), where the performer's silence leaves space to the organization of sounds to emanate from the audience and the environment in which it is performed.

4.2 Music perception and enactivism

The experience of listening to music involves the attribution of meaning to the musical information which is being perceived by the listener (Koelsch, 2011). Perception in itself can already be considered as a process of attributing meaning to the changes (or absence of change) of sensory inputs over time, by the organization of stable *percepts* – the psychological phenomenology, which is the outcome of perception (Zimbardo & Gerrig, 2002). Perception involves simultaneously bidirectional information processing mechanisms: *bottom-up* processes (data-driven), which extract and analyze relevant sensory information, and *top-down* processes (conceptually-driven), which interpret relevant sensory information in accordance to one's own experience, knowledge, motivation and cultural background (Zimbardo & Gerrig, 2002).

From the perspective of ecological theory, perception represents the process of understanding and adapting to one's environment, and thus according meaning to objects and events of this environment (Clarke, 2005). A central concept here is that of "affordances", understood as the range of opportunities for potential action that an object can provide - *afford* - to an organism (Gibson, 1966). Affordances are based on both the properties of the object, as well as the goals and capacities of the organism (Gibson, 1979). The detection of affordances underlies the perceptual experience, and has both biological (innate or learned through universal stages of development) and cultural (learned) origins (Gibson, 1966). The detection of affordances can be considered at the basis of the process of percept construal, mediating interactions between an organism and it's environment. Ecological theory of perception refers to an *enactivist* perspective, which adopts a strong interdependence between perception and action, conceptualizing them as "a unitary inseparable entity" (Schiavio, 2014). This coincides with the suggestions that "music perception is an active act of listening" (Koelsch et al., 2019, p. 1). Music perception is thus not considered as a passive process, but rather as being actively constructed by the listener (Matyja & Schiavio, 2013). As for perception in general, musical affordances depend on both object's properties (e.g. structural features of the music) and the individual's ability to engage with the object, i.e. music (e.g. sensorimotor capacities and affective sensibility) (Krueger, 2011). Among the various types of affordances that music can provide, which are also shaped by the listener's intentions, two notable examples are those of social coordination and emotion regulation (Clarke, 2005; Krueger, 2011). The opportunities for goal-oriented action expressed by musical affordances are here considered at the core of music perception, as "a dynamic process firmly rooted in the natural disposition of sounds and the human auditory and motor system" (Maes, 2016, p. 1).

Enactivism does not rely on traditional mental representations to account for cognition (Matyja & Schiavio, 2013; Schiavio, 2014; Witek, 2023). Instead, enacted cognition focuses on how meaning is inferred from sensorimotor action-perception coupling between the organism and the environment (Schiavio, 2014). This paradigm further develops the inextricable relation between cognition and perception, in a fundamentally interactionist view, where both are grounded in the interaction between body and environment (Leman et al., 2018). The concept of mind is redefined as "a distributed system that spans the body, the brain, and the environment, […] a consequence of mind being an active process driven by an organism" (Witek, 2023, p. 170). This views allows to conceptualize a philosophy of mind beyond a mind-body dualism, but rather as an inseparable entity constituted in interaction with its environment.

4.3 Rhythmic entrainment and mimetic behavior

Adopting an enactivist perspective implies adhering a broader view of embodied (music) cognition, whereby the body exerts a fundamental role in cognitive processing (of music) (Leman et al., 2018; Witek, 2023). A core aspect of this view is that of entrainment, the process by which two systems become synchronized as a consequence of their interaction (Trost et al., 2017). In the context of music cognition, it is the body of the listener which synchronizes to the rhythm of the music - *rhythmic entrainment*, characterized by four distinct, but inter-related, dimensions: perceptual, physiological, motor and social (Trost & Vuilleumier, 2013).

Perceptual entrainment refers to the integration of auditory patterns into a percept, through both bottom-up and top-down information processing (Trost et al., 2017). It is this percept which allows for the emergence of phenomena such as rhythm with it's associated affordances, from the detection of temporal information in a sequence of sounds. It has been proposed that predictive coding mechanisms are at the basis of perceptual processes coupled with action tendencies (Dell'Anna et al., 2021; Leman et al., 2018; Koelsch et al., 2019; Witek, 2023), where the brain is constantly constructing models and making assumptions about patterns which are being perceived. Action is intended as a means to minimize the prediction error between the model and the sensory information, a process referred to as active inference (Koelsch et al., 2019; Witek, 2023). Brain regions associated with motor activity are activated while passively listening to music (Anderson, 2020; Chen et al., 2008; Ross et al., 2022; Wu et al., 2016), which supports the claim that action is an inherently intertwined with perception, even when it is not overtly expressed. Music may "evoke motion in the listener not only in relation to strong rhythms, which induce clear motor responses (e.g. finger snapping or foot-tapping), but also to more subtle ones, which elicit more complex spatial trajectories that may or may not be acted out" (Shove & Repp, 1995).

Enactivism considers action, whether expressed or not, as a fundamental aspect of cognition. This claim resonates with the mimetic hypothesis (Cox, 2016), which postulates that we understand others' behaviors through covert and overt imitation of their actions, and that these same mechanisms also underlie our comprehension and experience of music. Cox distinguishes between *mimetic motor action* (MMA), to refer to imitation through overt body movements, and *mimetic motor imagery* (MMI), to refer to imitation through covert motor representations. The term representation is traditionally associated with a cognitivist account of cognition, where perceptual information is interpreted by the computation of mental representations (Matyja & Schiavio, 2013). Although Cox (2016) does not explicitly address this issue, his usage of the term is understood from a fundamentally enactivist perspective; closer to the notion of enactive sense making, rather than that of mental representations (Witek, 2023). In his view, *bodily representations* are simultaneously constituted by non-mimetic (perceptual) representations and mimetic (motor-related) representations (MMA, MMI), which emerge automatically as the individual accords attention to the sensory stimuli. This layer of attention, which Cox refers to as *mimetic* *engagement*, can be understood as the juncture between perception and cognition. Bodily mimetic representations extend from the motor-related brain areas, to the musculo-skeletal system, as well as other physiological systems (e.g. respiratory and cardio-vascular systems). Thus this concept refers simultaneously to physiological entrainment and motor entrainment, when actions are overtly expressed through movement (e.g. dancing).

At this point, it seems necessary to address the following question: how does the act of listening to music evoke mimetic behavior ? The process of imitation is to be understood as operating on two continuous dimensions. The first is the spectrum from mimetic (MMA and MMI) to non-mimetic behavior, and the second is that of imitable actions of observed entities (Cox, 2016). The latter ranges from observed human actions, to those of other animals, followed by inanimate entities (ex. the movement of the trees with the wind) and finally "non-actions" of inanimate entities (ex. architecture). In the case of music, it's understanding can involve mimetic behavior at different point of this spectrum. If music is being listened to in a social context, mimetic behavior can be directed towards actions of the performer and other people dancing. If music is not being listened to live, but still comprises instrumental or vocal elements, a similar mimetic behavior is present with respects to imagined actions necessary to production of these element. This may equally apply to cases where the production of sound is not attributed to another human (e.g. by playing an instrument or singing), but by other entities (animate or inanimate), for example an animal vocalization, a plant growing or a rock falling. Finally, mimetic behavior can also apply to non-actions, as for example when conceptualizing music as a landscape, where the listener' mentally imitates what it would 's bodily representations imitate what motion would be like within this landscape. At both ends of the spectrum, comprehension of music involves both non-mimetic representations, i.e. auditory perception of sound, and mimetic (overt and covert), i.e. motor-related representations emerging from the listener's attention to the sound. This paper focuses on the case where the experience of listening to music is neither live nor comprised of instrumental or vocal elements, through the particular case for electronic music. The subject is addressed through the prism of metaphors, emotions and entrainment, from an embodied and enactive perspective of cognition.

The last dimension of rhythmic entrainment is particularly important with respects to electronic music. Social entrainment refers to the interpersonal synchronization of movements, when music is experienced in a social context. This phenomenon promotes social interaction and coordination which has been valued of evolutionary relevance, considering music as a communication system parallel to that of spoken language which is present cross-culturally (Altenmüller et al., 2013; Cross, 2009; Dell'Anna et al., 2021).

4.4 Embodied musical meaning in terms of metaphors and emotions

Meaning is broadly understood as the psychological phenomenology "emerging as a property of animal-environment interaction the moment a sensory input becomes mapped to a behavioral output" (Wharton & Cornell, 2021, p. 182). Any perceptual experience is considered meaningful once affordances are cognitively mapped, enabling the animal to interact (or not) with it's environment. In line with the previous sections, the mapping of sensory input to behavioral output is fundamentally a process of sensorimotor activation, independently of whether an action is overtly expressed or not. In this view, listening to music conveys meaning when the listener attends to the "musical environment", through the process of mimetic engagement.

Music can be considered as a human communication system, parallel to verbal language (Cross, 2009; Dell'Anna, 2021). In this system however, the meaning that emerges in the listener can be completely independent of composers intentions: musical meaning is ideosyncractic (Koelsch, 2011). According to Koelsch (2011), musical meaning can emerge from both *extra-musical* and and *intra-musical* inferencial processes. The latter refers to the interpretation of structural relations between the elements of the music, while the former refers to the interpretation of musical information with respect to entities which are external to the music. This includes iconic meaning, where a musical pattern or form of resembles the sound of an object (e.g. "sounds like a thunderstorm"), the qualities of an object (e.g. "sounds round") or qualities of an abstract concept (e.g. "sounds colorful"), indexical meaning, indicating the presence of a psychological state, as an emotion of an intention, of an agent (e.g. "sounds happy"), and symbolic meaning which explicitly refers to an extra-musical association (e.g. national anthems). In both intra-musical and extra-musical meaning, an additional layer may emerge from the interpretation of the effects that music elicits in the listener: *musicogenic meaning*, referring to physical effects, emotional effects and self-related effects, associated with personal experience.

The empirical section of this paper focuses particularly on extra-musical (iconic and indexical) meaning, by “hearing music as…” (Schaerlaeken et al., 2019), in terms of conceptual metaphors which designate the cognitive process of understanding an abstract domain (e.g. music) in terms of another more concrete one (e.g. motion) (Johnson, 1997; Lakoff & Johnson, 1980). This process often relies on cross-domain mappings grounded in embodied experiences (Johnson, 1997; Zbikowski, 2008). Music and motion share fundamental properties related to temporal experience. As such, musical metaphors can also be viewed through the lens of conceptual blending theory (Fauconnier & Turner, 2008). In this case, the two domains are blended into a third (musical motion), which possesses the common properties of the two initial domains as well as supplementary concepts borrowed from either one. Motion is ubiquitous to the human experience and as such languages have evolved concepts to express experiences of motion. Such culture-specific concepts may be borrowed in order to understand the experience of music.

The conceptualization of emotions also heavily relies on metaphorical language (Lakoff & Johnson, 1980). In addition, both are temporally dynamic; both music and emotions unfold over time (Eliard et al., 2011). As such, certain concepts relative to motion can be relevant for both domains. On the premise that musical stimuli can elicit emotions (Scherer, 2004; Scherer & Zentner, 2001), music does not represent their specific content, but rather their dynamic form (Zbikowski, 2010). When music is conceptualized in terms of motion, that motion can also refer to emotional characteristics (Gabrielsson, 2016). It has been proposed that “the perception of emotion in music relies on the resemblance between music and bodily postures, gestures, or movements characteristic of emotions (e.g. sad music seems to walk slowly, as sad people do)” (Lauria, 2023).

An important distinction must be noted between *perceived* and *induced* emotions. The perception of emotion refers to the process of recognizing and representing the expression of feelings through music, whereas the induction of emotion refers to the experience of feelings, which may or may not coincide with those being expressed (Eliard et al., 2011; Pannese et al., 2016; Scherer & Zentner, 2001). According to Cox (2016), perception is a non-mimetic process, while mimetic engagement is associated with change in affective states, and thus induction. In the case of a mimetically engaged listener, cognitive inferential processes evaluating the auditory stimuli may or may not be appraised as affectively relevant (Sander et al., 2005). If so, the mimetic behavior may induce an emotional response. According to appraisal-driven theories, an emotional response is triggered when cognitive appraisals are synchronized with four other components: physiological, action tendencies, motor expression and subjective feeling (Sander et al., 2005).

The following empirical section focuses on the association between emotional and metaphorical labels, emerging from the conceptualization of the experience of listening to music. These labels are presumed to share common embodied experiences. The focus here is not whether an emotion is effectively induced in the listener, but rather that the process of sensorimotor mapping inherent to the perception of music underlies the foundation for the emergence of (emotional) musical meaning.

5 Analytical framework

The present study describes relationships between metaphors, emotions and subjective feelings of entrainment, evaluated while listening to electronic music. Furthermore, it explores interactions of such relationships with acoustic features of sound. The analytical approach involves a comparative examination with the results of Schaerlaeken et al. (2022) on classical music. Both studies collected data through an online survey which asked participants to rate a set of music excerpts on the basis of the following scales: the Geneva Musical Metaphors Scale (GEMMES) (Schaerlaeken et al., 2019) to assess musical metaphors, the Geneva Emotional Music Scale (GEMS) (Zenter et al., 2008) to assess musical emotions, the Musical Entrainment Questionnaire (MEQ) (Labbé & Grandjean, 2014) to assess feelings of entrainment, and the two-dimensional emotional model of valance and arousal (Russel, 1980). Furthermore, the music excerpts were previously analysed with the Music Information Retrieval (MIR) toolbox, a Matlab toolbox dedicated to extracting acoustic features (Lartillot & Toiviainen, 2007). These acoustic features refer to physical characteristics of sound, for example the mathematical descriptions of the frequency spectrum and rhythmic pulsation of sound waves. In the case of Schaerlaeken et al. (2022), these acoustic features were additionally related with a set of perceptual features (Aljanaki & Soleymani, 2018). This was not the case in this study, although the interpretations of the results were also considered in light of established associations between acoustic and perceptual features.

Thee results of Schaerlaeken et al. (2022) indicated two main groups of correlations. Group 1 included the metaphors of "Force", "Movement" and "Wandering", the emotions of "Power", "Tension", "Joyful activation" and "Wonder", as well as "Arousal", high subjective entrainment and perceptual features of articulation and intensity. Group 2 included metaphors of "Flow" and "Interior", emotions of "Peacefulness", "Tenderness", "Sadness" and "Nostalgia", as well as low subjective entrainment and perceptual features of melody and low dissonance. The 16 music excerpts used in this study represented two distinct categories (sub-genres) of electronic music: dance and ambient. The choice of these two sub-genres was based on the assumption that it would reflect two fundamental types of influence that music can have on body systems: stimulative or sedative (Bartlett, 1996; Gaston, 1951; Hodges, 2016). Each sub-genre was subdivided into two styles: techno and drum and bass (dnb), for dance music, and a rather atmospheric style (ambient 1) and a rather rhythmic style (ambient 2), for ambient music. These sub-genres and styles have been identified as relevant in MIR comparative studies, within the context of electronic music classification (Diakopoulos, 2009; Kriss, 2007; Lefaivre & Zhang, 2018).

The experimental design of this study aims at investigating several questions. Firstly, to what extent is acoustic feature analysis (with the MIR toolbox) an efficient method to classify electronic music sub-genres and styles? If so, are these distinctions based on stable descriptors independent on the genre of music, by comparison with classical music? My hypothesis are that both sub-genres (H1a) and styles (H1b) are distinguishable on the basis of common acoustic features descriptors, which are independent of the music genre (H1c). If this is the case, the resulting correlation patterns that emerge from the acoustic features analysis will show significant differences in how sub-genres and styles of electronic music relate to these patterns. Furthermore, these patterns will be similar to those observed with classical music (Schaerlaeken et al., 2022; Thibault De Beauregard, 2017). Secondly, do the groups of correlations (Group 1 and Group 2) observed by Schaerlaeken et al. (2022) also emerge when rating electronic music? If so, do differences in ratings for the two groups efficiently reflect the sub-genres and styles of electronic music? My hypotheses are that the groups of correlations observed by Schaerlaeken et al. (2022) are independent of the music genre (H2a), and reflect differences both between sub-genres (H2b) and styles (H2c). According to these hypothesis, the results of this study will show the same groups of correlations as Schaerlaeken et al. (2022), and these would correspond to the distinction between dance electronic music (Group 1) and ambient electronic music (Group 2), with more subtle differences in ratings for each group with respects to styles of each sub-genre. Lastly, do the patterns of acoustic features influence the subjective ratings of the music? I believe that this is the case (H3), and as such there will be an interaction effect between the acoustic features descriptors and the ratings of each of the two groups of correlations.

6 Method

6.1 Participants

The study included 172 participants (95 female, 72 male, 4 non-binary/third gender, 1 N/A), aged between 18 and 34 years old, out of which 116 are university students and 125 live in the french part of Switzerland (Appendix 1a, Table 1).

6.2 Materials

6.2.1 Audio stimuli

The 16 electronic music excerpts (Appendix 1b) originate from the Free Music Archive (FMA) (Defferrard, 2016). The two different sub-genres (dance and ambient) were represented by 8 tracks, subdivided into 2 different styles: dance music was split into 4 techno tracks and 4 drum and bass (dnb) tracks, while ambient music was split into 4 rather atmospheric tracks and 4 rather percussive tracks[1]. The music excerpts used in the survey represent ca. 30 seconds extracted from the original tracks, by using Audacity. The excerpts contained no lyrics and were intended to be as homogeneous as possible, in terms of containing few musical variations throughout their duration.

[1]The two dance styles correspond to FMA tags, whereas the two ambient styles were partly differentiated by ear, since FMA tags do not afford a clear distinction for ambient music.

The exact cutting meant to preserve the musical structure, which accounts for the approximative lengths, with 3 seconds of fade in and fade out at the start and at the end of each excerpt. The physical properties of these 16 music excerpts were characterized by set of 36 acoustic features (Appendix 2a, Figure 8) from the MIR toolbox.

6.2.2 Survey

The participants rated the music excerpts on the basis of several scales aimed at subjective evaluation of the music excerpts, in terms of of emotional, metaphorical, and entrainment labels. The two-dimensional model of valence and arousal (Russel, 1980) measures emotions on the basis of the subscales of “Valence”, as an indication of pleasantness, and “Arousal”, as an indication of intensity. The GEMS (Zenter, Grandjean & Scherer, 2008) measures musical emotions, and consists in 9 sub-scales: “Wonder”, “Transcendence”, “Tenderness”, “Nostalgia”, “Peacefulness”, “Power”, “Joyful Activation”, “Tension” and “Sadness”. The GEMMES (Schaerlaeken et al., 2019) measures musical metaphors, and consists of 5 sub-scales: “Flow”, “Movement”, “Force”, “Interior” and “Wandering”. The MEQ evaluates feelings of subjective entrainment, via 12 items which can subsequently be summarized into 2 main dimensions: Visceral Entrainment (VE) and Motor Entrainment (ME)[2] (Labbe & Grandjean, 2014). An additional variable, “Familiarity”, evaluated the extent to which participant’s were habituated to the style of music of each individual excerpts, on a 5-item Likert scale (from “Strongly disagree” to “Strongly agree”). Furthermore, participants were able express potential language difficulties and leave comments at the end of the survey. Appendix 1b contains all questions of the survey relative to the scales mentioned here.

6.3 Procedure

The survey was distributed, in the period between December 2022 and May 2023, through three different channels: printed flyers on public buildings of the University of Geneva (UniGE) and the Haute École d’Art et Design (HEAD), a mailing list to students of all faculties from the University of Neuchâtel (UniNE), and personal social networks either on social media (WhatsApp, Facebook and Instagram) or by direct

[2]Items 1, 6, 7, 8, 9, 11 and 12 load onto VE, while items 2 3, 4, 5 and 10 load onto VE.

contact. Appendix 1c contains both details concerning the building in which the flyers were displayed and the proportion of participants who responded via each one of the distribution channels (Table 2). The survey was administered through the Qualtrics online platform. Participants were instructed to respond in a quiet environment and preferably with the use headphones or earphones. After a few demographic questions, the survey consisted in listening to and rating 8 music excerpts. The initial 16 music excerpts were split into two equal groups, each one containing 2 excerpts of each style of music. Each participant was randomly assigned one of these two groups and the music excerpts were presented individually in a random order. After each one, participants were asked to assess its familiarity on a 5-item Likert scale and then rate each item of the scales described previously (GEMS, GEMMES, MEQ, valence and arousal), on a slider from 0 to 100. The listening period and the rating period were separated in order to avoid participants being focused on the evaluation instead of the experience of listening to music. The items to be rated were presented in one of two fixed orders, to control for a potential bias from the question's order. This order was kept the same throughout the whole survey, to maintain a similar structure for each music excerpt and thus facilitate rating for the participants. No reward was involved at completion and the estimated duration of the survey was of 15-20 minutes[3]. Data was still analysed even if the survey was not completed entirely. Appendix 1c displays the number of participants that rated each music excerpt (Table 3).

6.4 Statistical analyses

The following sections (results and discussion) are each composed by three subsections. These correspond to three distinct phases of analysis, with respects to the acoustic feature analysis (MIR toolbox), the subjective ratings data collected through the survey, and the combination of the two. Most of the variables' score distributions were not normally distributed, thus required non-parametrical statistical analysis, best described by a series of Wilcoxon Rank Sum tests.

Principal Component Analysis (PCA) was essential to reduce dimensionality among

[3]This was the limiting factor for the number of tracks per style that could be chosen. The pilot version of the survey included 8 excerpts per style (a total of 16 per participant), but it was too long for a non-remunerated study. Therefore, the total number of excerpts was cut by half, at the expense of loosing statistical power in the subsequent analyses

the acoustic features. This enabled to derive relevant groups of variables and to compare the two electronic music sub-genres (dance and ambient), as well as the four different styles (ambient 1-2, techno and dnb). The acoustic features loadings for each principal components also served as a tool for comparison between electronic and classical music (Schaerlaeken, et al., 2022; Thibault De Beauregard, 2017).

The second phase consisted in calculating correlations between the various items of the subjective ratings scales, as well as with the acoustic features principal components. These correlations served as basis for k-means clustering, which allowed to visualize the global structure of the relationships between all measures. A second PCA further confirmed the presence of clusters of correlations between the items of scales used to evaluate the music excerpts in the survey. The principal components served both as variables to compare the electronic music sub-genres and styles, as well as a tool for comparison with the correlation patterns found by Schaerlaeken, et al. (2022).

The third phase investigated interaction effects of acoustic features with subjective ratings, by calculating correlations between the principal components of both PCAs and computing linear mixed-effects models.

7 Results

7.1 Acoustic features

The PCA revealed that the first two components explained 57.4% of variance, 38.3% and 19.1% respectively (Appendix 2a, Figure 9, Tables 4-6). Each component was characterized by a subset of acoustic features, based on their saturation levels (Appendix 2a, Figure 10). The first component (MIR PC1) is strongly associated with spectral features of timbre (spectral centroid, spectral spread, spectral skewness, spectral kurtosis, spectral flatness, spectral entropie, rolloff, brightness 1000, brightness 1500, brightness 3000, zerocross, spectre flux, cepstre flux) and shows a strong negative association with tone (key clarity). In addition, it is also, to a lesser extent, positively associated with a rhythm/metric feature (tempo change) and negatively with dynamics/structure (novelty) features. The second component (MIR PC2) is strongly negatively associated with rms energy, and to a lesser extent with roughness, while it is strongly positively associated with cepstre centroid. In addition, it is also negatively associated with rhythm/metric features (pulse clarity, metrical strength) and positively with dynamics/structure (novelty) features.

The respective scores for these two components characterized the 16 music excerpts (Figure 1; Appendix 2a, Figure 11). This allowed to compare the two sub-genres (dance and ambient) (Appendix 2a, Tables 7-8) and the four different styles of electronic music (Appendix 2a, Tables 9-12). Whether comparing sub-genres or styles, the groups contained too few data points (8 and 4 respectively) to effectively determine if the score distributions were normally distributed. Thus, it was most adequate to compare groups through a series of non-parametric pairwise Wilcoxon Rank Sum tests. MIR PC1 draws a significant distinction between the two sub-genres (p-value < 0.01). Ambient music is characterized by negative values, with the exception of slight positive values for A6 and A8, whereas dance music is characterized by positive values. However, there was no significant difference between ambient styles (p-value $= 0.069$), even though visually ambient 1 is characterized by strong negative values, compared to ambient 2 excerpts, which are generally characterized by less negative and even slightly positive values. There was also no significant difference between techno and dnb (p-value $= 0.486$). Differences between techno or dnb and with both ambient styles were all significant (p-value $= 0.043$). MIR PC2 also draws a significant distinction between sub-genres (p-value < 0.01). Ambient music was characterized by positive values, with the exception of a slightly negative value for A8, whereas dance music was characterized by negative values, with the exception of slightly positive values for A13 and A16. The differences between styles were significant (p-value $= 0.043$), with the exception of the one between dnb and ambient 2 (p-value $= 0.2$) and the one between the two ambient styles (p-value $= 0.069$).

There were striking similarities between the two principal components and those found by Schaerlaeken, et al. (2022) and Thibault De Beauregard (2017). In all three studies, PC1 showed strong associations with spectral features of timbre, while PC2 with rms and roughness. What is novel, in this study, with regards to timbre is that spectre flux was positively associated with PC1, instead of being negatively associated with PC2, and cepstre mean was barely represented, instead of being positively associated with PC1. Furthermore, with respects to tone, only this study showed a strong negative association with key clarity for PC1, while Schaerlaeken et al. (2022) found a negative association with hcdf. With respects to rhythm/metric, tempo change is positively associated with PC1, while metrical strength and pulse clarity are negatively associated with PC2. In Schaerlaeken et al. (2022), tempo change is not relevant, while pulse clarity is negatively associated with PC1 and event density is, in addition, positively associated with PC2. Finally, in this study, novelty features are relevant for both components, whereas they are irrelevant in Schaerlaeken et al. (2022) and are only mentioned in a third component in Thibault De Beauregard (2017).

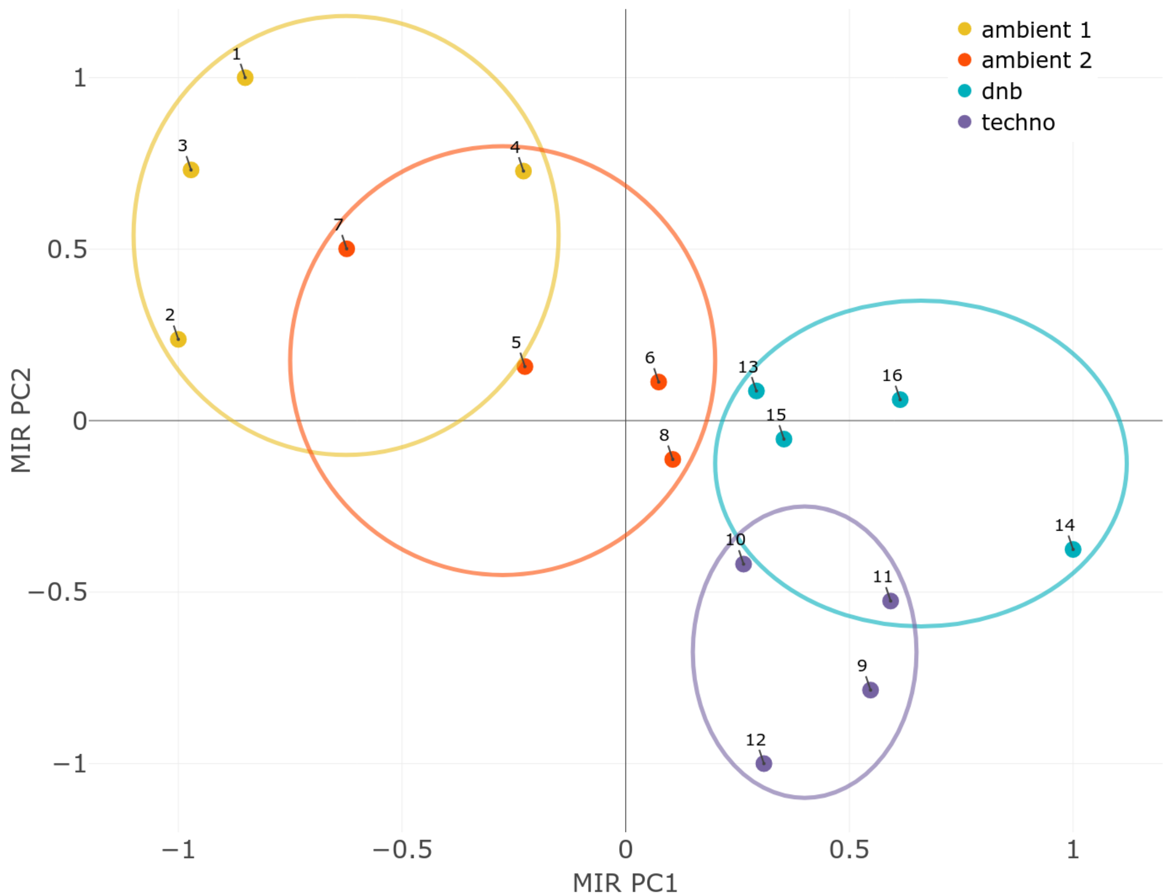

Figure 1: Acoustic features PCA scores for each music excerpt, ranging from -1 to 1. The numbers indicate the excerpt id and the coulours indicate the style. The scores belonging to the same style are grouped together within an ellipse.

7.2 Subjective ratings

Each music excerpt being characterized by a total of 19 variables from the participant's subjective ratings: 1 categorical ordinal variable (familiarity) and 18 numerical continuous variables (5 from GEMMES, 9 from GEMS, 2 from MEQ, valence and arousal), each ranging from 0 to 100. The score distributions of the 12 items of MEQ were not normally distributed, thus they were summarized into the two main variables (MEQ_VE and MEQ_ME) by taking the median values of their respective items (cf. Methods: Survey). Furthermore, all variables were normalized on a range from -1 to 1, including the acoustic features principal component scores and "Familiarity", which was first converted to numeric form. When considering all 16 music excerpts together, the order in which questions were presented had an effect on 13.5% of the ratings [4]. Nonetheless, these ratings were merged together in order to compute Spearman correlations between all variables (Figure 2), which ranged from r = -0.79 between the two acoustic features principal components (MIR PC1 and MIR PC2) to r = 0.84 between the two MEQ variables (MEQ_VE and MEQ_ME). The latter showed no particular differences between each other in correlations with all other variables.

[4] Percentage = Number of variables with Wilcoxon Rank Sum Test p-value < 0.05 (n = 41) divided by the total number of variables per the number of excerpts (N = 16*19 = 304).

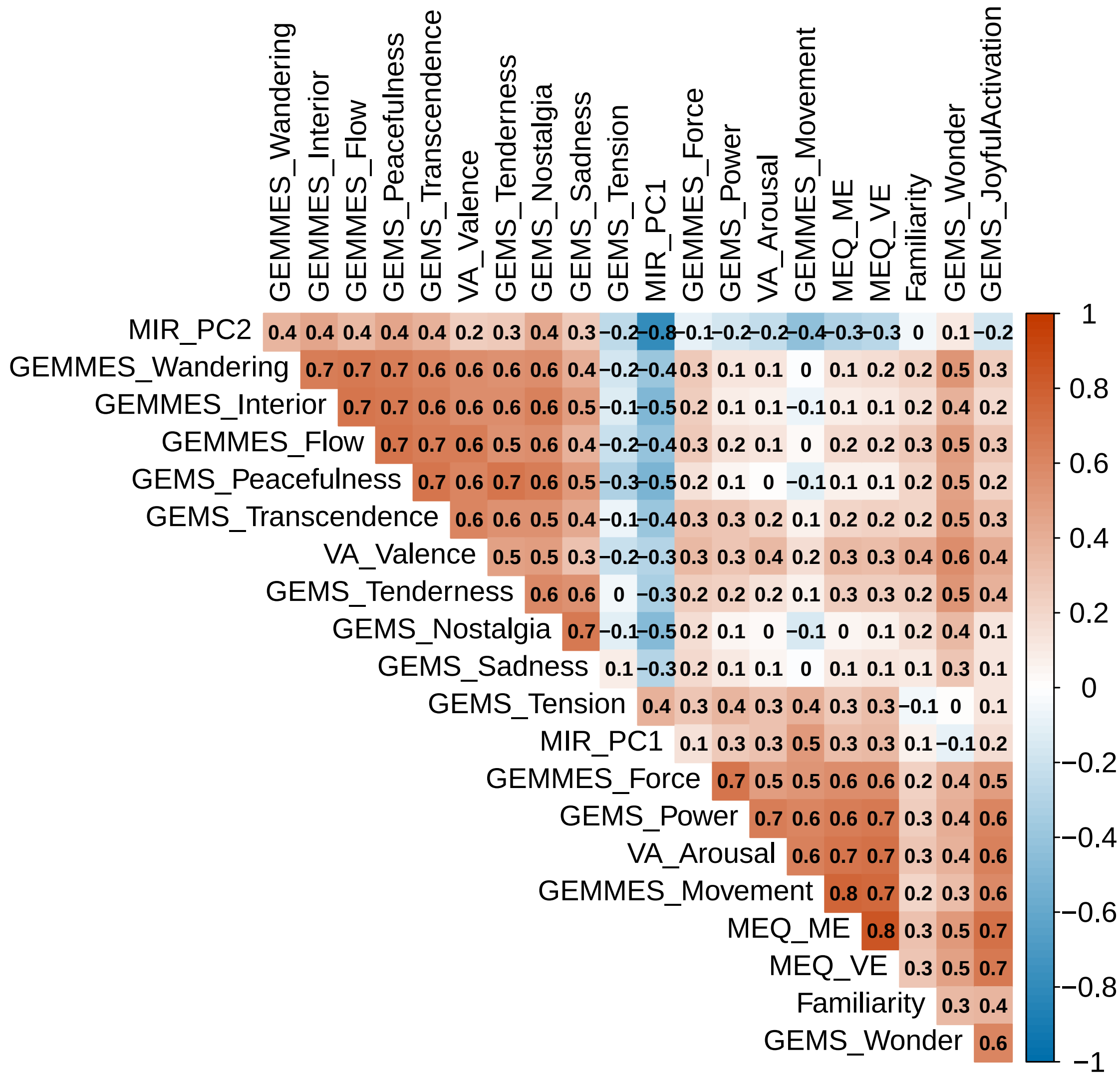

Figure 2: Spearman correlations matrix of subjective ratings variables and acoustic features PCs.

Visualizing these correlations through k-means clustering established two main clusters (Figure 3; Appendix 2b, Figure 12). The first cluster (red) is associated with both variables of subjective entrainment (MEQ VE and MEQ ME), arousal, metaphors of “Movement” and “Force”, emotions of “Power”, “Joyful Activation” and “Tension”, as well as MIR PC1 and “Familiarity”. The second cluster (blue) is associated with valence, metaphors of “Wandering”, “Flow” and “Interior”, emotions of “Wonder”, “Tenderness”, “Peacefulness”, “Nostalgia”, “Sadness” and “Transcendence”, as well as MIR PC2. Not all variables of these two clusters are equally represented and some variables are more strongly related with each other than others. “Tension” and MIR

PC1 are relatively weakly correlated with the other items of the first cluster, with the exception of MIR PC1 and "Movement" (r = 0.52), but are most notably negatively correlated with all items of the second cluster. "Familiarity" is relatively weakly correlated with all elements, with the strongest correlation being with "Valence" (r = 0.41) although they belong to two different clusters. "Valence" also correlates well with "Joyful Activation" (r = 0.44) and "Wonder" (r = 0.57), which both correlate relatively strongly together (r = 0.61) although also belonging to two different clusters. However they are also well associated with all other elements of their respective clusters. Finally, "Valence" is positively correlated with all variables, except "Tension" and MIR PC1.

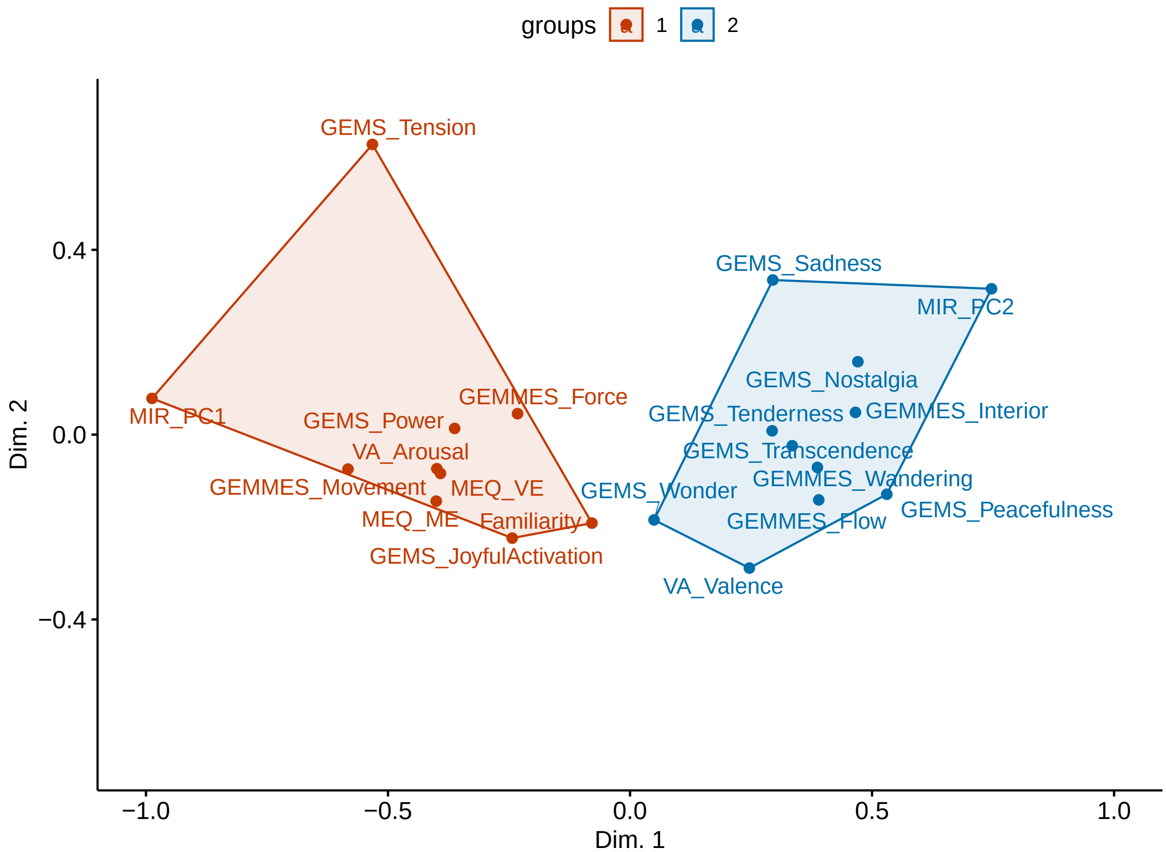

Figure 3: K-means clustering of subjective ratings variables and acoustic features PCs.

The PCA, computed on subjective ratings variables, revealed that the first two components explained 60% of variance, 31% and 29% respectively (Appendix 2b, Figures 13-14, Tables 13-15). The first component (ratings PC1) is associated with all variables from the second cluster (represented in blue), although "Sadness" and "Wonder" are slightly less represented. The second component (ratings PC2) is associated with all variables from the first cluster (represented in red), although "Tension" and "Familiarity" are less represented. The only exceptions concern the variables "Wonder" and "Familiarity", which are equally represented in both components. Furthermore, "Tension" is weakly positively associated with PC2 and weakly negatively associated with PC1.

The respective scores for these two components characterized the 16 music excerpts (Figure 4; Appendix 2b, Figure 15). This allowed to compare the two sub-genres (dance and ambient) (Appendix 2b, Tables 16-17) and the four different styles of electronic music (Appendix 2b, Tables 18-21). Whether comparing sub-genres or styles, all groups were not normally distributed. Thus, it was most adequate to compare groups through a series of non-parametric pairwise Wilcoxon Rank Sum tests. The first component draws a significant distinction between the two main sub-genres (p-value < 0.01). Ambient music was characterized positive values, with the exception of the slight negative value for A6, whereas dance music was characterized by strong negative values. Differences between styles were also significant for all comparisons (p-value < 0.01 between sub-genres; p-value = 0.011 within sub-genres). The second component also draws a significant distinction between sub-genres (p-value < 0.01). Ambient music was characterized by strong negative values, whereas dance music was characterized values close to zero, with the exception of a relatively strong negative value of A12. Once again, differences between styles were also significant between the two ambient styles and for all other comparisons (p-value < 0.01), except between techno and dnb (p = 0.275). The raw data showed a high proportion of ratings close to 0, on the original non-normalized scale that was proposed on the survey (0-100). As a consequence, the spread of the score distributions showed a strong positive skewness, for each negative association of a sub-genre or style with a principal component.

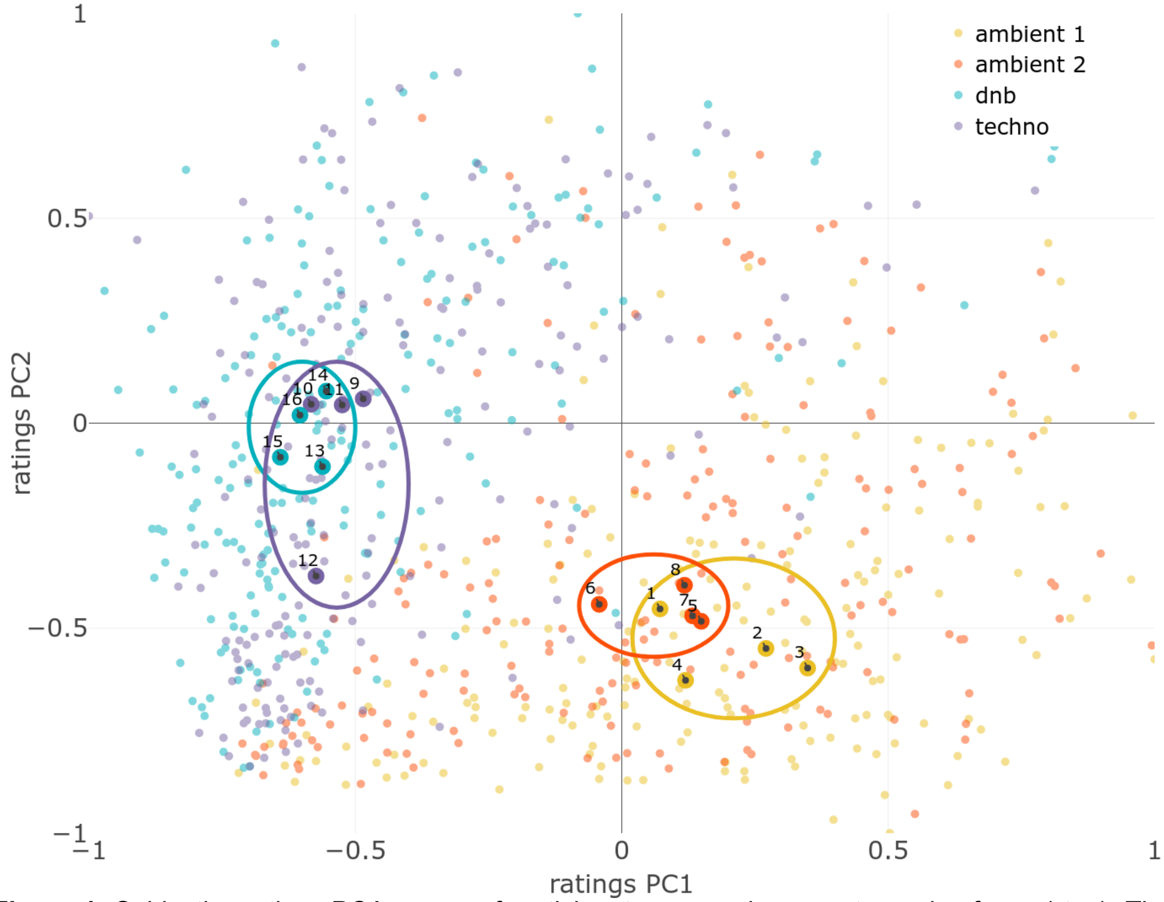

Figure 4: Subjective ratings PCA scores of participants per music excerpt, ranging from -1 to 1. The numbers indicate the excerpt id and the coulours indicate the style. The small dots represent individual participant PC scores for the respective excerpt, the larger dots represent the median of the PC scores distributions for that excerpt. Those belonging to the same style are grouped together within an ellipse.

The results of the k-means clustering, and of this second PCA, coincide with the results found by Schaerlaeken et al. (2022) with classical music. The first (red) cluster of correlations (ratings PC2) corresponds to the first group of correlations (Group 1). This group includes metaphors of “Force”, “Movement” and “Wandering”, emotions of “Joyful activation”, “Power”, “Tension” and, to a lesser extent, “Wonder”, as well as feelings of subjective entrainment and “Arousal”. When comparing this group of correlations with the results of this study, the only exception is that of the metaphor of “Wandering” which correlated best with the second cluster instead. This exception aside, the second (blue) cluster of correlations (ratings PC1) corresponds to the second group (Group 2) defined by Schaerlaeken et al. (2022). This group includes metaphors of “Flow” and “Interior”, and emotions of “Peacefulness”, “Tenderness”, “Sadness” and “Nostalgia”. An additional exception here is the emotion of “Transcendence” which, in this study, also correlated well with this group.

7.3 Interaction between acoustic features and subjective ratings

The previous two PCAs provide 4 dimensions: 2 acoustic features PCs and 2 subjective ratings PCs. These 4 components allowed to further investigate differences and similarities between conditions, styles and individual music excerpts. Spearman correlations (Figure 5) ranged from r = -0.79 between the 2 acoustic features components (MIR PC1 and MIR PC2) and r = 0.5 between ratings PC1 and MIR PC2. In between, ratings PC2 and MIR PC1 are also positively correlated (r = 0.42), while ratings PC1 and MIR PC1 are negatively correlated (r = -0.58) and so are ratings PC2 and MIR PC2 (r = -0.33). The two subjective ratings components are not correlated between each other (r = 0.02 for ratings PC1 and ratings PC2).

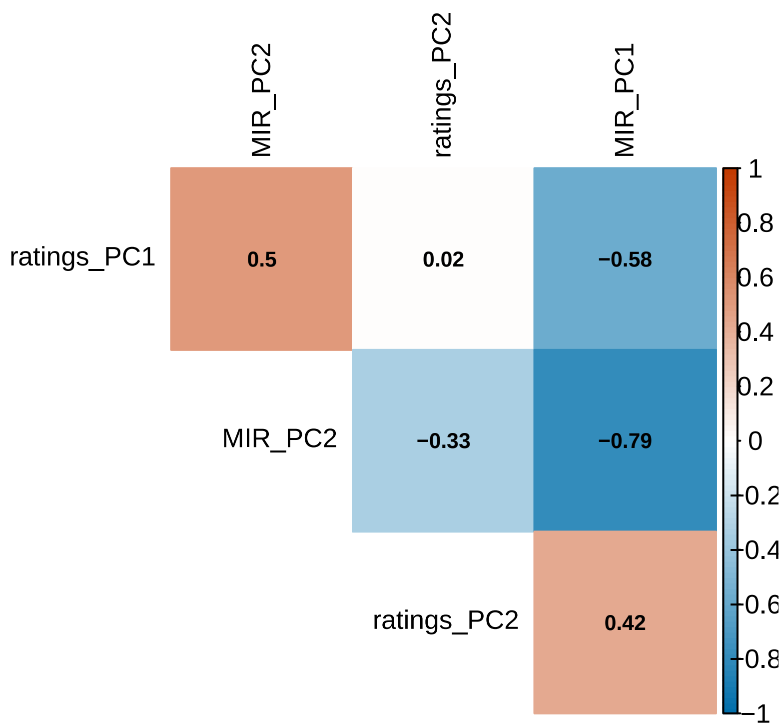

Figure 5: Spearman correlations matrix of PCs of subjective ratings and acoustic features.

The following step of analysis consisted in computing linear mixed-effect models to test for interactions between the acoustic features PCs and the subjective ratings PCs. Since differences between individual styles (ambient 1 vs. ambient 2; techno vs. dnb) were not significant for all PCs, the models only focused on the sub-genres (ambient and dance). Considering each acoustic features PC separately (MIR PC1, MIR PC2), in both cases the models which included the interaction with the music condition outperformed those without (MIR PC1: Chi squared = 183, p-value < 0.01, AIC_{PC1}∗ratings = 1260, AIC_{PC1}∗ratings∗condition = 1080, BIC_{PC1}∗ratings = 1290, BIC_{PC1}∗ratings∗condition = 1140; MIR PC2: Chi squared = 409, p-value < 0.01, AIC_{PC2}∗ratings = 1500, AIC_{PC2}∗ratings∗condition = 1090, BIC_{PC2}∗ratings = 1530, BIC_{PC2}∗ratings∗condition = 1150) (Appendix 2c, Tables 22-23). In both cases the music condition alone, as well as the interaction between the condition and the ratings PC are significant (p-value < 0.01). However, in the case of MIR PC1, the interaction with ratings PC is significant (p-value < 0.01) but the triple interaction is not (p-value = 0.705) (Appendix 2c, Tables 24-25). The opposite is true in the case of MIR PC2 (p-value = 0.9 and p-value < 0.01, respectively). Finally, the last step of analysis consisted in computing trends for each combination of condition and ratings PC, with respects to each of the two acoustic features PC, and compared the slope to that of a horizontal line (Figure 6) (Appendix 2c, Tables 26-27). For increasing values of MIR PC1, ambient PC1 ratings decrease (p-value < 0.01) while PC2 ratings increase (p-value < 0.01). For increasing values of MIR PC1, dance PC1 ratings decrease slightly but not significantly (p-value = 0.87) while PC2 ratings increase (p-value < 0.01). For increasing values of MIR PC2, ambient PC1 ratings increase but not significantly (p-value = 0.28) while PC2 ratings decrease (p-value < 0.05). For increasing values of MIR PC2, dance PC1 ratings increase (p-value < 0.5) while PC2 ratings decrease but not significantly (p-value = 0.069). It is important to note that for both acoustic features PCs, each music excerpt has distinct scores which is confined within a certain range specific to the condition (i.e. ambient: MIR PC1 < 0.1, MIR PC2 > -0.1; dance: MIR PC1 > 0.1, MIR PC2 < 0.1) (Figure 1).

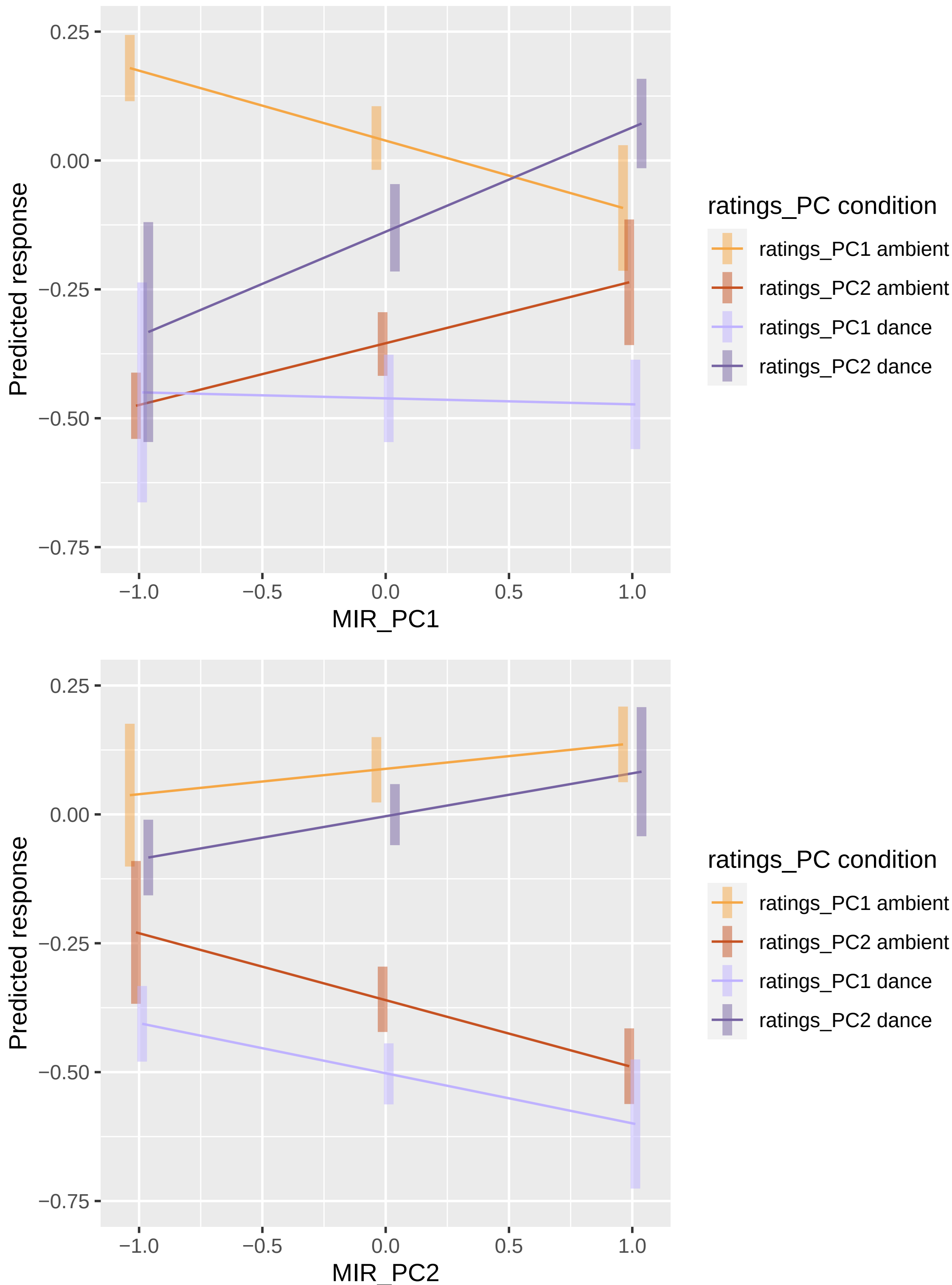

Figure 6: Trends for linear mixed-effects models of interaction between ratings PCs and music condition (sub-genre), along increasing values of MIR PC1 (above) and MIR PC2 (below).

8 Discussion

8.1 Acoustic features

The scores of individual music excerpts for both acoustic features PCs draw a significant distinction between ambient and dance music, however not necessarily between styles. Differences between ambient styles were not significant for either component. This may be a consequence of the fact that the original tracks were described with certain common FMA tags, and thus not representing sufficiently different styles. An alternative explanation is that only certain excerpts were not well fit into one or the other category. If willing to determine styles on the basis of acoustic features, an option would be to consider A7 as part of ambient 1, while leaving the other three excerpts of ambient 2 together. Alternatively, the two ambient 2 excepts with slightly positive values for PC1 (A6 and A8) could be categorized separately from the other two excerpts (A5 and A7), which could instead be categorized alongside the excerpts from the ambient 1 style. Interestingly, the lack of a significant difference for MIR PC2 between dnb and ambient 2 confirms the relevance of the latter category. This component is particularly associated with acoustic features related to energy, rhythm/metric and dynamics/structure. Thus the non-significant difference between dnb and ambient 2, in comparison to ambient 1, indicates that there is indeed a relevant criteria of distinction based on percussive (vs. atmospheric) elements (cf. Materials: Audio stimuli). With regards to dance music, the lack of significant differences for PC1 indicates that acoustic features of timbre are not a relevant predictor of style between techno and dnb. In contrast, PC2 does show a significant difference between the two styles. This can probably be explained by the dnb excerpts having a more irregular rhythm and structure, in comparison to the techno excerpts.

These results emerge from statistical analyses of merely 4 music excerpts per style, which is a too small sample size to establish a solid interpretation. However, they do begin to sketch a pattern of acoustic features and how they interact with electronic music styles. Based on these results, the first hypothesis (H1a) is confirmed, since the two sub-genres were significantly different from each other, but the second (H1b) is not, since the styles within each sub-genre were not significantly different from each other. This can partly be explained from a larger context of music classification. MIR technology is currently used to automatically classify enormous databases of music, but often fails to effectively reflect the human conception of genre, since the algorithms are based solely on acoustic feature analysis and omit entirely the cognitive processes involved in the formation of such categories (Aucouturier & Bigand, 2013; Siedenburg et al., 2016). Furthermore, music (sub-)genres and styles are not objective categories, but are cultural constructs with no clear cut boundaries and which are continuously changing as functions of arbitrary influences by artists, music industries and subjective evaluations (Aucouturier & Pachet, 2003; Defferrard et al., 2016; Nie, 2022). In particular, electronic music is an extremely rapidly growing genre with a rich variety of sub-genres (and styles) and it is often misrepresented and inaccurate in automated MIR classifications (Chen, 2014).

Although MIR technology is a powerful tool for the physical analysis of sound, it's characteristics are profoundly mathematical, which makes it a difficult task to derive cognitively or psychologically relevant interpretations (Aucouturier & Bigand, 2013; Siedenburg et al., 2016). In an attempt to do so, some insights can be deduced from the comparison of this study's PCs with those found by Schaerlaeken et al. (2022) and Thibault De Beauregard (2017) on classical music. Even though the music excerpts belonged to different music genres, there were striking similarities across these studies, which confirms the third hypothesis (H1c). Although it is beyond the scope of this study to provide an in depth interpretation of these acoustic features, the comparison of the results between these studies enables a preliminary interpretation. Low-level (acoustic) features are objective characteristics of sound, high-level features (ex. genre, style, mood) are complex subjective concepts, while perceptual (mid-level) features represent the gap in between (Aljanaki & Soleymani, 2018). Schaerlaeken et al. (2022) found Group 1 to be related to perceptual features of articulation and intensity, and Group 2 to be related to melody and low dissonance. The results of the subjective ratings correlations of this study are explored in the next section, thus for the moment the comparison is made with the assumption of dance music being rather stimulating and ambient music being rather sedative. "Stimulative music refers to music characterized by faster tempi, louder dynamics, angular or disjunct melodies with wide pitch ranges, and accented or staccato articulation. Rhythm is emphasized over melody and harmony. Sedative music has opposite characteristics — it is softer, slower, smoother, and more legato." (Hodges, 2016, p. 184). Although, Aljanaki & Soleymani (2018) emphasize the fact that acoustic features cannot efficiently predict perceptual features, they also point out certain correlations. Since this study is missing a perceptual features analysis, the following interpretation is based on these correlations. Firstly, articulation seems to be related with various spectral features, and as such PC1 may a good predictor for it. Secondly, dissonance seems to be related with roughness, and thus PC2 may be a good predictor this time. Finally, key clarity and hcdf may contribute the perception of tonal stability, and pulse clarity seems to be related to rhythmic stability. In both cases, together with other metric/rhythm and novelty features, this study shows differences in PC loadings in comparison to Schaerlaeken et al. (2022). This may indicate potentially relevant predictors of distinction in the perception of electronic or classical music. In summary, articulation and dissonance seem to be related to stimulative music, independently of it's genre. Whereas, patterns in tonal stability, rhythmic stability and dynamic/structure seems to distinguish electronic music from classical music.

8.2 Subjective ratings

The strong similarities between the results of this study and those of Schaerlaeken et al. (2022) indicate that indeed there are stable genre-independent correlations between subjective ratings of metaphors, emotions and entrainment. This effectively confirms the hypothesis H2a. Ambient music was found to be strongly negatively associated with the second principal component (ratings PC2), reflective of metaphors and emotions associated with overt movements, as well as feelings of entrainment and arousal. In contrast, dance music was found to be strongly negatively associated with the first principal component (ratings PC1), reflective of metaphors and emotions associated with covert bodily sensations. These differences validate the next hypothesis (H2b). With respects to four styles of music, the difference in ratings between techno and dnb was not found to be significant. This result is sufficient to invalidate the hypothesis H2c. It is interesting to note that ratings for all variables are spread along the entire range (from -1 to 1) (Appendix 3a, figures 16-19). This most likely indicates important influence of inter-individual and context-dependent differences in the evaluation of the subjective experience of listening to music. However, this spread is not uniformly distributed, as a consequence of the high proportion of ratings close to zero (on original range from 0 to 100 that was proposed on the survey), as was also the case in previous studies (Schaerlaeken et al., 2019, 2022). As a consequence, the distinction between dance and ambient music is based on strong negative association with each respective cluster, rather than on positive associations with their counterpart. An interpretation of this phenomenon is that GEMS and GEMMES may be perceived as categorical, instead of continuous, scales, and thus potentially prompting participants to rate the music according to only one or few categories instead of adopting a more complex mixed-feeling representation (Schaerlaeken et al., 2022). A complementary interpretation is that of music's general "ineffability", where participants may find it difficult to express their emotional experience with respects to the tools provided to them (Zentner & Eerola, 2010).

8.2.1 Metaphors

"We experience musical events as moving through time, with various musical forces influencing the direction and shape of the musical motion" (Johnson, 1997, p. 99). According to Cox (2016), a *musical event* "includes anything that can be anticipated, then felt as present, and then remembered, from a single chord to an entire concert" (p. 124), while *musical motion* refers to the "metaphoric conceptualization of non-spatial change in the auditory stimulus in terms of spatial change (i.e., motion)" (p. 132). *Temporal motion*, of which musical motion is a subset, is inevitably conceptualized in spatial terms (Casasanto & Boroditsky, 2008; Fauconnier & Turner, 2008; Nuñez & Cooperrider, 2006), which provides a perceptually concrete source domain (space) to an abstract target domain (time) (Boroditsky, 2000; Lakoff & Johnson, 1999). Although significant cultural variations exist about the specific ways in which humans conceptualize time, the underlying conceptual metaphor's source domain of space is consistent cross-culturally (Dancygier & Sweetser, 2014; Nuñez & Cooperrider, 2006). The experience of temporality is unavoidably understood in terms of space and motion, and thus also in terms of the phenomenological experiences of anticipation (location: *ahead*; motion: *approach*), presence (location: *here*; motion: *arrival*) and memory (location: *behind*; motion: *departure*) (Cox, 2016). Based on Conceptual blending theory (Fauconnier & Turner, 2008), with one domain being that of spatiality and the other that of musical events, which both share the phenomenological experiences of anticipation, presence and memory, we can establish the blended domain of musical motion and space, which thus inherits the respective concepts of location and motion. Following from this conceptually blended space, we can consider all 5 metaphors from the GEMMES, as defined by Schaerlaeken (2019), to be conceptual metaphors of motion. "Force" corresponds to the MUSIC AS A MOVING FORCE metaphor, where a moving entity, a physical force, has the ability to move the listener with it. "Movement" corresponds MUSIC AS MOTION metaphor, where once again a moving entity can induce motion in the listener. "Wandering" corresponds to the MUSIC IS A LANDSCAPE metaphor where the listener can move through a fixed space or observe another agent move through it. "Flow" corresponds to the MUSIC IS A LIQUID metaphor, where the listener can move through a fluid environment, observe another agent move through it, or observe the environment moving towards one's self or an imaginary agent/location. Finally, "Interior" corresponds to LISTENER AS A CONTAINER metaphor, where in this case the motion is directed in and out of the listener, in reference to either the music itself or one's own thoughts, feelings and sensations moving with the music. These metaphors are not conceptualized as mutually exclusive, but rather as a complex multi-dimensional spectrum, where the listener may experience different types of motion either simultaneously and/or unfolding with the temporal motion of the music. Furthermore, the listener may exert the capacity to chose from which perspective to experience music, although this choice is also prompted by the musical stimulus (Spitzer, 2021).

In light of these conceptual metaphors, the two clusters of correlations can now be re-evaluated as representing two different general types of musical motion. The red cluster (ratings PC2), associated with dance music, is characterized by metaphors that refer to musical motions that mainly affords overt movements, what Cox (2016) would refer to as MMA. Whereas the blue cluster (ratings PC1), associated with ambient music, is characterized by metaphors that refer to musical motions that mainly afford covert motor sensations, what Cox (2016) would refer to as MMI. The case of the "Wandering" metaphor is interesting one to point out, as it is the only one which is correlated with different groups between electronic and classical music: in this study it was associated with ambient music, whereas in Schaerlaeken et al. (2022) it was associated with Group 1. This may indicate that electronic dance music is less prone to be conceptualized as a fixed landscape in comparison to classical music that also affords overt movements for the listener, where the listener may more easily represent one's self as moving through this landscape. Instead, the fact that electronic ambient music is associated with this metaphor may indicate that listeners represent one's self as observing an agent moving through this landscape, but not necessarily representing one's self moving through it. An alternative explanation to the exception of the metaphor of Wandering, may be linked to methodological differences between the two studies. In Schaerlaeken et al. (2022) each participant only evaluated the music according to either metaphors or emotions, whereas in this study both were evaluated in parallel. The rating of emotion labels may influence the rating metaphors, and vice-versa.

8.2.2 Entrainment

A common feature with the concept of mimetic engagement is the ability of music to activate sensorimotor neural circuits (Leman, 2018; Witek, 2019; 2023). Although this study includes no neuroscientific evidence to support this claim, it does contribute to the argument with through the bias of subjective feelings of entrainement, which are believed to emerge from a complex interplay of the four dimensions (perceptual, physiological, motor and social) of rhythmic entrainment (Labbé & Grandjean, 2014; Trost et al., 2017). Higher levels of subjective entrainment are correlated with “Arousal” and movement-related conceptual metaphors. According to Labbé & Grandjean (2014), feelings of subjective entrainment can be subdivided in Visceral entrainment (VE), corresponding to the feeling of changes in internal bodily sensations, and Motor entrainment (ME), corresponding to the tendency to move to the rhythm of the music. Thus, it could be expected that VE may be more associated with music predominantly affording MMI, as they both represent internal bodily sensations, whereas ME might be associated with music predominantly affording MMA, since they both represent overt movements. Following this logic, in this study it would be expected for VE correlated with ambient music and ME with dance music. However, this was not the case and the separation of VE and ME was not relevant, as they were both strongly correlated together. This lack of differentiation might be due to the experimental design of the survey, which was relatively long and involved answering to a large amount of questions several times[5]. This may imply a tendency to answer quickly and overlook subtle differences between items. In fact, several comments from participants indicate that both the length of the survey and the apparent similarity of certain items (especially in the MEQ) influenced their ratings. Schaerlaeken et al. (2022) also reported a lack of differentiation between subjective entrainment components. This may indicate that the different items of MEQ are not easily distinguished when presented to participants alongside with a large number of other questions. Furthermore, the comment section of the survey also hint the possibility of carry-over effects, where the previous excerpts influenced the following ratings. Since the variation of types of music was relatively minimal, an additional explanation might be that participants may feel expected to rate dance music higher than ambient music, with respects to certain variables, as an effect of contrast.

[5]Participants had to rate a total of 26 items after each music excerpts.

Rhythmic entrainment has been found to interact with the reward system (Dell'Anna et al., 2021; Trost et al., 2017) and as such it would be expected to be associated with pleasantness. However this was not the case in this study as neither VE not ME showed any particular correlation with "Valence", which was instead associated with ambient music. The opposite was true for Schaerlaeken et al. (2022) with classical music, where both "Valence" and entrainment were correlated with Group 1. A straight-forward explanation is that the selection of dance music excerpts was not particularly appreciated by the participants. In this regard, there are a few comments from participants which confirm this statement. Beyond inter-individual preferences, it might also be the case that classical music and ambient electronic music are more easily pleasant for any listener, whereas dance electronic music may only be pleasant to a niche of the population which is already accustomed to this type of music. In fact, even though "Familiarity" was fairly weakly correlated with both clusters, it was most correlated with "Valence". This may indicate that dance electronic music requires more habituation in order to be appreciated, in comparison to ambient electronic music. It has been established that through mere exposure one learns to understand and appreciate a certain type of music, through the improvement of one's predictive abilities which is also related to neural reward systems (Huron, 2006). Furthermore, predictive abilities also improve with respects to one's own capacity of mimetic engagement (Cox, 2016). Exposure contributes to mimetic engagement and appreciation of the musical experience, as "we tend to move to music that allows for easy motoric resonance with our preferred tempo." (Toiviainen & Carlson, 2022). Following our previous line of thought, although mimetic engagement also concerns ambient music, it comprises less potential to afford MMA, in comparison to dance music. Therefore, ambient music should require less exposure in order to be appreciated, since it's appreciation does not comprise the processes of understanding and learning how to overtly engage with the music through bodily movements. Furthermore, social entrainment is an important aspect of dance music (Leman et al., 2018; Witek, 2019; Trost et al., 2017). Deprived of the social context, it seems logical to assume that the mimetic engagement would also be lesser and thus also it's appreciation.

8.2.3 Emotions

The emotions of "Power" and "Joyful activation" were associated with dance music, which is also characterized by "Arousal", subjective entrainment and metaphors of "Movement" and "Force". These emotions, as the other elements to which they are associated with, implicitly express the concept of overt motion, and can be considered as predominantly affording MMA. The emotions of “Peacefulness”, “Tenderness”, “Sadness”, “Nostalgia” and “Transcendence” were associated with ambient music, and metaphors of “Wandering”, “Flow” and “Interior”. These items, in contrast, implicitly express covert bodily sensations, and can be considered as predominantly affording MMI.

The fact that the emotion of “Wonder” was equally moderately correlated with both groups is consistent with the results from Schaerlaeken et al. (2022). In both cases, it was most closely correlated with “Joyful Activation” and “Valence”. This result may support the argument that wonder has a particular role in the appreciation of art (Fingerhut & Prinz, 2018) and in the case of music, independently of it’s (sub-)genre or style. The emotion of “Tension” has a somewhat special status in this study, as it is weakly positively correlated with the first group and weakly negatively correlated with the second. The fact that it was only weakly correlated with the other items associated with dance music might be due to the differences between techno and dnb styles. In fact, “Tension” was the item with most variance between the two styles, with dnb extracts being rated higher (Appendix 3a, Figures 18-19). This may emerge from the musical structure of dnb being more irregular and dissonant than techno.

8.3 Interaction between acoustic features and subjective ratings

The two acoustic features PCs (MIR PC1 and MIR PC2) showed a strong negative correlation between each other and were found to be inversely correlated with the subjective ratings PCs: MIR PC1 being positively correlated with ratings PC2 and negatively with ratings PC1, while the opposite is true for MIR PC2. Ambient music excerpts (associated with higher scores for ratings PC1) are generally characterized by negative values for MIR PC1 and positive values for MIR PC2, while the opposite is true for dance music excerpts (associated with higher scores for ratings PC2). A particularity of the second component (MIR PC2) is that the results of the ANOVA for the linear mixed-effects model showed significance for the three way interaction between MIR PC2, ratings PCs and the music condition (sub-genre) and not for the two-way interaction without the condition. The opposite was the case with the first component. This is probably a consequence of the fact the distinction between ambient and dance was less clear cut for MIR PC2 than for MIR PC1.

In comparison to Schaerlaeken et al. (2022), the first component of acoustic features (MIR PC1) showed a stronger association to movement-related labels. This may reflect a stronger association of electronic dance music with overt body movements. In both studies, the second component of acoustic features (MIR PC2) was positively correlated with the first cluster of subjective rating variables and a negatively with the second (movement-related). Following the argumentation from the previous sections, this consistently indicates a potential tight association of the acoustic features of this component with covert bodily sensations.

In both ambient and dance music, a trend was observed for ratings PC2 values increasing together with MIR PC1 values. These results seem to indicate that, in addition of being a good predictor for the distinction between dance and ambient music, MIR PC1 is also predictive of the extent to which variables from the first cluster of subjective ratings are associated with the musical experience, independent of the music's sub-genre. An increase in tempo and timbre, understood in the sense of higher harmonics, have been suggested to result in higher activation in the listener (Gabrielsson, 2016). Furthermore, high-frequency elements and percussive rhythmic structures have been correlated to a higher amount of movement in motion capture studies on electronic dance music (Burger & Toiviainen, 2020). These results are consistent with those of this study, where increasing values of MIR PC1 are related to an increase in subjective ratings of entrainment, "Arousal", the metaphors of "Movement" and "Force", as well as the emotions of "Power", "Joyful activation" and "Tension". Since this component is by far mostly associated with spectral features of timbre, it may very well indicate that timbre has a relevant influence on the affordance of MMA. In addition, ratings PC1 values decrease for ambient music as MIR PC1 increases, which reflects a potentially simultaneous decrease in the affordance of MMI, as that of MMA becomes more predominant. The fact that this pattern is not observed in dance music might be due to the fact that ratings PC1 scores are already low, and thus so is the affordance of MMI.

With increasing values of MIR PC2, subjective ratings PC2 show a decreasing trend for ambient music, while ratings PC1 show an increasing trend for dance music. This case, unlike previously, shows a clear distinction of how the acoustic features of this component are perceived differently for the two sub-genres of electronic music. This second component is negatively associated with energy, dissonance and rhythmic features, while being positively associated with novelty features. Rhythmic stability here is represented by metric strength and pulse clarity, which "is related to the ease with which listeners can pick up an underlying pulse and thus implies the matching of an internal temporal percept with an external rhythm." (Trost et al., 2017, p. 100). As this decreases, together with dissonance and energy, while novelty features increase, ambient music becomes even less prone to afford MMA, and dance music even less prone to afford MMI. The complementary trends also seem to be present, even though not significant: ambient music becoming more prone to afford MMI and dance music to afford MMA.

As shown here, the two components of acoustic features interact with subjective ratings components, thus validating the last hypothesis (H3). During the musical experience, however, acoustic features are not dissociable from each other, and are instead processed simultaneously. Therefore, the two components must be considered in symbiosis on a two-dimensional spectrum. The two categories of ambient and dance music are constrained by different ranges of values for each one of the two acoustic features components. Within these ranges, the combination of the components influence the subjective ratings in two different resulting trends (Figure 7). On one end there are ambient music excerpts with strong negative values for MIR PC1 and strong positive values for MIR PC2, characterized by high scores for ratings PC1 and low scores for ratings PC2. As timbre decreases (lower frequency spectrum and energy), with more complex rhythms and irregular structure, ambient music increasingly affords MMI and decreasingly affords MMA. On the other end of the spectrum are dance music excerpts with strong positive values for MIR PC1 and weak negative values for MIR PC2, characterized by low scores for ratings PC1 and high scores for ratings PC2. In comparison to ambient music, all dance music excerpts have higher timbre (spectrum and energy), simpler rhythms and more regular structures. Within this range, as spectral features of timbre increase, while energy, rhythm complexity and structure regularity decrease, dance music increasingly affords MMA and decreasingly affords MMI. The peak of maximum MMA and minimum MMI for dance music is not reached if the rhythm is too simple and the structure is too regular, as would be expected in opposition to the trend for ambient music. This may be a reflection of the importance of violations of musical expectations as a major component of musical meaning (Meyer, 1956). In this sense, dance music would have more potential to induce action, if containing some degree of complexity and irregularity.

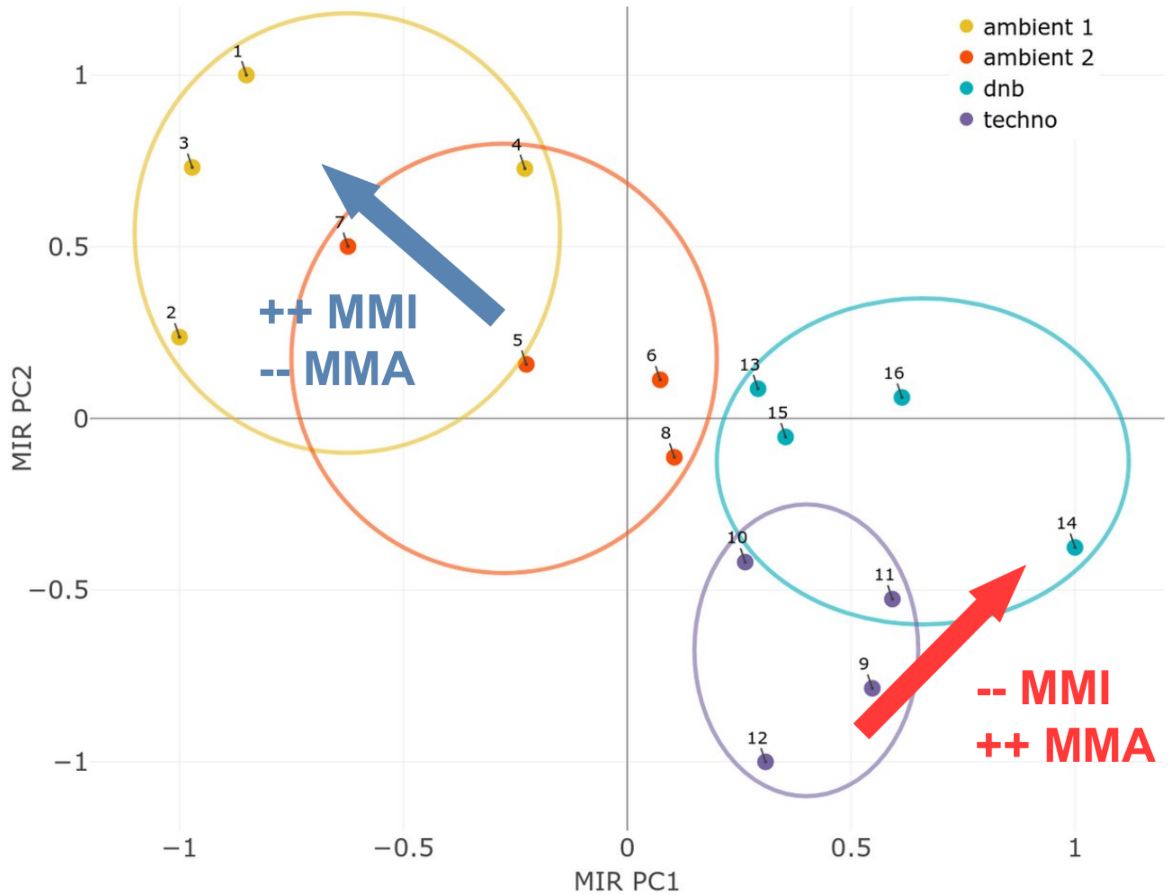

Figure 7: Figure 1 modified. Acoustic features PCA scores for each music excerpt, ranging from -1 to 1. The numbers indicate the excerpt id and the coulours indicate the style. The scores belonging to the same style are grouped together within an ellipse. The blue arrow indicates the trend of increasing Mimetic Motor Imagery (MMI) and decreasing Mimetic Motor Action (MMA) affordace, with respects to ambient music. The red arrow indicates the trend of increasing MMA and decreasing MMI affordace, with respects to dance music.

8.4 Strengths and weaknesses

This study aimed to provide evidence supporting an embodied music cognition framework and has successfully achieved this in several ways. First and foremost, the results are consistent with those of Schaerlaeken et al. (2022), linking feelings on entrainment with specific metaphors and emotions. This study explores another genre of music (electronic), and as such reinforces their ecological validity. Secondly, the conceptual framework provides a deepened conceptualization of the role of the body in the construal of musical meaning. The focus on an enactivist perspective of cognition appropriately accounts for differences in results for sub-genres of electronic music (ambient and dance). The focus on electronic music is also of scientific interest in itself, as it is a rapidly expanding contemporary genre of music, which is only marginally addressed in scientific research, in comparison with classical music. Furthermore, this study also contributes to bridging the gap between MIR technology and music psychology, promoting cross-disciplinary research and applications.

Alongside the scientific interests just mentioned, this work also suggests potential societal applications. Firstly, it contributes to the inclusion of perceptual and cognitive aspects in the classification of music. In fact, music providers could largely improve their services by adopting concepts as metaphors, emotions and entrainment in complement of a classification methodology purely based on acoustic features of sound. Music psychology has the potential to provide insight where MIR technology finds its limitations. Secondly, music composition and performance can surely benefit from a better understanding of how meaning is conveyed through music. Metaphorical and emotional concepts can aid musicians in the process of comprehending their own intentions of expression through music and how to do so through an embodied and enacted perspective. Similarly, such aspects can also have a positive influence in the field of music education. Metaphorical and emotional language is fundamental to transmit knowledge and insight from the teacher to the student.

It is worth mentioning that a few limitations emerged from the methodology of this study. The next three paragraphs are dedicated to different sets of limitations, concerning the sample population, the music excerpts and survey structure, and lastly the scales for subjective ratings.

The generalization of the results is relative to WEIRD population samples (Apicella, Norenzayan & Henrich, 2020). Given the culturally-dependent nature of metaphors (Dancygier & Sweetser, 2014; Nuñez & Cooperrider, 2006; Zbikowski, 2008), it would be of particular interest to address this aspect in further research. Furthermore, language difficulties were only evaluated at the end of the survey, thus there was no such information regarding the participants that did not reach the end of the survey. Among those that did (52.3%), 14.5% noted having some difficulties in rating the music excerpts due to language (Appendix 3b, Table 28). This issue was also explicitly addressed by a few comments by the participants, who mentioned having trouble understanding certain terms. In retrospect, an improvement would be to include language competence as an inclusion/exclusion criteria. An additional criteria should also concern any physical or psychological dysfunctions which may affect motor or auditory abilities. Concerning the subjective ratings, the results show a significant variation between participants (Appendix 3b, Figures 16-19). This probably indicates an important effect of inter-personal and contextual differences in the evaluation of music. A significant improvement would be to include several questionnaires to assess individual musical habits, personality and contextual information at the time of participation. Furthermore, it would also be favorable to offer a reward for participation, in order to incite completion and minimize and potential selection bias relative to attracting mostly individuals interested in electronic music. A distribution platform as Amazon Mechanical Turk (MTurk) could be a good candidate, which could also contribute to collecting data from a population beyond WEIRD.

Each style of music only comprised 4 music excerpts. This sample size is relatively small in order to establish robust interpretations from the statistical analysis of acoustic features. The main restrain for this aspect is the duration of the survey, which was already extensive. This could be overcome by including less excerpts per participant, over a larger sample size. Such a solution could also be applied to diversify the genres, sub-genres and styles of music. The choice of ambient and dance electronic music may have influenced polarization in ratings due to the distinct nature of these two sub-genres. In fact, a few comments from participants indicate a potential carry-over effect where ratings were influenced by those of previous excerpts. Other comments highlighted difficulties evaluating the several items on a numeric scale, it would thus be preferable to use a continuous slider referring to categorical limits (e.g. from “Strongly disagree” to “Strongly agree”). Furthermore, in contrast to Shaerlaeken et al. (2022), each participant evaluated both metaphors and emotions, which allowed to establish inter-individual correlations between items of these scales, but also increased the potential of exposure effect. In addition, this significantly increased the length of the survey. As suggested by Shaerlaeken et al. (2022), a repeated measurements design could be an alternative. Such a design would assess a fewer number of items, and so it could also be conceivable to use full music tracks, instead of excerpts, coupled with continuous ratings. This would allow to consider unfolding dynamics (Eliard et al., 2011), as well as potentially inducing a richer experience for the participant. Finally, the survey was also structured in a way that the ratings could only be made after having listened to the full length of the excerpt, in order to avoid participants skipping through rapidly without according attention to the music. However, this also implied that the items were assessed on the basis of memory. In the case of continuous ratings, this would issue would not be relevant.

Another improvement would be to investigate the distinction between *felt* and *perceived* emotions (Zentner et al., 2008). Even if participants were instructed to indicate to what extent they *felt* the particular emotion while listening to the excerpt, there was no means to effectively verify this claim. As such, it may be best to evaluate both in parallel. Concerning emotions, a further limitation is the GEMS itself (Juslin, 2016), which lacks means to evaluate negative emotions such as *boredom*, which is especially important when the music is imposed by the researcher, as well as other potentially relevant categories, such as *interest*, *surprise* and *awe*. This critique was also pointed out by participants in the comments section, who did not particularly appreciate the music nor felt any of the GEMS emotions while listening. Finally, the familiarity variable was too vague and it would be preferable to incorporate multiple items (e.g. “do you know the style of music”, “do you often hear it”, “do you chose to listen to it”) in order to properly evaluate its effect.

8.5 Future research

Aside from the improvements listed in the previous section, there are several interesting directions to explore in future research. The first concerns the acoustic feature analysis. It would be relevant to explore different methodologies of MIR analysis which are currently used to categorize large databases of music (Aucouturier & Pachet, 2003; Chen, 2014; Defferrand et al., 2016; Diakopoulos, 2009; Kirss, 2007; Lefaivre & Zhang, 2018; Nie, 2022). In doing so, it would further contribute to bridging the gap between MIR technology and music psychology (Acounturier & Bigand, 2013; Siedenburg et al., 2016). This should also be related with perceptual features (Aljanaki & Soleymani; 2018), as they provide a good mediator between the physical properties of sounds and the psychological phenomenology of listening to music. A second direction worth pursuing is in the laboratory. Firstly, it would be relevant to investigate the neural correlates of rhythmic entrainment and action-perception coupling. This can be achieved for instance by collecting EEG data of *mu waves* over the motor-related brain areas (Anderson, 2020; Fox et al., 2016; Hari, 2006; Pineda, 2005; Ross, 2022; Wu et al., 2016). This could further be complemented with physiological data (Bartlett, 1996; Gaston, 1951; Hodges, 2016; Matyja, 2016), in order to establish correlations between neural, physiological and subjective entrainment as well as emotional reactions. An additional path would be to analyse movements while dancing (Toiviainen & Carlson, 2022). Finally, these previous orientations could be combined with ethnographical fieldwork in a fundamentally inter-disciplinary approach. Firstly, intercultural studies could shed light on universal vs. cultural aspects of music cognition. A potential path could be to adopt bottom-up methodologies similar to those involved in the creation of the GEMS and GEMMES, in order to establish cross-cultural comparisons in musically-relevant metaphorical and emotional labels. Secondly, a collection of contextually-situated qualitative research could shed light on phenomena of social entrainment (Garcia, 2020; Witek, 2019).

9 Conclusion

This paper explored the role of the body in the perceptual and cognitive processes involved in the experience of listening to music. The conceptual framework presented in this paper provided a theoretical basis of embodied music cognition, to interpret the results of the empirical study. Music perception is considered from an enactivist perspective, where musical affordances represent the potential of action tendencies as a consequence of the interaction between the listener and the music. The mapping of the physical properties of sounds to sensorimotor representations provide the basis of the phenomena of rhythmic entrainment and of the emergence of embodied musical meaning. The empirical study addressed musical meaning in terms of musical metaphors, musical emotions and subjective feelings of entrainment. The patterns of correlations between these elements were found to be consistent across electronic and classical music, suggesting that the difference between stimulative versus sedative bodily influences of music is reflected by underlying embodied experiences. In the case of electronic music, this difference is represented by dance music, being associated with action tendencies for overt bodily motion, and ambient music with action tendencies for covert bodily motion. Furthermore, the physical properties of music associated with the perception of timbre, rhythm and structure influence action tendencies for overt/covert motion differently with respects to ambient or dance music. Linking the physical properties of sounds with the subjective experience of listening to music is an attempt to bridge the disciplinary gap between computational sciences and music psychology. Such interdisciplinary approaches are fundamental to further understand the human experience of listening to music from a holistic perspective.

10 Aknowledgements

This research project would not have been possible without the contribution of several people, to whom I am very grateful. First and foremost, I would like to thank my supervisors. Prof. Fabrice Clément, from the University of Neuchâtel, who has always been very supportive and encouraged me to pursue my interest in music cognition at the Neuroscience of Emotions and Affective Dynamics (NEAD) laboratory at the University of Geneva. It is there that Prof. Didier Grandjean inspired me to work on a novel research project of my own. His contributions were essential to the elaboration of the conceptual and experimental design. During my internship at NAED, I had the chance to meet and collaborate with a few people, and I am very thankful to them for the precious help that they provided me. Firstly, the contributions of Andrés Posoda were greatly appreciated in the initial phase of my project, when I needed guidance to learn how to use the Qualtrics platform to administer the online survey. Secondly, I was able to discuss my data with Dr. Damien Benis, who provided me with essential insight for the statistical analyses and interpretation of the results. I greatly appreciated his presence and his help. Finally, the main source of inspiration for this project originates from the research of Dr. Simon Schaerlaeken, in particular his publication in 2022, "Linking musical metaphors and emotions evoked by the sound of classical music". I would like to thank him for the source of inspiration and for his help with the manipulation of the MIR toolbox. My interest in this paper was greatly influenced by a course given by prof. Ana Piata, during my master program in the University of Neuchâtel, which revolved around the role of metaphors in meaning and cognition. I really appreciated her course, as well as her engaging and pleasant pedagogy. I would like to thank her for the personal support and for her suggestions of relevant literature for my project. I would also like to thank prof. Steve Moran, who's course on "Data Science" gave me the basis of data analyses, without which I would not have been able to pursue this project. In addition, I also appreciated his personal support in the initial phases of my analyses, when he provided me with some suggestions on how to approach my data. For all that concerns the administrative aspects of this research project, I would like to thank Rachel Sanchez, from the University of Neuchâtel. Finally, I cannot express enough gratitude to all the participants of this study, who took the time to complete my survey even without any reward. A particular thank you to my friends and family who tested the pilot version of the survey, gave me feedback, helped me with its distribution and provided me with precious personal support.

11 Appendix

11.1 Appendix 1: Method

11.1.1 Appendix 1a: Participants

Country	n
Switzerland (french)	123
Switzerland (german)	7
Italy	21
France	7
England	5
Netherlands	1
United States	1
Canada	1
Argentina	1
Australia	1
Japan	1
NA	1

Table 1: Number of participants (n) per country of residency.

11.1.2 Appendix 1b: Materials

List of music excerpts:

https://www.dropbox.com/scl/fo/k4ppfhyjaze9xeyq3imlx/h?rlkey=sbk6iml98c8u6yujejoydzvhx&dl=0

List of original music tracks:

- Ambient 1 tracks (A1-A4):
  - A1: https://freemusicarchive.org/music/4T_Thieves/Electro_Cool/04_4t_thieves_-_bleep/
  - A2: https://freemusicarchive.org/music/aAirial/Winterkauen/aAirial_-_Winterkauen_EP_-_05_Lonely_Landscape/
  - A3: https://freemusicarchive.org/music/aAirial/Winterkauen/aAirial_-_Winterkauen_EP_-_06_Rainbow/
  - A4: https://freemusicarchive.org/music/kirk-osamayo/season-two-yellow/an-early-morning-deep-breath/

- Ambient 2 tracks (A5-A8):
  - A5: https://freemusicarchive.org/music/cryptic-scenery/interstate/valediction-1/
  - A6: https://freemusicarchive.org/music/maarten-schellekens/chillout-lounge/up-from-here/
  - A7: https://freemusicarchive.org/music/Daniel_Birch/in-the-presence-of-trees-original-score/daniel-birch-in-the-presence-of-trees-original-score-04-protection/
  - A8: https://freemusicarchive.org/music/Daniel_Birch/in-the-presence-of-trees-original-score/daniel-birch-in-the-presence-of-trees-original-score-03-community/
- Techno tracks (A9-A12):
  - A9: https://freemusicarchive.org/music/Cronicool_aka_029/Backwards_EP/01Cronicool-Adira_029_Rework-Backwards_EP-SLC34/
  - A10: https://freemusicarchive.org/music/Cronicool_aka_029/Backwards_EP/04029-De_la_Rue-Backwards_EP-SLC34/
  - A11: https://freemusicarchive.org/music/ngel_Garca/Happy_Jambo_1368/Angel_Garcia-Happy_Jambo/
  - A12: https://freemusicarchive.org/music/viscid/flux/dundun/
- Drum and Bass (dnb) tracks (A13-A16):
  - A13: https://freemusicarchive.org/music/Egg_Nebula/We_Are_Humanity_And_We_Are_Lost_In_The_Cosmic_Infinity/15_Egg_Nebula_-_Vulve/
  - A14: https://freemusicarchive.org/music/Hobotek/Mutual_Kitchen/03_Parkan_Flow/
  - A15: https://freemusicarchive.org/music/Le_Perche_Oreille/Contes_Et_Histoires_Grsillantes/11_-_Le_Perche_Oreille_-_Fin/
  - A16: https://freemusicarchive.org/music/Skyzo_Maniak/Skyzo_Land/02_-_Skyzo_Maniak_-_PL1_IQ/

List of questions presented after each music excerpt:

- “Are you familiar with this style of music?”
- “To what extent would you use the following metaphors to describe this music excerpt ? Each metaphor is followed by a short list of characteristic terms and must be rated on a scale 0-100.”
  - Flow (harmonious, rocking/swaying, to float, to dream, to glide)
  - Movement (to move, body movement, rhythm, jumps/hops)
  - Force (impressive, empowerment, to amplify/magnify, immense/vast, grandeur/magnitude, intense)
  - Interior (to come undone/overwhelmed, deep/profound, inside one’s self, to internalize)
  - Wandering (large spaces, to leave, to go for a walk, to wander around, journey/trip)
- “To what extent did listening to this music excerpt evoke the following emotions ? Each emotion is followed by a short list of characteristic terms and must be rated on a scale 0-100.”
  - Wonder (happy, amazed, dazzled, allumed, moved)
  - Transcendence (inspired, feeling of spirituality, thrills)
  - Tenderness (in love, affectionate, sensual, softened-up)
  - Nostalgia (sentimental, dreamy, melancholic)
  - Peacefulness (calm, relaxed, serene, soothed, meditative)
  - Power (energetic, triumphant, fiery, strong, heroic)
  - Joyful Activation (simulated, joyful, animated, dancing, amused)
  - Tension (agitated, nervous, tense, impatient, imitated)
  - Sadness (sad, sorrowful)
- “To what extent did you feel … Each slider must be rated on a scale 0-100.”
  - physically stimulated
  - like dancing
  - entrained/driven
  - energized
  - like moving
  - animated
  - physically excited

- the rhythm in your body
- bodily agitated
- like beating time, tempo or rhythm
- your own bodily rhythms change
- your own body resonate with the music

- “How would you rate this excerpt on the basis of the following scales? Each slider must be rated on a scale 0-100.”
  - Valence (pleasantness)
  - Arousal (excitement, activation)

11.1.3 Appendix 1c: Procedure

List of campus buildings:

- Université de Genève (UniGE):
  - Site UniMail (Bd du Pont-d’Arve 40, 1205 Genève)
  - Site Sciences II et III (Quai Ernest-Ansermet 30, 1205 Genève)
- Haute École d’Art et de Design (HEAD):
  - Bâtiment A (Rte des Franchises 2, 1203 Genève)
  - Bâtiment D (Bd James-Fazy 15, 1201 Genève)
  - Bâtiment principal (Av. de Châtelaine 5, 1203 Genève)

Distribution Channel	n
Direct link from Stefano Politi	53
Direct link from friends of Stefano Politi	40
Facebook/Instagram	17
UniGE QR code	13
UniNE email	46
NA	3

Table 2: Number of participants (n) per distribution channel.

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16
49	52	49	48	51	49	52	48	49	53	51	49	53	48	48	54

Table 3: Number of participant rating per music excerpt (A1-A16).

11.2 Appendix 2: Results

11.2.1 Appendix 2a: Acoustic features

MIRtoolbox predictors	Length	Description	Label
BRIGHTNESS 1000	350 ms	Energy percentage above 1000 Hz	Timbre
BRIGHTNESS 1500	350 ms	Energy percentage above 1500 Hz	
BRIGHTNESS 3000	350 ms	Energy percentage above 3000 Hz	
CEPSTRUM CENTROID	350 ms	Measure of the geometric center of the cepstrum	
CEPSTRUM FLUX	50 ms	Measure of the distance between successive frames in the cepstrum	
CEPSTRUM MAX	350 ms	Maximal value of the cepstrum	
CEPSTRUM MEAN	350 ms	Average value of the cepstrum	
HCDF	750 ms	Harmonic Change Detection Function : detection of the flow of the centroid tonal	
RMS ENERGY	350 ms	Root mean square energy	
ROLLOFF	350 ms	Estimation of the energy distribution in the high frequencies	
ROUGHNESS	50 ms	Sensory dissonance index	
SPECTRUM CENTROID	350 ms	Measure of the geometric center of the quantity of energy in the different band frequencies	
SPECTRUM ENTROPY	350 ms	Characterization of the degree of disorganization in the spectrum	
SPECTRUM FLATNESS	350 ms	Indicates whether the distribution is smooth or spiky, in the spectrum	
SPECTRUM FLUX	50 ms	Measure the distance between successive frames in the spectrum	
SPECTRUM KURTOSIS	350 ms	Measure of the flattening of data, in the spectrum	
SPECTRUM SKEWNESS	350 ms	Measure of the symmetry of the distribution	
SPECTRUM SPREAD	350 ms	Standard deviation index	
ZEROCROSS	350 ms	Number of sign change in the signal (noise index)	
KEYCLARITY	1000 ms	Wide estimation of the positions of the tonal center and their respective clarity	Tone
MODE	1000 ms	Mode indicator (minor, major)	
EVENTDENSITY	1000 ms	Estimate of the average frequency of events, i.e., the number of note onsets per second	Rhythm / Metric
METRICAL CENTROID	1000 ms	Measure of the geometric center of the metric	
METRICAL STRENGTH	1000 ms	Measure of the accuracy of the dominant metric	
NOVELTY METRIC	1000 ms	Change indicator of the tempo	
PULSECLARITY	5000 ms	Evaluation of the rhythmic clarity	
TEMPO	1000 ms	Speed indicator	
TEMPO CHANGE	1000 ms	Time difference between two successive frames	
NOVELTY CEPSTRUM	2000 ms	Indicator of change in the cepstral texture	Dynamics/Structure
NOVELTY CHROMAGRAM_1	2000 ms	Indicator of change in the energy distribution in the different pitches classes	
NOVELTY CHROMAGRAM_2	5000 ms	Indicator of change in the energy distribution in the different pitches classes 2	
NOVELTY KEYSTRENGTH_1	2000 ms	Indicator of change in the tone strength	
NOVELTY KEYSTRENGTH_2	5000 ms	Indicator of change in the tone strength 2	
NOVELTY MFCC	3000 ms	Indicator of change in the cepstral form of the sound	
NOVELTY SPECTRUM	2000 ms	Indicator of change in the spectral texture	
NOVELTY WAVEFORM	2000 ms	Indicator of change in the waveform	

Figure 8: Set of 36 acoustic features from the MIR toolbox (Thibault De Beauregard, 2017), used in the present study and in Schaerlaeken et al. (2019; 2022). For more detailed descriptions see Latrillot (2021) or Thibault De Beauregard (2017, p. 215-218).

Two principal components (MIR PC1, MIR PC2) cumulatively explained 57.4% of variance (43.8% and 13.5% respectively) (Table 4, Figure 9). After varimax rotation[6], they explained 38.3% and 19.1% respectively (Table 5). Rotation statistics are available in Table 6. Variables are considered representative of a component if the absolute value of their saturation level is at least above (-)0.5 (Figure 10). Each music excerpt possess a unique score for each component (Figure 11). A series of Wilcoxon Rank Sum Tests enabled comparisons of these scores between sub-genres (ambient, dance) (MIR PC1: p-value = 0.000155; MIR PC2: p-value = 0.00466) (see Tables 7-8 for descriptive statistics) and styles (ambient 1, ambient 2, techno, dnb) (see Tables 9-10 for descriptive statistics and Tables 11-12 for Wilcoxon Rank Sum

[6] Maximizes the sum of the variance of the squared loadings

Tests p-values).

	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	15.800	43.800	43.8
Dim.2	4.870	13.500	57.4
Dim.3	3.200	8.890	66.3
Dim.4	2.750	7.640	73.9
Dim.5	2.190	6.080	80.0
Dim.6	1.680	4.670	84.6
Dim.7	1.320	3.670	88.3
Dim.8	1.220	3.400	91.7
Dim.9	0.891	2.470	94.2
Dim.10	0.746	2.070	96.2
Dim.11	0.429	1.190	97.4
Dim.12	0.351	0.976	98.4
Dim.13	0.258	0.718	99.1
Dim.14	0.197	0.548	99.7
Dim.15	0.117	0.325	100.0
Dim.16	0.000	0.000	100.0

Table 4: Eigenvalues and variance explained for each dimension of the PCA for acoustic features.

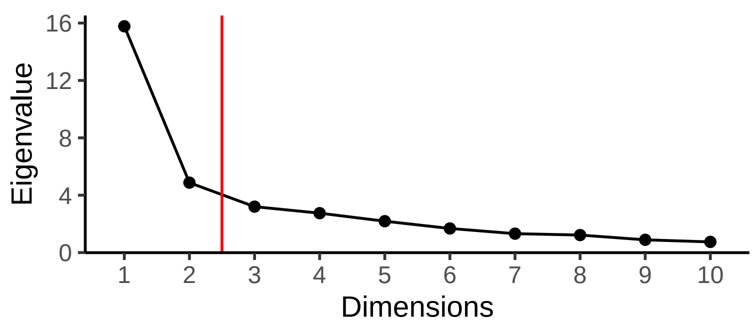

Figure 9: Screeplot of the eigenvalues for each dimension of the PCA for acoustic features. The dimensions to the left of the vertical (red) line are those considered as principal components.

	RC1	RC2
SS loadings	13.800	6.880
Proportion Var	0.383	0.191
Cumulative Var	0.383	0.574
Proportion Explained	0.667	0.333
Cumulative Proportion	0.667	1.000

Table 5: Variance explained after varimax rotation (RC1, RC2).

df	obj fun	n° obs	chi square	p.value	RMSR	fit
559	448	16	337	1	0.129	0.921

Table 6: Varimax rotation statistics for the PCA of acoustic features.

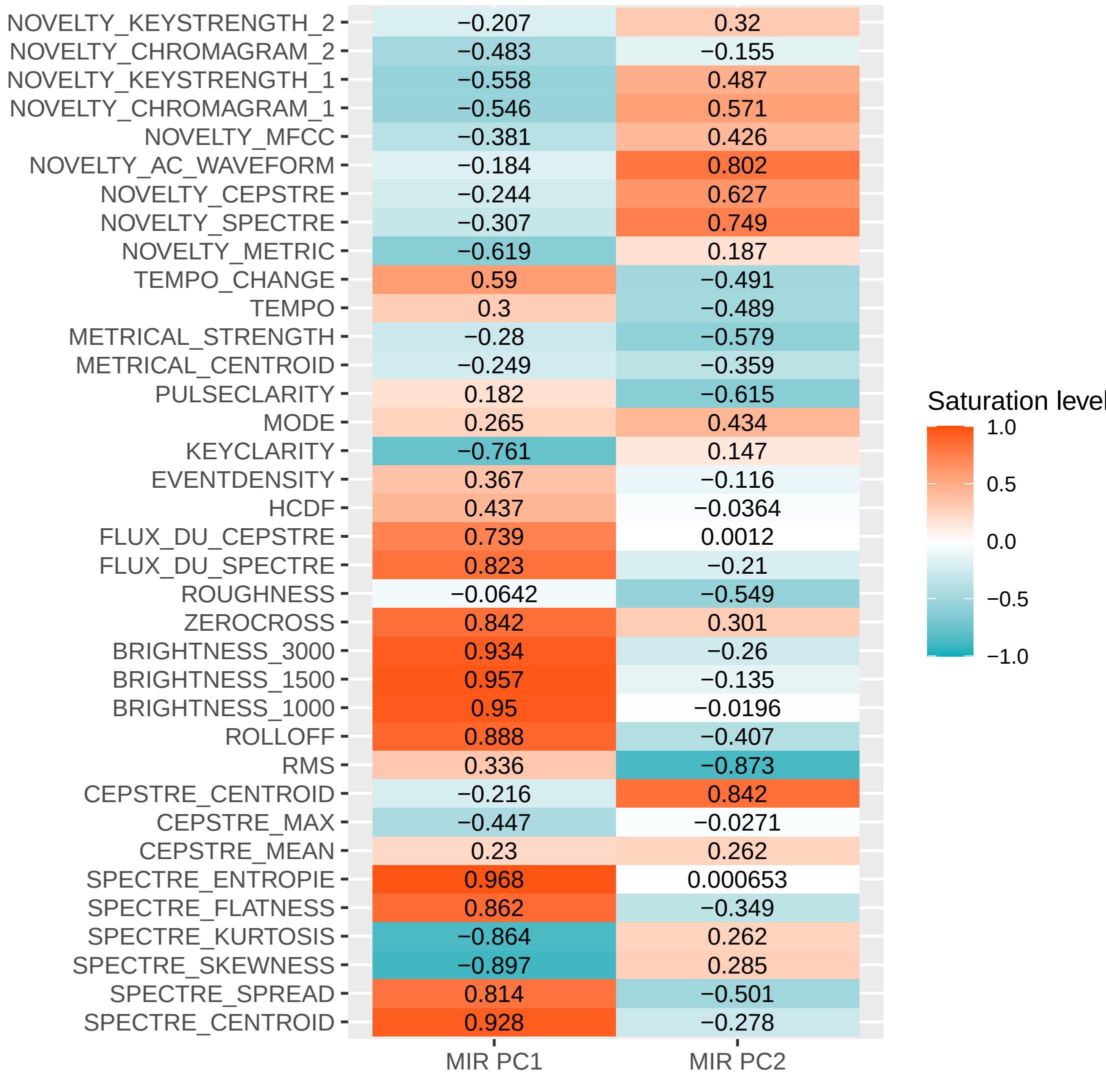

Figure 10: Acoustic features principal component loadings

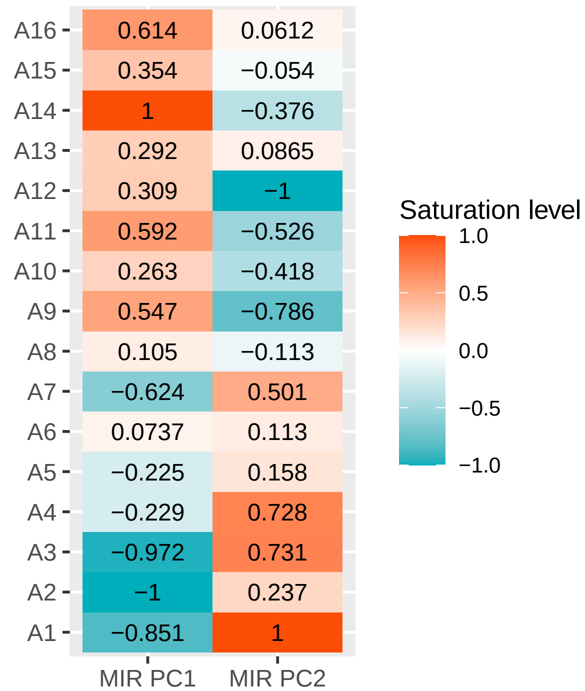

Figure 11: Acoustic features principal component scores per music excerpts

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient	8	-0.465	0.454	-0.426	-1.000	0.105	1.110	0.0329	-1.920	0.1610
dance	8	0.496	0.248	0.451	0.263	1.000	0.737	0.8170	-0.657	0.0877

Table 7: Descriptive statistics for MIR PC1 scores grouped by sub-genre.

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient	8	0.419	0.381	0.369	-0.113	1.0000	1.11	0.119	-1.65	0.135
dance	8	-0.377	0.395	-0.397	-1.000	0.0865	1.09	-0.199	-1.59	0.140

Table 8: Descriptive statistics for MIR PC2 scores grouped by sub-genre.

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient 1	4	-0.763	0.362	-0.9110	-1.000	-0.229	0.771	0.681000	-1.74	0.1810
ambient 2	4	-0.168	0.339	-0.0758	-0.624	0.105	0.729	-0.392000	-2.01	0.1690
techno	4	0.428	0.166	0.4280	0.263	0.592	0.329	-0.000655	-2.38	0.0829
dnb	4	0.565	0.322	0.4840	0.292	1.000	0.708	0.396000	-1.99	0.1610

Table 9: Descriptive statistics for MIR PC1 scores grouped by style.

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient 1	4	0.6740	0.318	0.72900	0.237	1.0000	0.763	-0.376	-1.83	0.159
ambient 2	4	0.1640	0.254	0.13500	-0.113	0.5010	0.614	0.254	-1.87	0.127
techno	4	-0.6830	0.262	-0.65600	-1.000	-0.4180	0.582	-0.151	-2.16	0.131
dnb	4	-0.0705	0.212	0.00361	-0.376	0.0865	0.462	-0.581	-1.82	0.106

Table 10: Descriptive statistics for MIR PC2 scores grouped by style.

	ambient 1	ambient 2	techno
ambient 2	0.0686	NA	NA
techno	0.0429	0.0429	NA
dnb	0.0429	0.0429	0.486

Table 11: P-values (fdr-adjusted) for Pairwise Wilcoxon Rank Sum Test comparison of MIR PC1 scores between styles.

	ambient 1	ambient 2	techno
ambient 2	0.0686	NA	NA
techno	0.0429	0.0429	NA
dnb	0.0429	0.2000	0.0429

Table 12: P-values (fdr-adjusted) for Pairwise Wilcoxon Rank Sum Test comparison of MIR PC2 scores between styles.

11.2.2 Appendix 2b: Subjective ratings

The correlations between the 19 subjective ratings variables were optimally grouped into two clusters (Figure 12), for k-means clustering visualization. These groups were very closely related to those that arose from the PCA. Two principal components (ratings PC1, ratings PC2) cumulatively explained 59.0% of variance (36.2% and 22.7% respectively) (Table 13, Figure 13). After varimax rotation, they explained 30.7% and 28.3% respectively (Table 14). Rotation statistics are available in Table 15. Variables are considered representative of a component if the absolute value of their saturation level is at least above 0.5 (Figure 14). Each music excerpt was characterized by a score for each PC (Figure 15). These scores enabled comparisons between the two sub-genres (ambient, dance) with Wilcoxon Rank Sum tests ($p < 0.01$ for both PCs) (see Tables 16-17 for descriptive statistics) and between styles (ambient 1, ambient 2, techno, dnb) (see Tables 18-19 for descriptive statistics and Tables 20-21 for Wilcoxon Rank Sum Tests p-values).

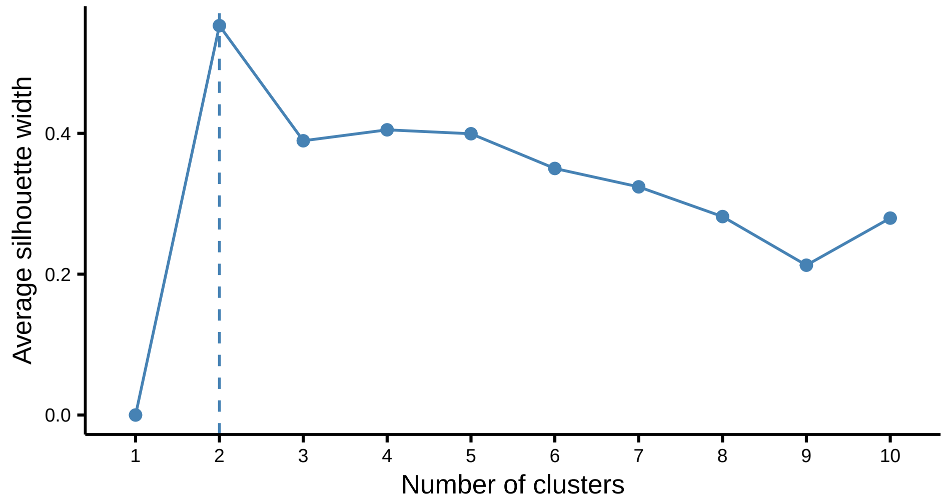

Figure 12: Optimal number of clusters as indicated by maximum value for average silhouette width.

	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	6.860	36.100	36.1
Dim.2	4.440	23.400	59.5
Dim.3	1.300	6.860	66.3
Dim.4	0.837	4.400	70.7
Dim.5	0.730	3.840	74.6
Dim.6	0.593	3.120	77.7
Dim.7	0.528	2.780	80.5
Dim.8	0.504	2.650	83.1
Dim.9	0.469	2.470	85.6
Dim.10	0.401	2.110	87.7
Dim.11	0.339	1.780	89.5
Dim.12	0.326	1.720	91.2
Dim.13	0.310	1.630	92.8
Dim.14	0.278	1.460	94.3
Dim.15	0.263	1.380	95.7
Dim.16	0.242	1.270	96.9
Dim.17	0.224	1.180	98.1
Dim.18	0.211	1.110	99.2
Dim.19	0.148	0.777	100.0

Table 14: Eigenvalues and variance explained for each dimension of the PCA for subjective ratings.

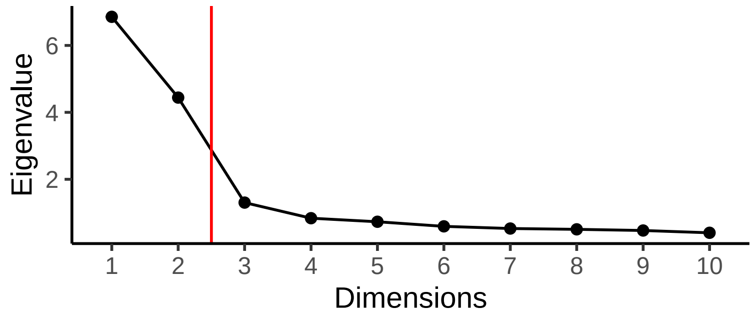

Figure 13: Screeplot of the eigenvalues for each dimension of the PCA for subjective ratings variables. The components to the left of the vertical (red) line are those considered as principal components.

	RC1	RC2
SS loadings	5.840	5.460
Proportion Var	0.307	0.287
Cumulative Var	0.307	0.595
Proportion Explained	0.517	0.483
Cumulative Proportion	0.517	1.000

Table 15: Variance explained after varimax rotations (RC1, RC2)

df	obj fun	n° obs	chi square	p.value	RMSR	fit
134	1.64	803	967	0	0.0593	0.977

Table 16: Varimax rotation statistics for the PCA of subjective ratings variables.

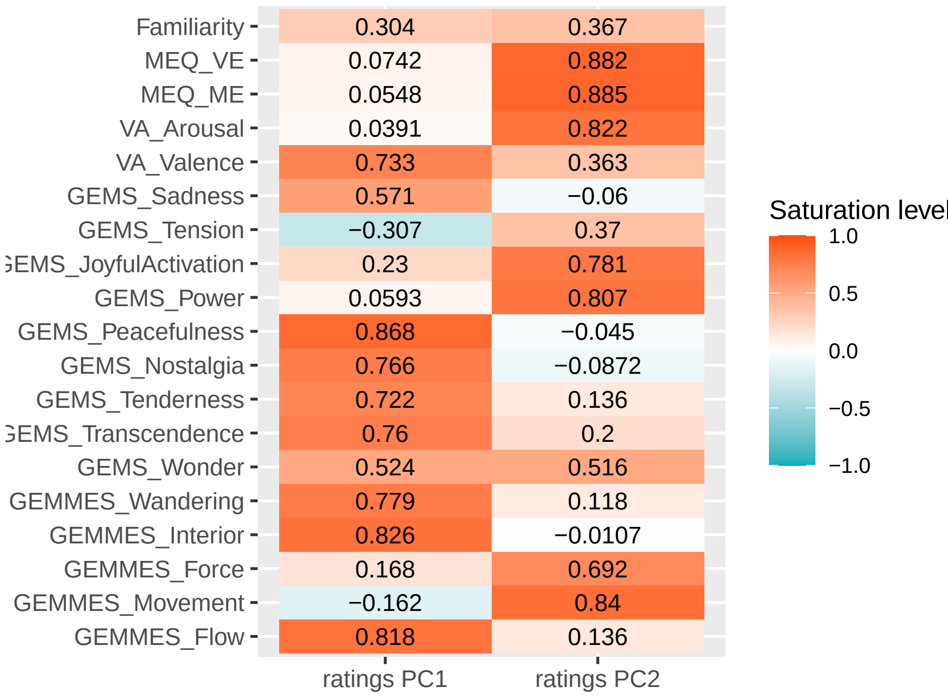

Figure 14: Subjective ratings principal component loadings

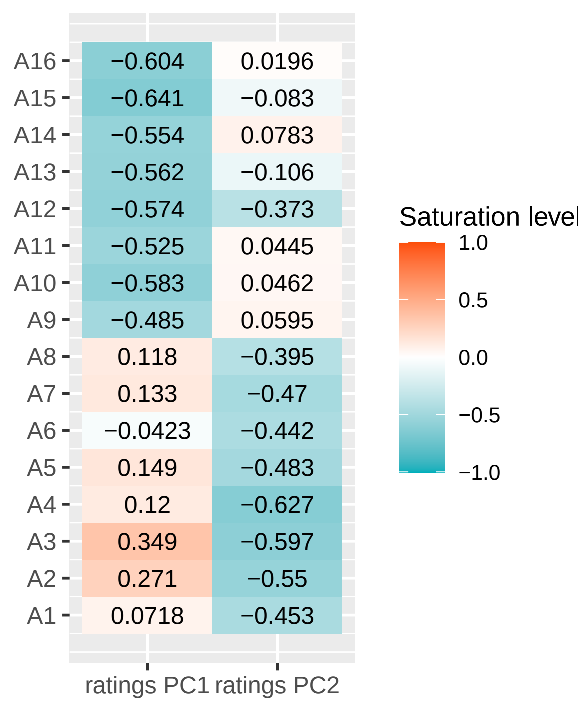

Figure 15: Subjective ratings principal component scores per music excerpts

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient	398	0.108	0.406	0.133	-0.719	1.000	1.72	-0.104	-0.763	0.0204
dance	405	-0.467	0.312	-0.564	-1.000	0.812	1.81	1.410	2.170	0.0155

Table 16: Descriptive statistics for ratings PC1 scores grouped by sub-genre.

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient	398	-0.4120	0.374	-0.4870	-1.00	0.745	1.74	0.9090	0.137	0.0187
dance	405	-0.0303	0.440	-0.0278	-0.84	1.000	1.84	0.0627	-1.010	0.0219

Table 17: Descriptive statistics for ratings PC2 scores grouped by sub-genre.

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient 1	198	0.164	0.394	0.1940	-0.719	1.000	1.72	-0.18600	-0.625	0.0280
ambient 2	200	0.054	0.412	0.0638	-0.712	0.995	1.71	-0.00329	-0.872	0.0292
techno	202	-0.437	0.308	-0.5460	-1.000	0.775	1.78	1.26000	1.360	0.0216
dnb	203	-0.497	0.314	-0.5850	-0.971	0.812	1.78	1.59000	3.090	0.0221

Table 18: Descriptive statistics for ratings PC1 scores grouped by style.

	n	mean	sd	median	min	max	range	skew	kurtosis	se
ambient 1	198	-0.47000	0.341	-0.54100	-1.000	0.740	1.74	0.9760	0.591	0.0242
ambient 2	200	-0.35400	0.396	-0.46100	-0.953	0.745	1.70	0.7920	-0.323	0.0280
techno	202	-0.05560	0.459	-0.04910	-0.837	0.869	1.71	0.0439	-1.210	0.0323
dnb	203	-0.00515	0.419	-0.00324	-0.840	1.000	1.84	0.1230	-0.815	0.0294

Table 19: Descriptive statistics for ratings PC2 scores grouped by style.

	ambient 1	ambient 2	techno
ambient 2	0.0107	NA	NA
techno	0.0000	0	NA
dnb	0.0000	0	0.0107

Table 20 P-values (fdr-adjusted) for Pairwise Wilcoxon Rank Sum Test comparison of ratings PCq scores between styles.

	ambient 1	ambient 2	techno
ambient 2	0.00418	NA	NA
techno	0.00000	0	NA
dnb	0.00000	0	0.275

Table 21: P-values (fdr-adjusted) for Pairwise Wilcoxon Rank Sum Test comparison of ratings PC2 scores between styles.

11.2.3 Appendix 2c: Interaction between acoustic features and subjective ratings

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
lmerMIR1.1	6	1260	1290	-622	1240	NA	NA	NA
lmerMIR1.2	10	1080	1140	-531	1060	183	4	0

Table 22: Models comparison, with sub-genre as a fixed factor for lmerMIR1.2.

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
lmerMIR2.1	6	1500	1530	-742	1480	NA	NA	NA
lmerMIR2.2	10	1090	1150	-537	1070	409	4	0

Table 23: Models comparison, with sub-genre as a fixed factor for lmerMIR2.2.

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
MIR_PC1	0.2400	0.2400	1	1520	2.420	1.20e-01
ratings_PC	0.1210	0.1210	1	1490	1.220	2.69e-01
condition	1.9200	1.9200	1	1520	19.400	1.13e-05
MIR_PC1:ratings_PC	1.8100	1.8100	1	1490	18.300	2.04e-05
MIR_PC1:condition	0.2770	0.2770	1	1580	2.800	9.46e-02
ratings_PC:condition	13.7000	13.7000	1	1490	138.000	0.00e+00
MIR_PC1:ratings_PC:condition	0.0142	0.0142	1	1490	0.143	7.05e-01

Table 24: ANOVA for interaction between subjective MIR PC1, ratings PCs and sub-genre.

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
MIR_PC2	1.04e-01	1.04e-01	1	1540	1.04e+00	3.08e-01
ratings_PC	1.52e-01	1.52e-01	1	1490	1.52e+00	2.17e-01
condition	2.26e+00	2.26e+00	1	1530	2.27e+01	2.10e-06
MIR_PC2:ratings_PC	2.65e-05	2.65e-05	1	1490	2.65e-04	9.87e-01
MIR_PC2:condition	5.56e-02	5.56e-02	1	1510	5.57e-01	4.56e-01
ratings_PC:condition	4.10e+01	4.10e+01	1	1490	4.11e+02	0.00e+00
MIR_PC2:ratings_PC:condition	1.69e+00	1.69e+00	1	1490	1.69e+01	4.11e-05

Table 25: ANOVA for interaction between subjective MIR PC2, ratings PCs and sub-genre.

ratings_PC	MIR_PC1	condition	MIR_PC1.trend	SE	df	t.ratio	p.value
1	0.0151	1	-0.1360	0.0384	1520	-3.530	0.00171
2	0.0151	1	0.1200	0.0384	1520	3.120	0.00374
1	0.0151	2	-0.0116	0.0709	1520	-0.164	0.87000
2	0.0151	2	0.2020	0.0709	1520	2.850	0.00582

Table 26: Comparison of MIR PC1 emtrends to horizontal slope, with respects to ratings PCs and sub-genre (condition 1 = ambient; condition 2 = dance). P-values are fdr adjusted.

ratings_PC	MIR_PC2	condition	MIR_PC2.trend	SE	df	t.ratio	p.value
1	0.0213	1	0.0492	0.0464	1520	1.06	0.2890
2	0.0213	1	-0.1300	0.0464	1520	-2.80	0.0208
1	0.0213	2	-0.0972	0.0428	1490	-2.27	0.0465
2	0.0213	2	0.0833	0.0428	1490	1.95	0.0690

Table 27: Comparison of MIR PC2 emtrends to horizontal slope, with respects to ratings PCs and sub-genre (condition 1 = ambient; condition 2 = dance). P-values are fdr adjusted.

11.3 Appendix 3: Discussion

11.3.1 Appendix 3a: Score distributions for music styles

The following figures illustrate the score distributions for all variables of subjective ratings, alongside the two acoustic features PCs. These distributions are portrayed separately for each style of electronic music. Color groups are based on Schaerlaeken et al. (2022) (Figure 8, p. 12), where red and orange correspond to Group 1 and the other colors to Group 2, with the exception of Valance, Familiarity and MIR PC1-2.

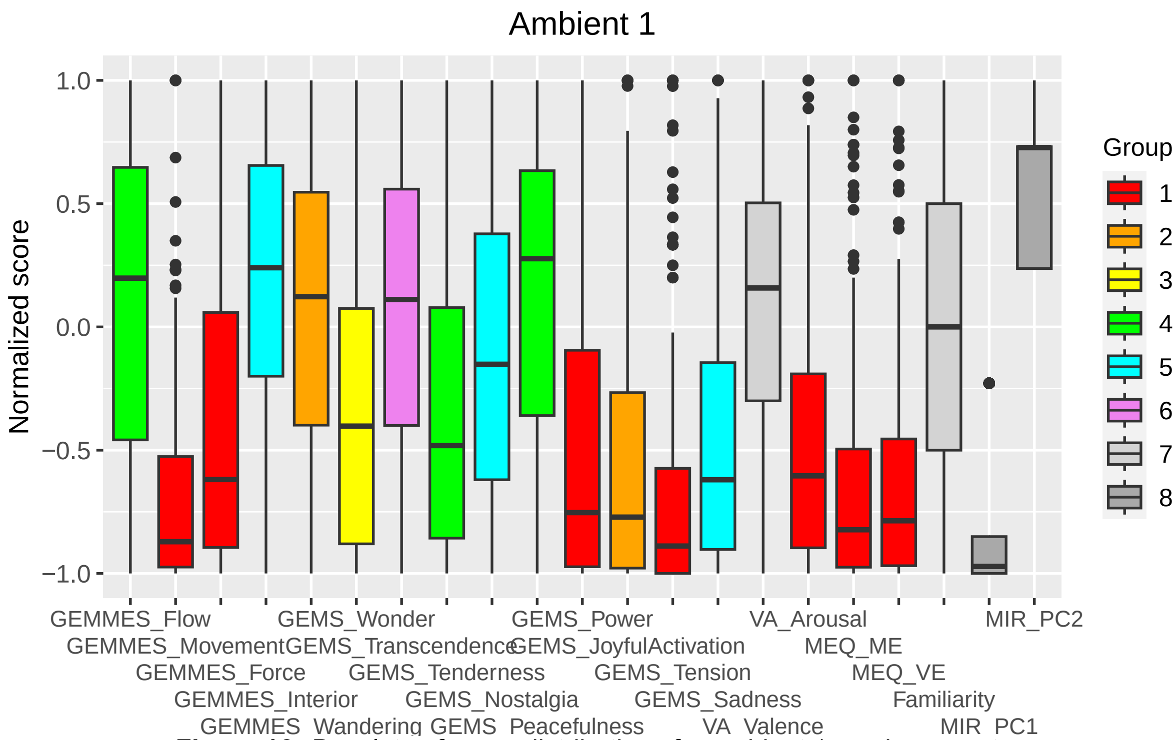

Figure 16: Boxplots of score distributions for ambient 1 music excerpts.

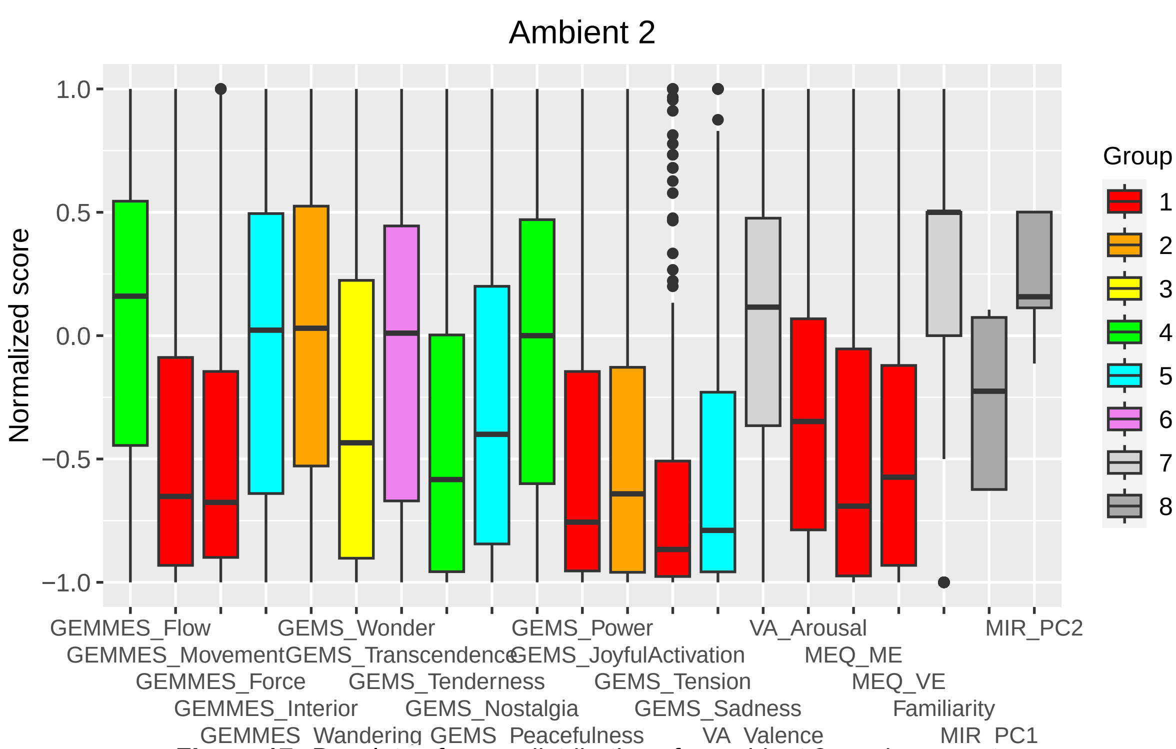

Figure 17: Boxplots of score distributions for ambient 2 music excerpts.

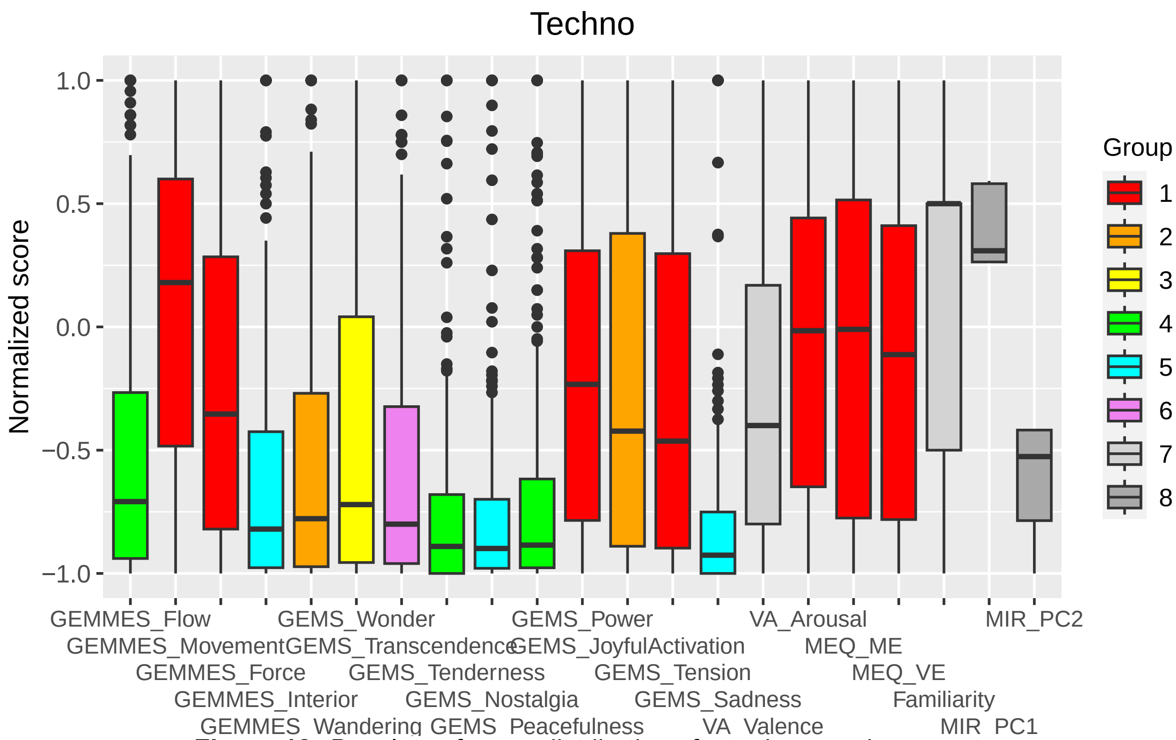

Figure 18: Boxplots of score distributions for techno music excerpts.

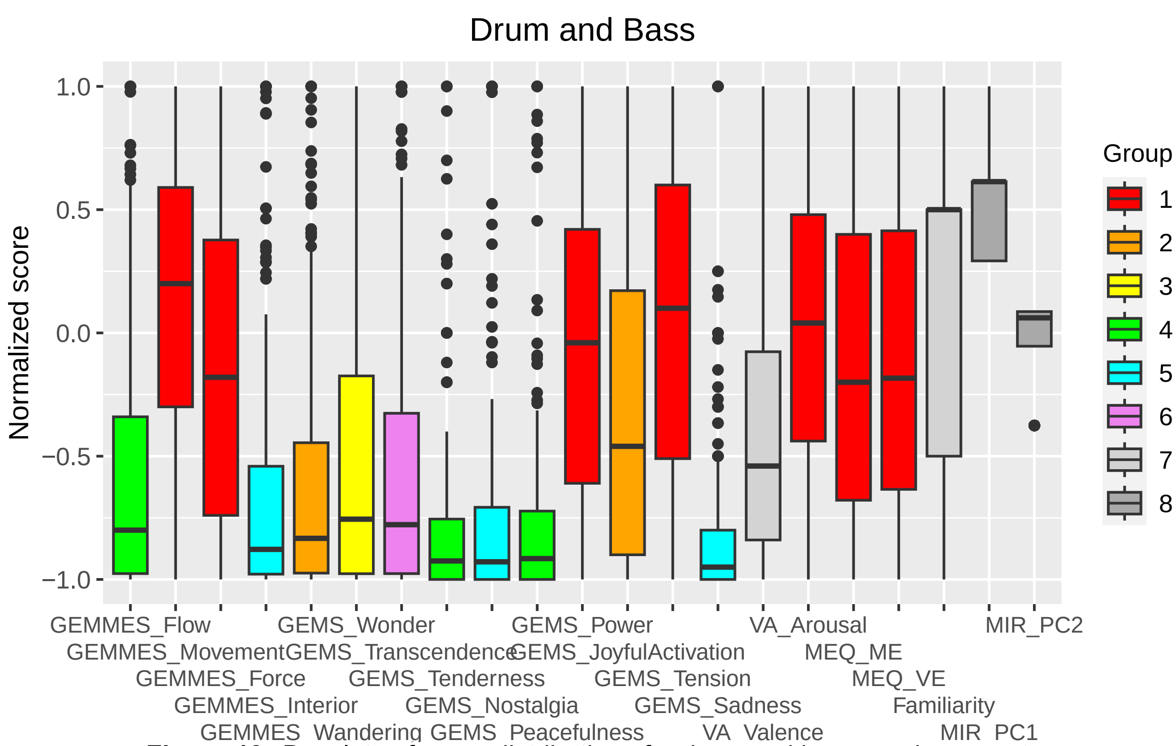

Figure 19: Boxplots of score distributions for drum and bass music excerpts.

11.3.2 Appendix 3b: Language difficulties

Language difficulties were self assessed by the participants at the end of the survey, by answering to the following question: “Did you have any difficulties answering the questions due to fact that they are in English?”

LanguageDifficulty	n
Strongly disagree	50
Somewhat disagree	17
Neither agree nor disagree	10
Somewhat agree	13
NA	82

Table 28: Number of participants (n) with language difficulties.

12 References

Aljanaki, A., & Soleymani, M. (2018). A data-driven approach to mid-level perceptual musical feature modeling. *arXiv preprint arXiv:1806.04903*.

Altenmüller, E., Kopiez, R., & Grewe, O. (2013). A contribution to the evolutionary basis of music: lessons from the chill. In Altenmüller, E., Schmidt, S., & Zimmermann, E. (Eds.), *The evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man*. Oxford University Press.

Anderson, J. (2020). The Action-Perception of Musical Rhythm: A Review of EEG Findings.

Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior, 41*(5), 319-329.

Aucouturier, J. J., & Bigand, E. (2013). Seven problems that keep MIR from attracting the interest of cognition and neuroscience. *Journal of Intelligent Information Systems, 41*(3), 483-497.

Aucouturier, J. J., & Pachet, F. (2003). Representing musical genre: A state of the art. *Journal of new music research, 32*(1), 83-93.

Bartlett, D. L. (1996). Physiological responses to music and sound stimuli. In Hodges, H. (Ed.), *Handbook of music psychology* (pp. 343-385). IMR Press.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition, 75*(1), 1-28.

Burger, B., & Toiviainen, P. (2020). Embodiment in electronic dance music: Effects of musical content and structure on body movement. *Musicae Scientiae, 24*(2), 186-205.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition, 106*(2), 579-593.

Chen, A. (2014). Automatic classification of electronic music and speech/music audio content (Doctoral dissertation, University of Illinois).

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral cortex, 18*(12), 2844-2854.

Clarke, E. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford University Press.

Cox, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Indiana University Press.

Cross, I. (2009). The evolutionary nature of musical meaning. *Musicae scientiae, 13*(2_suppl), 179-200.

Dancygier, B., & Sweetser, E. (2014). *Figurative language*. Cambridge University Press.

Davies, S. (1997). John Cage's 4'33": Is it music? *Australasian Journal of Philosophy, 75*(4), 448-462.

Defferrard, M., Benzi, K., Vandergheynst, P., & Bresson, X. (2016). FMA: A dataset for music analysis. *arXiv preprint arXiv:1612.01840*.

Dell'Anna, A., Leman, M., & Berti, A. (2021). Musical interaction reveals music as embodied language. *Frontiers in Neuroscience, 15*, 667838.

Diakopoulos, D., Vallis, O., Hochenbaum, J., Murphy, J. W., & Kapur, A. (2009). 21st Century Electronica: MIR Techniques for Classification and Performance. In *ISMIR* (pp. 465-470).

Eliard, K., Labbé, C., & Grandjean, D. (2011, May). Towards a dynamic approach to the study of emotions expressed by music. In *Proceedings 4th International Conference on Intelligent Technologies for Interactive Entertainment*, Genoa, Italy.

Fauconnier, G., & Turner, M. (2008). Rethinking metaphor. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 53-66). Cambridge University Press.

Fingerhut, J., & Prinz, J. J. (2018). Wonder, appreciation, and the value of art. *Progress in brain research, 237*, 107-128.

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., Ferrari, P. F., & Van IJzendoorn, M. H. (2016). Assessing human mirror activity with EEG mu rhythm: A meta-analysis. *Psychological bulletin, 142*(3), 291-313.

Gabrielsson, A. (2016). The relationship between musical structure and perceived expression. In Hallam, S., Cross, I., & Thaut, M. (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 215-232). Oxford University Press.

Garcia, L. M. (2020, January). Feeling the vibe: sound, vibration, and affective attunement in electronic dance music scenes. In *Ethnomusicology Forum* (Vol. 29, No. 1, pp. 21-39). Routledge.

Gaston, E. T. (1951). Dynamic music factors in mood change. *Music Educators Journal, 37*(4), 42-44.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Houghton Mifflin.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum.

Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and cognition, 17*(2), 484-495.

Hari, R. (2006). Action–perception connection and the cortical mu rhythm. *Progress in brain research, 159*, 253-260.

Hodges, D. A. (2016). Bodily responses to music. In Hallam, S., Cross, I., & Thaut, M. (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 183-196). Oxford University Press.

Huron, D. (1996). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press.

Johnson, M. L. (1997). Embodied musical meaning. *Theory and Practice, 22*, 95-102.

Juslin, P. N. (2016). Emotional reactions to music. In Hallam, S., Cross, I., & Thaut, M. (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 197-213). Oxford University Press.

Kania, A. (2017). The Philosophy of Music. In *Stanford Encyclopedia of Philosophy*. Stanford.edu. Retrieved 11 Jan. 2024, from https://plato.stanford.edu/entries/music/.

Kirss, P. (2007). Audio based genre classification of electronic music (Master thesis, University of Jyväskylä).

Koelsch, S. (2011). Towards a neural basis of processing musical semantics. *Physics of life reviews, 8*(2), 89-105.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in cognitive sciences, 23*(1), 63-77.

Krueger, J. W. (2011). Doing things with music. *Phenomenology and the cognitive sciences, 10*, 1-22.

Labbé, C., & Grandjean, D. (2014). Musical emotions predicted by feelings of entrainment. *Music Perception: An Interdisciplinary Journal, 32*(2), 170-185.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh* (Vol. 4). Basic Books.

Lartillot, O. (2021). MIRtoolbox 1.8.1 User's Manual.

Lartillot, O., & Toiviainen, P. (2007, September 10–15). A Matlab toolbox for musical feature extraction from audio. In *Proceedings of the 10th International Conference on Digital Audio Effects*, Bordeaux, France.

Lauria, F. (2023). Affective Responses to Music: An Affective Science Perspective. *Philosophies, 8*(2), 16.

Lefaivre, A., & Zhang, J. (2018). Characterizing and classifying music subgenres. In *Conference of Open Innovations Association, FRUCT* (No. 23, pp. 505-509).

Leman, M., Maes, P. J., Nijs, L., & Van Dyck, E. (2018). What is embodied music cognition?. In Bader, R. (Ed.), *Springer handbook of systematic musicology* (pp. 747-760). Springer.

Levinson, J. (1998). Music, aesthetics of. In *The Routledge Encyclopedia of Philosophy*. Taylor and Francis. Retrieved 11 Jan. 2024, from https://www.rep.routledge.com/articles/thematic/music-aesthetics-of/v-1.

Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. Penguin.

Maes, P. J. (2016). Sensorimotor grounding of musical embodiment and the role of prediction: A review. *Frontiers in psychology, 7*, 308.

Matyja, J. R. (2016). Embodied Music Cognition: Trouble Ahead, Trouble Behind. *Frontiers in Psychology, 7*, 1891.

Matyja, J. R., & Schiavio, A. (2013). Enactive music cognition: background and research themes. *Constructivist foundations, 8*(3), 351-357.

Meyer L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Nie, K. (2022, December). Inaccurate Prediction or Genre Evolution? Rethinking

Genre Classification. In *Proceedings of the 23rd International Society for Music Information Retrieval Conference*, Bengaluru, India.

Núñez, R., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in cognitive sciences, 17*(5), 220-229.

Pannese, A., Rappaz, M. A., & Grandjean, D. (2016). Metaphor and music emotion: Ancient views and future directions. *Consciousness and Cognition, 44*, 61-71.

Pineda, J. A. (2005). The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing”. *Brain research reviews, 50*(1), 57-68.

Pinel, J. P. (2011). *Biopsychology* (8th ed.). Pearson education.

Ross, J. M., Comstock, D. C., Iversen, J. R., Makeig, S., & Balasubramaniam, R. (2022). Cortical mu rhythms during action and passive music listening. *Journal of neurophysiology, 127*(1), 213-224.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural networks, 18*(4), 317-352.

Schaerlaeken, S. (2019). Being moved by music: literally and figuratively (Doctoral dissertation, University of Geneva).

Schaerlaeken, S., Glowinski, D., & Grandjean, D. (2022). Linking musical metaphors and emotions evoked by the sound of classical music. *Psychology of Music, 50*(1), 245-264.

Schaerlaeken, S., Glowinski, D., Rappaz, M. A., & Grandjean, D. (2019). “Hearing music as…”: Metaphors evoked by the sound of classical music. *Psychomusicology: Music, Mind, and Brain, 29*(2-3), 100-115.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them?. *Journal of new music research, 33*(3), 239-251.

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In Juslin, P. N. & Sloboda, J.A. (eds.), *Music and emotion: Theory and research* (pp. 361-392). Oxford University Press.

Schiavio, A. (2014). Action, enaction, inter (en) action. *Empirical Musicology Review, 9*(3-4), 254-262.

Shove, P., & Repp, B. H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), *The practice of performance: Studies in musical interpretation* (pp. 55–83). Cambridge University Press.

Siedenburg, K., Fujinaga, I., & McAdams, S. (2016). A comparison of approaches to timbre descriptors in music information retrieval and music psychology. *Journal of New Music Research, 45*(1), 27-41.

Spitzer, M. (2021). *The Musical Human: A History of Life on Earth*. Bloomsbury.

Thibault De Beauregard, K. J. (2017). Dynamiques temporelles des émotions exprimées par la musique (Doctoral dissertation, University of Geneva).

Toiviainen, P., & Carlson, E. (2022). Embodied meter revisited: Entrainment, musical content, and genre in music-induced movement. *Music Perception: An Interdisciplinary Journal, 39*(3), 249-267.

Trost, W., & Vuilleumier, P. (2013). 'Rhythmic entrainment' as a mechanism for emotion induction and contagion by music: a neurophysiological perspective. In Cochrane, T., Fantini, B. & Scherer, K. R. (Eds.), *The Emotional Power of Music* (pp. 213-225). Oxford University Press.

Trost, W. J., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia, 96*, 96-110.

Wharton, T., & Cornell, L. (2021). Before meaning: creature construction, seasponges, lizards and Humean projection. In Ifantidou, E., Wharton, T. & Saussure, L. (Eds.), *Beyond Meaning* (pp. 177-198). John Benjamins.

Witek, M. (2019). Feeling at one: socio-affective distribution, vibe, and dance-music consciousness. In Herbert, R., Clarke, D., & Clarke, E. (Eds.), *Music and Consciousness 2: Worlds, Practices, Modalities* (pp. 93-112). Oxford University Press.

Witek, M. (2023). Rhythmic entrainment and embodied cognition. In Margulis, E. H., Loui, P., & Loughridge, De. (Eds.) *The Science-Music Borderlands: Reckoning with the Past and Imagining the Future* (pp. 161-182). MIT Press.

Wu, C. C., Hamm, J. P., Lim, V. K., & Kirk, I. J. (2016). Mu rhythm suppression demonstrates action representation in pianists during passive listening of piano melodies. *Experimental brain research, 234*, 2133-2139.

Zentner, M., & Eerola, T. (2010). Self-report measures and models. In Juslin, P. N. & Sloboda, J.A. (eds.), *Music and emotion: Theory, research, applications* (pp. 187–221.). Oxford University Press.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion, 8*(4), 494-521.

Zbikowski, L. M. (2008). Metaphor and music. In Gibbs Jr, R. W. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 502-524). Cambridge University Press.

Zbikowski, L. M. (2010). Music, emotion, analysis. *Music Analysis, 29*(1‐3), 37-60.

Zimbardo, G. P., & Gerrig, J. R. (2002). Perception. In D. J. Levitin (Ed.), *Foundations of cognitive psychology: core readings* (pp. 129-184). MIT Press.