Tell Me What You See

An Investigation on Visual Mental Imagery Evoked by Instrumental Music

Antonio Stella

Sound and Music Computing, 2019-06

Master's Thesis

Architecture and Media Technology
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Tell Me What You See,
an investigation on
Visual Mental Imagery
evoked by instrumental music

Theme:
Scientific Theme

Project Period:
Spring Semester 2019

Project Group:

Participant(s):
Antonio Stella

Supervisor(s):
Sofia Dahl

Copies: 0

Page Numbers: 96

Date of Completion:
May 28, 2019

Abstract:

Visual Mental Imagery (VMI) consists of 'seeing' in the absence of a sensory stimulus [43]. This thesis aims to analyze if instrumental music can foster VMI and if it is possible to associate sonic elements of the music with visual features of the imagery.
The investigation was conducted through an online survey consisting of a listening test with one song (among four possible ones, two created by us and two commercial recordings) and some questions. Participants were asked if they experienced VMI, to describe it and to select descriptive elements from a list.
72.6% of 135 effective participants reported experiencing VMI ($\chi^2 = 27.6$; df = 1; $p<0.002$). Results show that instrumental music fosters VMI on different levels of vividness. The descriptions and the elements-selection task confirmed that mixing sonic elements from two different songs leads to a third song whose imagery features visual elements belonging to the imageries of both original songs.

Contents

Preface **v**

1 Introduction **1**

2 Pilot Study **6**
2.1 Stimuli Selection 6
2.2 Experiment Setup 6
2.3 Results 7

3 Factors Influencing Visual Mental Imagery **10**

4 Stimuli Creation **13**
4.1 Songs' Feature Analysis 13
4.2 New Stimuli Creation 16

5 Survey Design **23**
5.1 Introduction, Candidates Background and Context 23
5.2 Listening Test 24
5.3 Familiarity 24
5.4 Visual Mental Imagery 25
5.5 Emotional State 26
5.6 Empathy 26
5.7 Listening Behavior 27
5.8 Musical Sophistication 27
5.9 Conclusion 28

6 Results **29**
6.1 Participants 29
6.2 Visual Mental Imagery 30
6.3 Descriptive Elements Selection 36
6.4 Sonic Features 42

7 Discussion and Future Research **48**

8 Conclusion **53**

Bibliography **54**

A Pilot Survey **58**

B Survey Setup **60**

C Gear Used **75**

D Imagery Descriptions **76**
D.1 Deep Blue Day . . . 76
D.1.1 Descriptions . . . 76
D.1.2 Words Count . . . 79
D.1.3 Sonic Features . . . 79
D.2 Test Song . . . 82
D.2.1 Descriptions . . . 82
D.2.2 Words Count . . . 84
D.2.3 Sonic Features . . . 84
D.3 Tanca . . . 86
D.3.1 Descriptions . . . 86
D.3.2 Words Count . . . 89
D.3.3 Sonic Features . . . 90
D.4 Modified Song . . . 92
D.4.1 Descriptions . . . 92
D.4.2 Words Count . . . 94
D.4.3 Sonic Features . . . 94

Preface

I have always been interested in the evocative power of music and how it can foster images in the listener's mind. The first memory I have of experiencing mental imagery dates back to elementary school. In a late summer evening I was with my friend Daniele and his family in his father's car. We were driving back home through the countryside listening to Echoes, from the album Meddle by Pink Floyd. In the middle of the song, when it gets more psychedelic (around minute 11), I experienced the first Visual Mental Imagery that I can remember. I was looking through the car windows, into the dark flat land, clearly picturing lasers and alien spaceships in my mind. Back then I did not really thought about it, but the images were probably evoked by the high-pitched notes of the guitar together with the spooky synthesizer sounds.

Some years later I started playing guitar, and my huge passion for guitar pedals also began. Since then I tried to direct my musical projects towards the creation of music that could as well evoke images in listeners. Therefore, I started writing longer instrumental songs and incorporating visual tools–like live painting–to foster the creation of a mental imagery in the audience, giving them instruments to facilitate the picturing of a setting for the music being played.

This interest translated into a more formal research with my master degree at Aalborg University, first with a poster presentation at the 2018 KOSMOS Conference in Berlin, and now with this Master's Thesis.

Copenhagen, May 28, 2019

Antonio Stella
<astell17@student.aau.dk>

Chapter 1

Introduction

> "*For me music is such a visual thing, when I connect to a piece of music, it's almost like there's a location for it. [...] it's all been about the sound per se. [...] It's the sound that connects me with being in a place, in the energy.*"
> Ed O' Brien [1] - Radiohead.

Nowadays listening to music has become easier and easier. Since the advent of the internet and streaming platforms, music's role has evolved from being an entertaining event only enjoyable in music venues [9] or through an individual's small home collections of physical supports to being constantly present on every personal computer or handheld device [46]. Music has involuntarily traded the high level of attention and importance given by its relative scarceness with the possibility of being omnipresent in every moment of our life. Since less effort is made to retrieve music [11], consumers put less thoughts on the artistic side of it or the work behind music production, evolving in the trend of wanting it for free [27]. If on one hand this might be seen as a negative aspect of the internet era, on the other hand it led to a major opening (on the consumer side) towards the discovery of new and different genres [2] and in a way to the internalization of the idea of music being coherent with the different moments of an individual's routine. Therefore concepts like "relaxing music" or "party music" begun to be more present and unequivocal in the common jargon together with the idea of music having its own imagery. In fact, many people have experienced listening to music and spontaneously imagining visual scenes. Perhaps a dark forest, a sunny beach or a desert landscape. Such visual mental imagery instilled by music can be explicit and detailed with people and events. However, little is known of what arises these images in the first place and furthermore what musical features appear to foster these specific images.

[1] From the 5th of April That Pedal Show's episode `https://youtu.be/YK4Fmrlqz3I`

Definitions of Mental Imagery

Through the years different definitions of mental imagery were given by different researchers, each of them highlighting one specific aspect more than another, therefore there is not a universally accepted definition for it. In a nutshell, we experience mental imagery every time we have a specific perception when the event that usually cause this perception is not occurring. In 1969, Alan Richardson's definition of mental imagery was: "Mental imagery refers to all those quasi-sensory or quasi-perceptual experiences of which we are self-consciously aware, and which exist for us in the absence of those stimulus conditions that are known to produce their genuine sensory or perceptual counterparts, and which may be expected to have different consequences from their sensory or perceptual counterparts." [41]. This definition takes into consideration the basic idea of sensory-perceptual experience without the stimuli, but is circumscribing the topic only to conscious experiences. It is different, for instance, if we consider Nanay's article on multimodal mental imagery where he hypothesizes that mental imagery can affect different senses and it is often conscious but it can also be unconscious [31]. In 2001, Kosslyn et al. defined mental imagery from the point of view of episodic memory by writing that it "occurs when perceptual information is accessed from memory, giving rise to the experience of *seeing with the mind's eye*, *hearing with the mind's ear*" [23], this is of course in juxtaposition with the usual definitions of perception as experienced directly through the senses. However, to include a broader range of cases they also assume that mental imagery can be originated from the combination and modification of perceptual information already stored in memory. Among the different kinds of imageries [31] this thesis will focus on Visual Mental Imagery (VMI), defined by Taruffi and Küssner as "'seeing' in the absence of a sensory stimulus" [43]. In fact, according to Küssner and Eerola [43], mental imagery triggered by music is predominantly experienced as visual or a mix of other imageries with the visual.

Visual Mental Imagery fostered by music

Specifically related to music listening, there are a number of applications and definitions of Visual Mental Imagery (VMI). Composers have been using VMI as a tool for centuries. Especially in the Romantic period in the 19th century with 'Program music' [44]. Program music was commonly composed with the purpose of creating an extra-musical narrative through music. The narrative was supplied through program notes that listeners could follow to make their imagery evolve accordingly to a plot. Richard Strauss' Don Quixote [22] is an example of that.
In more recent times, Taruffi and Küssner affirmed that "[i]n relation to music, visual mental imagery refers to the mechanism whereby music stimulates internal images in the listener consisting of pictorial representations (e.g., natural land-

scape, colors), embodied image-schemata (e.g., picturing a melodic movement as an ascending or descending image), or complex visual narratives (e.g., similar to that of a movie)." [43]. From a different perspective, namely that of explaining emotional responses to music, Juslin and Västfjäll included VMI as one mechanism in how music can induce emotion. They define VMI as "a process whereby an emotion is induced in a listener because he or she conjures up visual images (e.g. of a beautiful landscape) while listening to the music. The emotions experienced are the result of a close interaction between the music and the images." [20]. In the same study the authors wonder whether VMI involves a pictorial representation of events or it reflects a propositional representation. In his book Image and Mind, Kosslyn [24] considers the images themselves as being quasi-pictorial representations, while the long-term structure of imagery is propositional.
The actual process that brings the listeners' mind to the formation of VMI is still obscure. In his 2007 article Bonde observes that "listeners seem to conceptualize the musical structure through a metaphorical nonverbal mapping between the music and so-called image-schemata grounded in bodily experience; for example, hearing melodic movement as "upward"." [6]. Metaphors is also the focus of Jungaberle et al. in their 2001 article. Right from the introduction they argue whether music can really be "dark", "heavy", "floating" or a sound be "mellow" or a rhythm "push", explaining that all these expression make use of the metaphorical concept of "musical space" [19]. Where the musical space is a situation that involves internal patterns (body, emotion, thought, images) and external stimulus patterns (physics-based like the acoustics, social like group dynamics and informative like communication and meaning) and it is never generated by means of the music as acoustical structure alone [19]. Among the rich diversity of points of view, different papers suggested that VMI is effectively stimulated by musical stimuli [35] [38]. Besides being object of academic research, Visual Mental Imagery is used in fields like the following.

Psychology

In psychology, Music and Imagery (MI) are used to induce a relaxed state in the patients. Their focus on the music listened is increased through the introduction of one or more images [14]. A slightly different approach is adopted in Guided Imagery and Music (GIM). It was perfected by Helen Bonny [7] and in therapy sessions patients "are invited to "share" images as they are experienced in real time during a pre-programmed sequence of music." [20]. The music induced experiences are shared "verbally using metaphors [. . .] [that] have clinical value and significance as information on the client's sense of self" [6]. The "guided" term refers to the fact that "therapist may respond to the imagery with affirming comments or questions to help the client experience it more fully." [14], in MI there are no verbal interventions. The state of relaxation consequent a music-induced im-

agery session usually also involves health benefits such as reduced cortisol levels [29]. GIM is considered a complete psychotherapeutic approach in and of itself [17] and has been used to treat different disorders like eating disorders and post traumatic stress disorders (PTSD) [14].

Film Music

In film music it is a bit more complicated to talk about VMI because of the nature of the format: music score is always combined with the visual side of the film. Therefore the tendency is that of associate film music to mood and emotion; as Cohen argues music is one of the strongest sources of emotions in films [10]. It is undeniable though that film music has its own character regardless the motion pictures it is coupled with. Pauli [36] proposes a theoretical framework with three basic categories that put in relation film music and its visual counterpart. The first one is called "paraphrasing" and the specific character of the music corresponds with the specific content of the picture; the effects are presumably additive. The second is "polarization" and the specific character of the music moves the ambiguous or indefinite content of the picture towards the character of the music. The third category is "counterpoint" and the specific character of the music contradicts the specific content of the picture; thus, the music conveys irony or comments on the content of the picture in another way. It appears clear that even if there is no explicit imagery connected, music's character shifts the perception of the picture [8]. Bullerjahn et al. in their research presented to different candidates the same short movie with different associated music, results showed that "each musical soundtrack creates its own particular type of film and plot", bringing to the conclusion that "film music polarizes the emotional atmosphere and influences the understanding of the plot" [8].

Taruffi and Kussner [43] end their review stating that in the field of VMI different questions remain untreated. What kind of images are fostered by music (e.g., colours, abstract shapes, etc.)? Does cultural background of listeners shape VMI? Can music features be associated with specific elements of the images? How are emotions involved in the mechanism of creation of the imagery?
Other questions could also be added. Does musical sophistication play a role in the assimilation of the music listened and in the image created? Is the context in which the music is listened shaping the VMI evoked? How does the listening behavior change the perception of music? Does the empathy of listeners play a role in perceiving a song's imagery?
The two main questions that this thesis will try to answer are:

- Can instrumental music foster Visual Mental Imagery?

- Is it possible to reconnect sonic elements of a piece of music to specific features of the VMI?

The research places its base on the poster 'Tell Me What You See' presented at the KOSMOS Conference in Berlin in 2018. In the next chapters we will report the pilot test and its results, talk about factors that could influence the VMI, explain how the stimuli for the current research were created, describe the experiment and how data were gathered, analyze the result and discuss them. In the conclusion a little wrap up will be done.

Chapter 2

Pilot Study

A pilot study on VMI was created and conducted between December 2017 and January 2018. The aim of this study was to investigate whether mental images are evoked in candidates' mind when they are presented with a 2-minute extract of instrumental music recordings.

2.1 Stimuli Selection

Four songs were selected with the purpose of generating VMI in participants. The choice was made trying to avoid episodic memory [24] to allow the candidates to focus on the music itself, on the sound, and the composition's layers. Hence the decision to use four instrumental pieces from commercial recordings. The selection was carried amidst songs considered to be likely unfamiliar to most listeners while still belonging to popular music taste. The music needed to be completely instrumental –to avoid any bias caused by lyrics [4]– and showcase a range of different musical features in terms of tempo, sound textures, instruments played and sound effects. Testers were not aware of the title of the track nor the artist. The songs selected were Hunting Bears[1] by Radiohead, Everest[2] by Modular Sonic Explorations, Tanca[3] by Iosonouncane and Deep Blue Day[4] by Brian Eno. One of the four would be randomly assigned to every participant starting the online test.

2.2 Experiment Setup

The experiment consisted of an online survey composed by a listening test followed by few questions. First, the testers were asked whether they were familiar with

[1] Hunting Bears, album Amnesiac, Radiohead, 2001
[2] Everest, album Modular Sonic Explorations, Matt Chamberlain, Viktor Krauss, Dan Phelps, 2011
[3] Tanca, album DIE, Iosonouncane, 2015
[4] Deep Blue Day, album Apollo, Brian Eno, 1983

the song. After, they had to assert if they experienced VMI and to describe the image evoked, if any. Every participant had to choose from 1 to 12 words between 35 descriptive elements (displayed in table 2.1) that could fit the song listened. Among the questions participants were requested to identify the features of the song that reminded them specific elements of the images they pictured.

Pilot Test: 35 descriptive elements	
Setting	On open water, On the beach, In a car, On a mountain, In the desert, On a train, In outer space, In the city, On a ship, In the forest, Inside a confined space
Weather conditions	Rainy, Cloudy, Sunny, Stormy, Warm, Snowy, Foggy, Frozen, Cold, Hot
Colour and consistency	Blue, Green, Red, Yellow, Shiny, Light, Dark, Dense, Empty, Opaque
Time of the day	Morning, Evening, Afternoon, Night

Table 2.1: Words that could be selected in the descriptive elements' selection task of the pilot test.

2.3 Results

The questionnaire was completed by 77 anonymous. 52 of them indicated that they saw some kind of VMI when listening to the piece of music. However a chi squared test showed significant results only for Tanca ($\chi^2 = 4.45$; df = 1; $p<0.05$). Nevertheless the results obtained from the descriptive elements selection were satisfying and relevant. The range of chosen words was often homogeneous and semantically coherent with the title of the song or the album. The same can be said for the descriptions given by candidates who experienced imagery. Participants were sometimes very meticulous with imageries and their details, they occasionally referred to specific scenes in films. The characteristics of the descriptions differed much between stimuli, this is evident for Deep Blue Day and Tanca.

Deep Blue Day

Deep Blue Day (DBD) was perceived as a warm and peaceful atmosphere. Some people described it as the sun rising (or shining) on a valley between mountains.

For others, the repetitiveness of the patterns was reminiscent of waves in the seaside. Peaceful and bright scenes were described, the prevalent colors were shades of green and light blue. The most selected terms in the descriptive elements task are shown in table 2.2.

Regarding the correspondence between sonic features and visual elements, *synthesizers* and *steel guitars* gave an ethereal feeling to the song. This was especially due to the *shimmer reverb* that gave a fluctuating motion to the sonic landscape, like waves on the shore or a spaceship in outer space. The overall mood was described as dreamy, calm, relaxed and peaceful but some others felt that this calmness gave them some sort of existential anxiety.

Tanca

Tanca (TNC) was quite unanimously described as the soundtrack of some sort of ritual performed by a tribe such as people dancing around a campfire, with weapons and armors, the night or morning before a battle or a war: overall definitely a dark and anxious atmosphere. The most selected terms in the descriptive elements task are shown in table 2.2.

In the sonic elements description it emerged that the idea of rituals and war was fostered by *timpani* and *percussion*. The *drone voices* reminded some candidates of monk chants. The *bass synthesizer* was said to give the whole scene a futuristic and artificial touch.

The mood generally perceived was anxiety, high-adrenaline, fear, sadness. A slim number of people also perceived calmness.

Considering these clear differences, we decided to use TNC and DBD as starting point for the stimuli adopted in the test developed of this current study. But first, in order to investigate more thoroughly on VMI we needed to identify what factors might influence the experience of imagery and the scenes evoked.

Pilot Test: most selected descriptive elements		
Song(candidates)	Deep Blue Day(26)	Tanca(19)
Setting	In outer space(13) On the beach(7) On a mountain(6)	On a mountain(5) In the forest(5)
Weather conditions	Warm(12) Sunny(8)	Hot(7) Stormy(5) Rainy(5) Cold(5)
Colour and consistency	Light(14) Blue(13) Yellow(8) Shiny(7) Opaque(6)	Dark(10) Dense(7)
Time of the day	Morning(10) Afternoon(9)	Night(8)

Table 2.2: The table shows the terms that wer mostly selected for DBD and TNC in the descriptive elements task of the pilot test. In brackets you find the number of participants that selected that specific term. Notice how –a part from 'On a mountain'– the words selected for both songs depict different imageries. One being positive and basically enjoyable, the other being negative and generally unpleasant.

Chapter 3

Factors Influencing Visual Mental Imagery

In scientific literature, it is hard to find studies that directly concern VMI. Therefore, it was difficult to state what could be the factors that actually influence imagery.

Yang et al. [49] in their study explained that *mood*, *context* and *music listened* are three factors to be considered interconnected when investigating on music's effects on listeners. They argued how "[t]he user mood is influenced by the user context and the personal traits; the user context is determined by factors such as the daily experience, social factors (e.g., listening alone or with friends), time and location; the music listening behavior is influenced by the user mood and user context, but is also conditioned on the individual's music taste and the audio, lyrics, and affective content of music.".

Another reason for which the *context* where candidates find themselves while listening to music is important is the fact that a noisy environment would distract more the candidates, lowering their attention level. Furthermore, according to Jungaberle [19] individuals tend to actively merge internal and external context with the perception of the music they are listening.

The fact that the listener's *mood* is important is also pointed out by Kamalzadeh et al. [21] in their 2012 research. Analyzing the results of their test on music listening on management behaviours, Kamalzadeh argued that "while it is important that the music during attention activities does not distract the listener, a large fraction of participants expressed a need for matching moods". It appears that listeners prefer music that matches their mood even while doing activities where their main focus is not going to be on the music itself. Thus, it could be assumed that proposing a song that does not match the candidate's mood could prevent VMI to be experienced. Moreover, according to Hunter et al. [16] if the listener is in a sad mood, the perception of music being sad even when the music does

not clearly sound happy or sad is increased. Nevertheless, Kamalzadeh's study [21] also showed evidence that among all participants "still nearly 40% preferred various moods".

Listening habits of individuals shape their approach to new music. Dunn et al. in their 2012 study [13] tried to find relations between music preferences, listening behavior and personality. Results "indicates that reported music preferences are generally related to listening behavior". Their research though was penalized by the categorization of music through musical genres, that inevitably are continuously evolving and with ambiguous boundaries.

Regardless of musical genres, how can our minds so often associate tight drum rhythms with rituals and tribes? How can they reconnect outer space and stars with reverberated pad synthesizers? Taruffi and Kussner in their framework [43] considered spontaneous images to be symbols or archetypes. The concept of archetype was used by Carl Gustav Jung [18] who affirmed that they "embody emotions and are collective and trans-personal, yet when activated, they are experienced by each individual in a unique personal way." [43]. Therefore even if individuals are tuned in with the same archetypes, it is likely that there will be differences from listener to listener and that diverse *cultural backgrounds* could accentuate these differences. Furthermore, North et al. and Tarrant et al. in two different studies say that English and American adolescents consider listening to music an important activity. This is because it "portrays an image of the outside world and satisfy their emotional needs" [32] and it helps with "self-actualising, fulfilling emotional needs, fulfilling social needs" [42]. Thus, considering music a relevant factor for adolescents' development it can be assumed that cultural background –from a musical point of view– also depends on the area of origin of individuals.

When considering listeners' reactions to music, not only individuals cultural background plays an important role but also their *empathy*. It was defined by Davis [12] as "the reaction of one individual to the observed experiences of another.". In a 2014 study, Balteș and Miu [4] explained that "individuals differ in their capacity to understand and respond to the emotions they perceive in the environment (i.e., empathy), as well as to associate stimuli from any sensory modality with vivid images in the mind (i.e., visual imagery)". Later they added that "empathy and visual imagery [have been identified] as two of the most important central 'routes' or mechanisms by which music may induce emotions in listeners". Similarly, in their 2012 paper on sad music, Vuoskoski and Eerola [45] found that there is a "significant connection between empathy and sadness induced by unfamiliar sad music", specifying that individuals with high trait empathy tend to experience stronger imageries due to their pronounced ability to enter the song's mood. Likewise, in his study from 2012 Wöllner [48] demonstrated that empathy –through the capacity of understanding expressive intentions of music performers– has an effect on the appreciation of performing arts such as music.

Even without considering empathy, for the purpose of this study it was crucial that participants were able to *distinguish the different musical instruments* that were present in a song. One might think that these kind of skills belongs more to candidates with a formal music training but Bigand and Poulin-Charronnat in their 2006 review [5] explained that "some musical capacities are acquired through exposure to music without the help of explicit training". However, it is true that training allows the development of skills in expert musicians, together with processing strategies and declarative knowledge that were never verified in non musicians. Nevertheless, even if "non-musicians do not learn a formal system with which they can describe and think about musical structures [...] they have a considerable amount of experience with music." [5]. Unfortunately the review was not comprehensive of information regarding the identification of musical elements that compose a song. Nonetheless it is reasonable to assume that this is as well a skill that develops over listening experiences more than formal musical education.

Another relevant factor to be considered is the *familiarity* that listeners might have with the music. Vuoskoski and Eerola [45], in their study about sad music explained that "music selected by participants is always more effective in inducing the intended emotional state than music selected by the experimenters.". According to the design of their experiment, participants were randomly assigned four different tasks: listening to self selected sad music, unfamiliar sad music, neutral music or recalling an autobiographical event and writing about it, the affective state induced by that task was measured indirectly. Therefore, we can assume that similar circumstances could also verify in the case of VMI, keeping in mind that episodic memory could constitute a bias.

To summarize, from the review of literature regarding music and its relation with different aspects of individuals life, we suppose that VMI can be influenced by the following factors:

- Context
- Emotional state
- Listening behavior
- Cultural background
- Empathy
- Musical sophistication
- Familiarity with the music

Therefore, the investigation focused on these aspects in connection to the musical stimuli. Chapter 4 will explain how stimuli were created.

Chapter 4

Stimuli Creation

The aim of this study is to establish whether individuals experience VMI when listening to instrumental music and determine if sonic features can be traced back to specific elements of the imagery. McKinney and Tims [29] suggested that a vivid imagery can be successfully stimulated through definite musical features. Repetitive and predictable melodies, harmony, rhythmic elements, slow tempo are musical elements that tend to foster VMI. All the four songs used in the pilot test were characterized by a relatively slow tempo, the most repetitive were most likely Deep Blue Day and Hunting Bears. Nonetheless, (as mentioned in Section 2.3) Deep Blue Day and Tanca proved to foster significantly different imageries, one being more positive and calm (DBD) and the other more anxious and negative (TNC). Therefore, they were chosen for further explorations. Acknowledging this, for the new test will be created two stimuli, one whose sonic features resemble DBD and one that reports both sonic features from DBD and TNC. This approach will allow to verify two main points: -songs with similar musical features can foster similar VMI. -mixing the musical elements of two different songs can lead to a VMI that features elements characteristics of both songs.
DBD and TNC will still be among the stimuli to confront their imagery with the ones evoked by the new stimuli. For the sake of creating the new songs we will proceed by performing a feature analysis over time of DBD and TNC. The analysis will then be used as guidelines for the new songs.

4.1 Songs' Feature Analysis

The basic parameters that we will consider for the analysis are key, tempo, instruments involved and harmony of the phrasing.

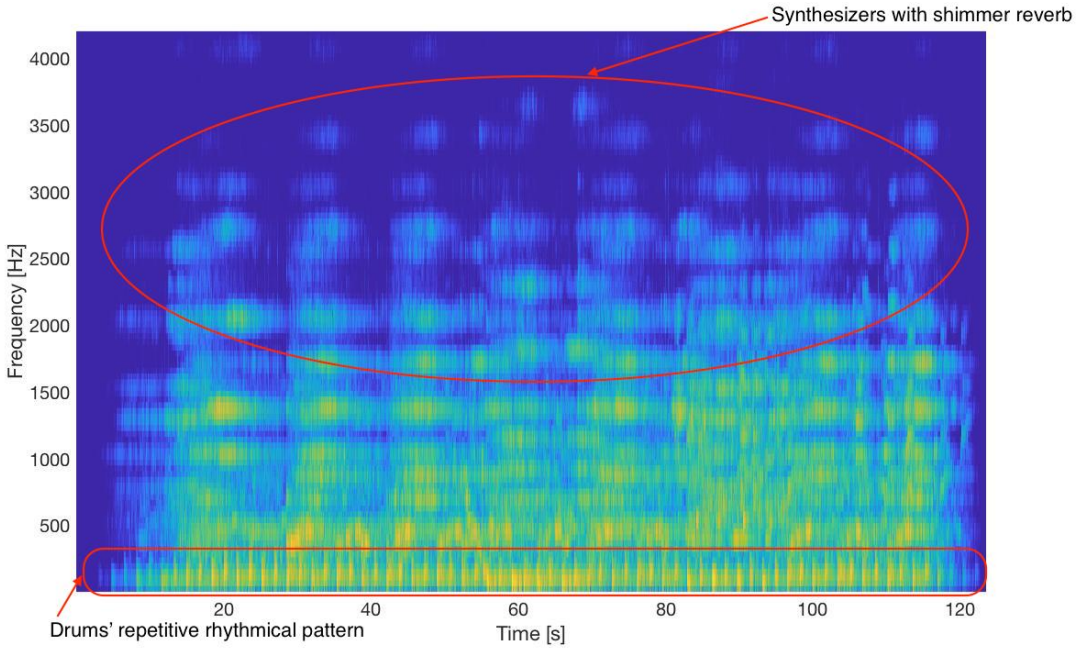

Figure 4.1: Deep Blue Day's spectrogram, Frequency[Hz] over Time[s].

Deep Blue Day

The song's tempo is 70 bpm and the key is E major. Chord progression starts with B in the initial fade in and then is E - A - E - B - E. Figure 4.1 shows the spectrogram of DBD, here it is easy to notice the repetitive rhythmical pattern of the drums and the "waves" of synthesizer's sound wet by the shimmer reverb that characterizes this song and gives a floating sensation.

0:00-0:28 The song starts with a fade in of all the 'background' instruments. Firstly we hear the shimmer reverb of the synthesizers, then the rhythm guitar's chords come up together with the bass and a muffled drum set sound. The rhythm is calm and peaceful, the bass line alternates between the dominant and the 5th of the cord being played. Most sonic 'space' is occupied by the pad of the synthesizers.

0:29-1:22 All the background instruments keep playing while the steel guitar starts its riff. The riff's notes are in ascending progression on the major scale and the last note is the same as the starting note.

1:22-end The lead steel guitar is added, its sound is brighter than the one playing the riff, this guitar is the only instrument that is not repeating over and over

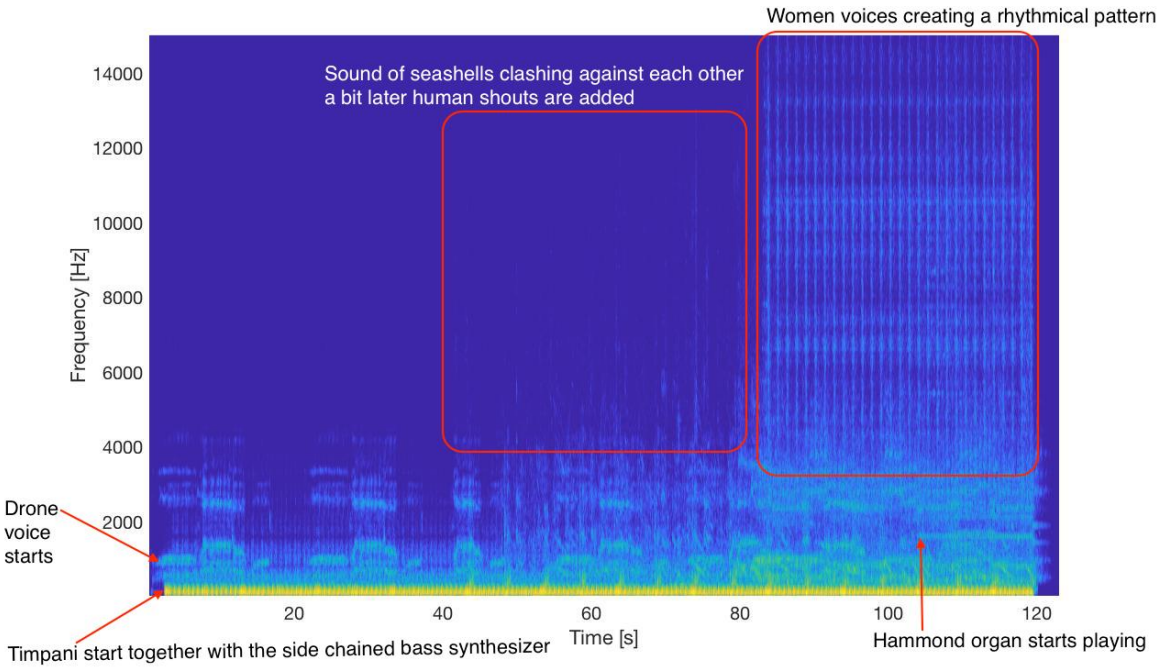

Figure 4.2: Tanca's spectrogram, Frequency[Hz] over Time[s].

the same notes.

1:58-end The background instruments slowly start fading out shortly followed by the two steel guitars.

Tanca

The song tempo is 95 bpm and the key is F minor, but the central chord of the extract is C7. Figure 4.2 shows the spectrogram of TNC where –thanks to the song's structure– it is easy to spot the different instruments that gradually are added to the music.

0:00-0:40 The song starts with two drones of low pitch human voices that resembles monk chants, one is panned to the right and the other one is to the left. Right after the drone, timpani with a regular rhythm start playing together with a bass synthesizer whose compression is side chained with the timpani's pattern, giving a 'breathing' effect to the bass. At 0:14 a wooden percussion starts dubbing the rhythm of the timpani adding some higher frequencies to the rhythmical section. At 0:23 another medium range frequency percussion is added on the right side of the audio panorama to give a more layered feeling to the rhythmical pattern.

0:41-0:48 Percussion that sound like seashells clashing against each other, or metal plates, start playing without a precise rhythmical pattern. At 00:43 clean guitar plays single notes that are slided towards the neck.

0:49-1:19 A layer of humans shouting and clamouring is added, there is not a recognizable pattern that the voices follow.

1:20-1:44 A higher pitched voice drone starts playing introducing (at 1:22) a new layer of women voices that create a rhythmical pattern and a melodic line at the same time. Together with the voices the rhythm of seashells intensifies, this keeps building up tension.

1:44- end An Hammond organ-like sound starts playing accentuating the tension.

4.2 New Stimuli Creation

The new stimuli were created using Logic Pro X. The first song created was the one with similar features to DBD, we will call this song the Test Song (TST). As mentioned above, another song was created taking TST as a base and adding sonic elements that are characteristic of TNC, we will call this song the Modified Song (MOD). For more information regarding the gear used to record these songs check Appendix C.

Test Song

The Test Song (TST) was created with in mind DBD, the idea was to make it similar, but not obviously the same. Each element of DBD was reproduced using instruments with a timber that reminded closely the original song and placed in an almost identical structure as in DBD. The bpm were kept at 70 as they are in DBD, but the key was shifted to C. The chord progression then became C - F - C - G - C.

Elements of TST

- The *rhythmical elements* (rhythm guitar, bass, drums) were kept similar to the original song, to avoid that the general perceived image could be too afar from DBD. The rhythm guitar was distorted and heavy reverberated with a spring-like sound. The bass was dry and equalized so that most of the high and hi-mid frequencies were cut out. The bass line alternates between the dominant and the 5th of the chord played by the guitar. The drums

were low-pass filtered in order to give them the "muffled" sound that all the rhythmical section has in DBD.

- The *synthesizers* have pad-like sounds, very dilated. They were recorded using Ableton Live native synthesizers. The attempt was to create a sound texture that resembles the shimmer reverb that is characteristic of DBD. One of the synthesizers has a piano/glockenspiel-like sound, the other one plays the same notes but has a more diffuse sound, with a longer attack time and richer in high frequencies. Both synthesizers play notes present on the C major scale without following a precise pattern.

- The *lead electric guitar* sound (that was replicating the steel guitar) had a long attack thanks to the envelope section in the POG 2, it also featured a good amount of high frequencies since the guitar signal was mixed with the digital signal created by the POG2 that was 1 and 2 octaves above. After the octave generator, the signal was processed with a reverse delay and a plate reverb. The reverse delay was set to a short delay time with long decay, this contributed in giving the sound texture a shimmery and watery character that recalls DBD. Thus we tried to mix (and in a way confuse) the lead guitar's sound with those of the diffuse synthesizers. The lead guitar follows a melodic line on the major C scale that for most of the song is ascending.

- An *over driven and reverberated guitar* was used to dub the lead electric guitar in order to accentuate the melodic line and stress the ascending phrasing by enlarging the perception of spaciousness through the reverb.

Time analysis of TST

Figure 4.3 shows the spectrogram of the Test Song.

0:00-0:28 To establish a clear difference from DBD, the song does not start with a fade in but rather with four sudden snare hits. The snare signal (only for these 4 hits) is patched to two very characteristic reverbs that are panned one to the left and the other to the right. The first one creates a cymbals-like sound while the second sounds like a pad synthesizer. Right after these four hits, all drums are low-pass filtered in order to give them the "muffled" perception of DBD. At the same moment all the instruments start, the ones that can easily be noticed are the smooth bass, the distorted and reverberated lead guitar and the synthesizers. In the background there is also the rhythm guitar.

0:28-0:40 There is a chord change from C to F, but the structure stays unchanged. Around 0:39 the lead guitars vary a bit the melodic line from the ascending pro-

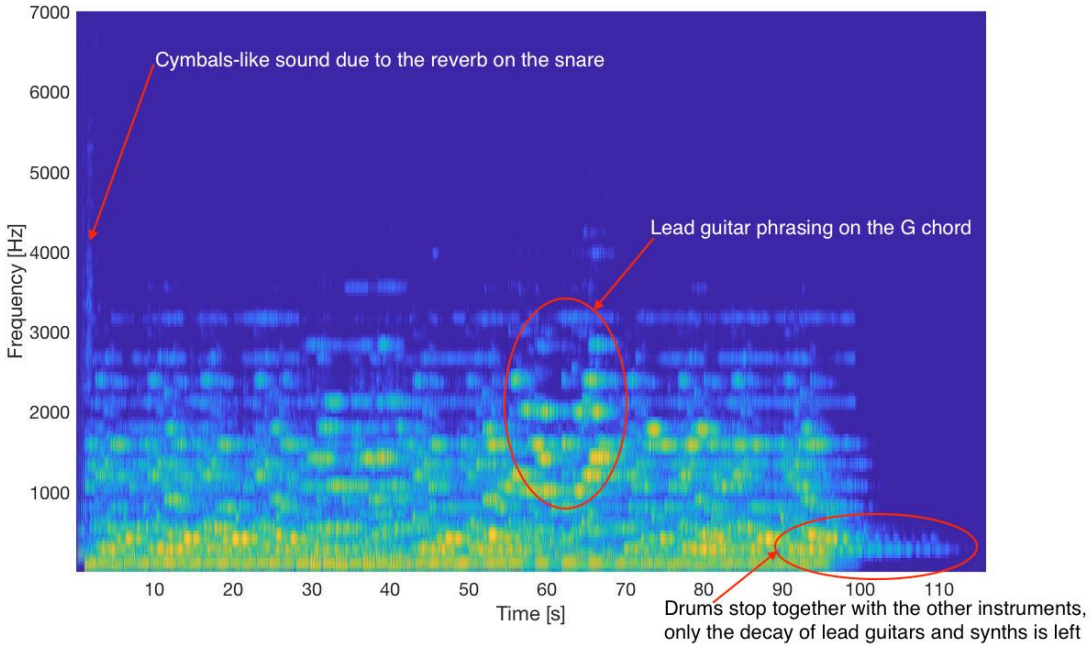

Figure 4.3: Test Song's spectrogram, Frequency[Hz] over Time[s].

gression, but always sticking to the C major scale.

0:41-1:08 The harmony goes back to C until 0:55 where there is a further change to G. On the G chord the lead guitars vary again the ascending melodic line with a short phrasing that differs a bit from DBD giving this stimuli a slightly different taste.

1:09-1:36 The harmony is back to C and all the instruments are back to the same notes of the beginning.

1:37-1:56 The drums stop and the sound of all the instruments sound are let to naturally decay. The synthesizers and the clean lead guitar are more evident than the others since their sound was designed to have long decay time.

While recording TST we supposed that candidates would generally not know DBD and even if they did, they would not directly associate TST to it since they did not have DBD fresh in mind. Nevertheless, in case they saw a connection the biggest risk for them would be to overlap DBD's imagery with the one created by TST, making the resulting description biased and not accurate. Yet, this unwanted result can be unveiled by asking if they found the song familiar (like we did in

the pilot test) and deciding to discard the answers from participants that saw this connection. Nonetheless, to forestall this eventuality we decided to test the finished TST with four selected listeners that we were sure knew Brian Eno's work and Apollo (the album that contains DBD). We asked every tester to listen to TST trying to guess any musical influence that they could find in it. After they answered, they were told that it was tailored on DBD to see if that flipped a switch. These were the comments:

- the first tester did not associate TST with DBD. When told that it was recorded to sound similar to DBD, he added that he knew Apollo but ambient records can be difficult to remember.
- the second listener (only after being told) said that it reminded him of Deep Blue Day and that the bass line was the element that he considered most similar between the two songs. He also added that he has been working with reverbs recently and that he had listened to DBD just few days before as it is the first example of shimmer reverb.
- the third answered that TST reminded him of U2's Joshua Tree (in particular Running To Stand Still), a bit of Radiohead's The Bends and Motion Picture Soundtrack and King Crimson's Islands. It is interesting to notice that Joshua Tree was produced by Brian Eno and Daniel Lanois, who wrote and recorded DBD.
- the fourth tester answered that it reminded him of Dream Pop/Bubble Gum Pop atmospheres and specifically of Kinky Love by Pale Saints and some Mogwai's songs, especially for the glockenspiel-like sound.

This led us to the conclusion that TST was not too obviously recognizable and that could be fit for our purpose.

Modified Song

As previously mentioned, the idea behind the Modified Song (MOD) is to mix sonic features that characterize DBD and TNC to verify if any change is detectable in the resulting VMI. Specifically, we want the imagery fostered by MOD to present elements from both the DBD imagery and TNC imagery. In order to create MOD, we took TST –created to foster a VMI that resembles the one of DBD– and we added musical features that characterize TNC, from approximately the middle of the song. According to the pilot test in Chapter 2, the sonic elements that characterize the most TNC are the timpani, the drone voices and the bass synthesizer. For MOD we reproduced these elements and added them to TST. The decision of adding the modifying elements around the middle of TST was taken assuming that candidates would form an imagery tuned with TST in the first half of MOD and then have enough time to tune with the MOD elements in the second half.

Elements of MOD

Since MOD consists of TST plus some added sonic elements, it is clear that it will feature all of the instruments that were listed for TST. Here are explained the added sonic elements of MOD.

- Two *drone voices*, one panned to the left and the other one to the right. They were created starting from samples of a male chamber choir available on the XS24, a sampler that is native in Logic Pro X. The notes sung are long and constant, very low in terms of frequency and from the timber point of view they could be described as growly.
- *Timpani* sound were created as well with the XS24, in order to make them sound darker and more emphasized, they were dubbed with drum kick sounds. The timpani follow a very simple rhythmical pattern throughout the song.
- *Bass synthesizer*, was created using Logic native synthesizers. Also in this case to achieve a full and 'fat' sounding bass, more virtual synthesizers sound were layered. They were equalized so that the resulting sound was deep and gritty. The sound of the bass synthesizer, like the drone voices, needed to be designed so that the sustain could be very long without loosing character.
- *Percussion* were also from different ethnic percussion kits native in Logic. A very important part is played by the different tambourines that resemble the clashing seashells sound that is distinctive of TNC. The presence of different percussive instruments that followed different rhythms than the main of the timpani also gives the impression of a multitude of players. This was to counterbalance the human shouts present in TNC.

Side chain compression synchronized to the timpani's track was applied to all the other tracks, except for the one of the percussion. This was necessary to reproduce the "breathing" sound that is characteristic of TNC.

Time analysis of MOD

Figure 4.4 shows the spectrogram of the Modified Song.

0:00-0:27 Same as TST.

0:28-0:40 Drone voice on the left side starts, the one on the right starts around 0:32, they both start with a C on the F chord. C is the 5th of the F chord (F A C), but since it starts and is kept sustained, the resulting feeling is unpleasant.

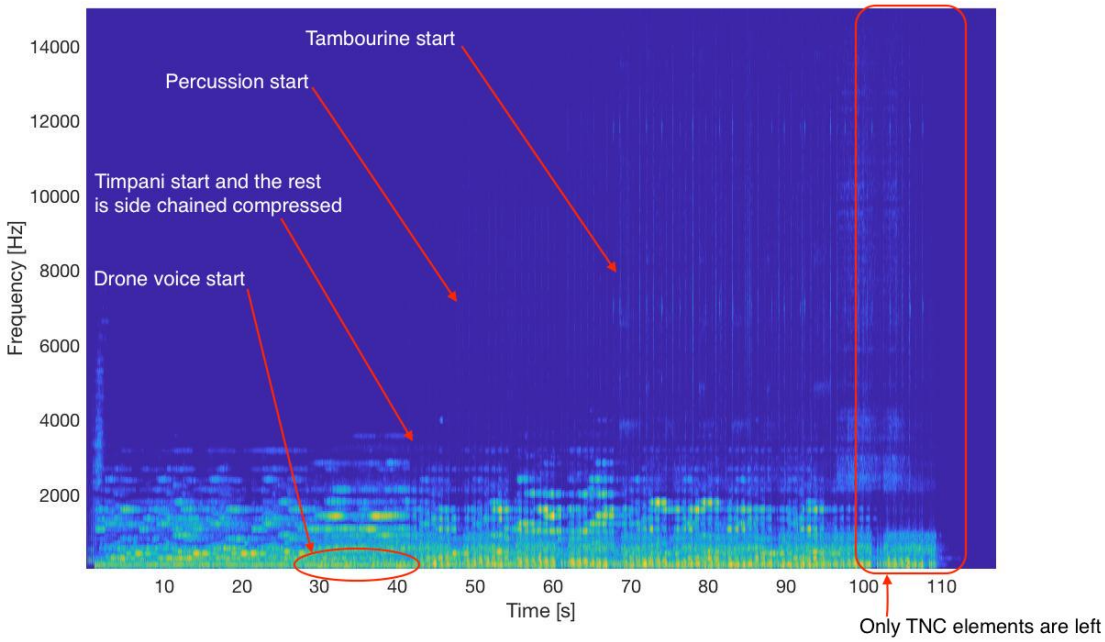

Figure 4.4: Modified Song's spectrogram, Frequenct[Hz] over Time[s].

0:41-0:47 Timpani start together with the bass synthesizers. The side chain compression on all the playing tracks is triggered by the timpani to start the 'breathing' effect. The bass synthesizers play notes that are present in the chords of TST, but alternates them with the ones half a tone above or below. This resembles TNC and at the same time creates tension through dissonance [37]. In this section the chord in TST is C and the bass synths play C and C#. The drone voices follow the bass synthesizers alternating C and C#.

0:48-0:55 Percussion start, they are not part of the side chain compression.

0:55-1.07 The TST harmony goes on the G chord, the bass synthesizers and the drone voices play D and C#.

1:08-1:41 TST goes back to C. Tambourine start boosting the high frequency content of the song.

1:42-1:57 All the TST elements fade out (freeing the mid range) and only TNC elements remain.

To summarize, the four stimuli that will be used for this study are DBD, TNC, TST and MOD. As above mentioned, we expect the imageries that were depicted for DBD and TNC in the pilot test (Chapter 2) to be confirmed by the results of this test. The overall imagery described for TST is expected to be similar to DBD. Therefore inclusive of vast landscapes, space, sunlight of some sort (e.g. sunrise, sunset, etc.), water (e.g. sea, river, lake, etc.), with blue and yellow as predominant colours. The general mood is expected to be positive, relaxed, not necessarily happy or cheerful, but definitely mellow and peaceful. It will be especially interesting to notice if elements that were created for TST –in line with features of DBD– will be linked to the same visual scenes as in the original song. Regarding MOD we expect imageries to present a turning point when TNC elements start. Specifically, we suppose that the scenery will be characterized by DBD or TST elements in the first half of the song and then that TNC visual elements will gradually become predominant. Thus, we expect the descriptions to start with a positive image (e.g. natural elements, bright, nothing to worry about) and then that the image is disturbed by unsettling elements (e.g. a threat, something dangerous approaching, a fight). These elements will contribute in turning the overall perception of the song towards a more anxious environment therefore closer to TNC imagery.

In the next chapter it will be illustrated how we planned to retrieve these information through an online questionnaire.

Chapter 5

Survey Design

Once created the new stimuli we developed a new online survey using soScisurvey [26]. The test was conceived keeping in mind the guidelines explained by Reips et al. in their 2002 article [39]. Using the internet as a medium for the experiment potentially grants the vastest amount of participants and allows to have a wider distribution of background characteristics. Reips suggests that "the behavior observed [...] may be more authentic and therefore can be generalized to a larger set of situations"[39]. This is due to the fact that internet based surveys give to candidates a higher impression of voluntariness. As consequence of this, participants feel less affected from the perceived threat of an experimenter that expects results from them, thus reducing the level of psychological reactance –change in candidates behavior caused by the awareness of participating to a study [39] [11].

The survey was distributed through social medias (Facebook, Linkedin) and auditory lists. In the test candidates were encouraged to spread the word to the people around them without anticipating the real purpose of the questionnaire in order to avoid any specific bias. The questionnaire was pilot tested by 4 selected candidates before being sent out.

In the following sections it will be explained more thoroughly how the questionnaire is structured. We will give reasons for the questions asked as well as highlight the studies that were used as basis for formulating them.

5.1 Introduction, Candidates Background and Context

In the introduction a short presentation of the study was given. It was explained that the experiment included a 2 minutes listening test and that the total length of the survey would have been of 15 minutes. Participants were asked to take the test in a calm listening environment. After a couple of warnings from the candidates about the not complete compatibility of one of the tasks with mobile phones, we added the recommendation of taking the survey on a laptop.

Candidates were advised that their participation was completely voluntary, that answers were completely anonymous and that data will be stored only for the length of the research time. They were asked to not retake the test once finished because that wound taint the data [39]. We decided not to mention VMI until before the actual song listening, so that the candidates would not be biased.

Participants age was asked to have the possibility to check if the capacity to experience VMI varies with age. We asked about the area of origin of participant to investigate on their background and to evaluate the data diversification making sure that they were not all from the same geographical area. If testers provenance is different, we will have the opportunity to verify that similar imageries can be fostered in people with supposedly different cultural background. Therefore, it is crucial that not all the participants are from so called western countries [1] and that we have also participants from Africa, Asia and Oceania. This would give more validity to the assumptions that will be drawn after the results' analysis. Candidates were asked about the context they found themselves in while taking the questionnaire.

In order for the participant to adjust the volume of their sound reproducing peripheral device a player with a test chord was placed before the listening test's slide.

5.2 Listening Test

One of the four stimuli (DBD, TNC, TST, MOD) was randomly assigned to the tester when the survey's link was opened. The instruction of this page was "Please carefully listen to this two minute song. Feel free to close your eyes if it helps you concentrate.". We supposed that encouraging people to close their eyes would reduce the possible distractions due to external factors and increase their focus on the song. As mentioned in the previous chapter, it was decided to use DBD and TNC (also present in the pilot test) in this survey and not to restrict the stimuli only to the newly created.

5.3 Familiarity

We asked candidates if they were somehow familiar with the song listened in order to avoid any possible bias for the imagery perceived. As mentioned before, episodic memory and any other kind of connection that could be related to each individual's experience with the song might lead to an overlap of imageries.

5.4 Visual Mental Imagery

Participants were asked to state if the music made them visualize anything in their mind. 'Yes' and 'No' were the only possible answers. If they answered positively, they were asked to describe as best and detailed as they could what they saw. On the same page it was specified that there were no right or wrong answers to encourage a more truthful answer. Every candidate —independently from what they answered to the previous question— was asked to select the words that could best describe a scene or setting that would fit the music they heard. They could choose among 47 descriptive elements. As mentioned in chapter 2, these elements were not only the 35 used in the pilot test but also the nouns, verbs and adjectives most frequently used in the descriptions given by participants for TNC and DBD in the pilot test. These terms were added to the previous 35 reaching the total of 47 words that are displayed in table 5.1. Compared to table 2.1 we added: Planet Earth, In heaven, In the sky, In a valley, In a tunnel, Around a fire, During a ritual, Before a war, Floating, Peaceful, Tribal, Contrast.

Current Test: 47 descriptive elements			
Setting	On open water	On the beach	In a car
	On a mountain	In the desert	On a train
	In outer space	In the city	On a ship
	In the forest	Inside a confined space	In heaven
	Plane Earth	Before a war	In the sky
	In a valley	In a tunnel	around a fire
	During a ritual		
Weather conditions	Rainy	Cloudy	Sunny
	Snowy	Foggy	Frozen
	Stormy	Warm	
	Cold	Hot	
Colour and consistency	Blue	Green	Red
	Yellow	Shiny	
	Light	Dark	Dense
	Empty	Opaque	
Time of the day	Morning	Evening	Afternoon
	Night		
Miscellaneous	Floating	Peaceful	Tribal
	Contrast		

Table 5.1: Descriptive elements that could be selected in the current test.

5.5 Emotional State

In Chapter 3 we highlighted how individuals emotional state might influence their perception of music. The Positive and Negative Affect Schedule (PANAS) [47] can be used to measure the participants' affective state. We suppose that the PANAS could clarify whether the candidates response to the song was influenced by their emotional state. However, after familiarizing with this scale we decided that the range of moods that it targeted were too diverse from what we would expect from someone participating to an online test. In fact, most candidates would probably not participate to an internet survey if they were very agitated, in a hostile mood, or occupied being ashamed of a recent event. Therefore the risk was of having a scale that was too intense compared to what we were looking for.

This was the main reason we opted for the Geneva Emotiotional Music Scale (GEMS) scale [50]. Its description explains that the model "accounts for music-elicited emotions better than the basic emotion and dimensional emotion models." [50]. This scale shifted the focus from the participant's affective state to the emotion that music elicited in the participant. This might be seen as a deviation from the original purpose, however it is not since –in order to investigate VMI– we are mainly interested in the emotions of participants in relation to the song. Furthermore, in the same paper we found confirmation for our choice to discard the PANAS scale, namely: "Among many emotions habitually experienced in day-to-day life, several were reported only very rarely in response to music. Specifically, guilt, shame, jealousy, disgust, contempt, embarrassment, anger, and fear—these and other negative emotions—were reported to be regularly experienced in everyday life but to be practically never aroused by music."[50]. The GEMS scale was formed after finding that the affect aroused by music can be discerned into several units. Specifically, this model presents nine emotions factors (most of them are positive) and contains emotion categories (e.g. wonder, nostalgia, and transcendence).

Unfortunately we could not get our hands on the original scale. Therefore, we decided to adopt the one used by Lahdelma and Eerola in their paper on emotional qualities conveyed by chords [25]. Similarly to the PANAS, this scale was composed by 9 items that were adapted for the purpose of our test. It was particularly appreciated that "liking" is one of the parameters that are kept into consideration in the scale. In fact through it participants can express their opinion towards the song. The specific scale can be found in Appendix B on Figure B.13.

5.6 Empathy

Empathy was also introduced in Chapter 3 as a factor that played a role when individuals listen to music. Vuoskoski and Eerola in their 2012 paper [45] argued that

people with high trait empathy tended to have stronger imageries. This was due to their inclination to better enter the music's mood. Empathy can be measured with the Interpersonal Reactivity Index (IRI) [12]. The IRI purpose is to measure the dispositional empathy of an individual. It takes into consideration the concept that empathy is a mixture of separate but related components. The IRI consists four sub-scales that gather the separate aspects of empathy, namely: "The Perspective Taking (PT) scale measures the reported tendency to spontaneously adopt the psychological point of view of others in everyday life. The Empathic Concern (EC) scale assesses the tendency to experience feelings of sympathy and compassion for unfortunate others. The Personal Distress (PD) scale taps the tendency to experience distress and discomfort in response to extreme distress in others. The Fantasy Scale (FS) measures the tendency to imaginatively transpose oneself into fictional situations."[12].

In order to reduce the overall timing of the survey we decided to reduce the number of components of the scale to 8 (2 for each facet of empathy). The specific scale can be found in Appendix B on Figures B.14 and B.15.

5.7 Listening Behavior

Participants listening behavior was investigated using Kamalzadeh's research [21] as guideline. We supposed that if the participant was used to listen to music in conditions where the focus was not completely absorbed by the task then he/she would be more prone to experience VMI. In the research before mentioned [21] activities performed while listening to music follow a distinction based on the level of attention they need. For example "work and commuting (if it is not driving) can lie on two opposing ends of this spectrum, with work needing very high attention from the listener and commuting needing much less."[21]. The specific scale can be found in Appendix B in Figures B.16 and B.17.

5.8 Musical Sophistication

Musical sophistication was another factor worth looking into for exploring the nature of VMI. It was important that candidates were capable of distinguishing the different musical elements of a song.

In 1953 Revesz [40] used the term 'musicality' to denote "the need and the capacity to understand and to experience the autonomous effects of music and to appraise musical utterances on the score of their objective quality (aesthetic content)". Hallam and Prince in their investigation on 'musical ability' [15] found that participants affirmed that being actively involved in music making by singing or playing an instrument was an indicator of musical ability. However, "receptive activities such as listening, appreciating and responding to music" were also

acknowledged as indicative of musical ability. Their final definition for musical ability was that it encompassed "aural, receptive, and generative skills which are integrated to enable music to be composed, performed and listened to.". Therefore, it was not strictly connected with the ability to play a musical instrument, but more with the relationship an individual had with music itself.

We considered adopting Ollen Musical Sophistication Index (OMSI) [33] as scale. In this case musical sophistication was considered a synonym of musical ability. It was decided to discard this index because it focused on the experience of candidates with musical instruments and voice. The Goldsmiths Musical Sophistication Index [30] was finally adopted. It is considered a short and reliable instrument that allows to evaluate testers without sacrificing its properties. The original Goldsmith's questionnaire is composed of 39 questions, for time reasons we decided to reduce the number of elements to 10 and to adapt them for our study. The aim of these questions was to understand whether if testers had a developed musicality, independently from their formal training.

The specific scale can be found in Appendix B in Figures B.18, B.19, B.20, and B.21.

5.9 Conclusion

On the conclusive slide participants were thanked for completing the survey and were informed regarding the real focus of it through these lines: "The survey incorporated questions from existing surveys related to empathy and emotional responses. However, an additional focus which was not mentioned at the start is to investigate whether an instrumental piece of music can foster Visual Mental Imagery." A link for the results was also provided in case they were particularly interested.

Participants were asked again to not re-take the test, but since there was no reward and the whole survey is quite long, it is very unlikely that multiple submission will be a problem [26]. We also encouraged to spread the survey to their network paying attention to not anticipate anything about the real objective of the research. Finally we asked to provide any comment regard the questionnaire and their experience.

In order avoid problems related to the length of the survey [39] we tried to keep it shorter than 15 minutes. To make sure that candidates would answer the most relevant questions in the best way possible, we decided to place the tasks regarding VMI in the first half of it[39]. This also would allow us to gather the most important data even if candidates decided to drop out before the end.

The whole survey can be found in Appendix B.

Chapter 6

Results

The survey was launched online on April 15th, 2019 and data analyzed was collected until of May 10th. It was completed by 154 candidates. As mentioned before, the questionnaires corresponding to different songs were randomly assigned by soScisurvey[26]. Each questionnaire had equal probability (weight) to be assigned at the beginning. Weights were then modified so that almost the same amount of candidates completed the surveys with different stimuli with a focus on those created for this experiment. Among the participants, 19 indicated that they took part to the pilot test or were not sure about it. To avoid the possibility of gathering biased data, the answers from these candidates were excluded from the analysis –that was carried out for the remaining 135 testers. Table 6.1 shows the distribution of participants in the different surveys.

Song	Listeners	Familiarity	VMI Yes	VMI NO
DBD	36	0	29 (80.6%)	7 (19.4%)
TST	35	0	21 (60%)	14 (40%)
TNC	36	0	28 (77.8%)	8 (22.2%)
MOD	28	1	20 (71.4%)	8 (28.6%)
Total	**135**	**1**	**98 (72.6%)**	**37 (27.4%)**

Table 6.1: Overview of Gathered Data. Participants experienced VMI mostly with DBD (% of yes over song's total listeners). Only one participant reported to be familiar with the listened song.

6.1 Participants

The average age of those who completed the test was 36.9, with a standard deviation of 12.2. The youngest candidate was 14 years old and the oldest 73.

As the pie-chart in figure 6.1 shows, the survey was completed by candidates

that grew up in very different countries. Most of participants were originally from Italy (n 19) and India (n 19), followed by US (n 14), UK (n 11) and Germany (n 11). The rest of participants were from Sweden (n 7), France (n 6), Serbia (n 6), Mexico (n 5), Brazil (n 4), Australia (n 4), Turkey (n 4), Canada (n 2), South Korea (n 2), Greece (n 2), Switzerland (n 2), Austria, Estonia, Norway, Slovakia, Tunisia, Ireland, Malta, Poland, Spain, Colombia, Costa Rica, Chile, Palestine, Nigeria. Despite this wide spectrum of countries the participant that grew up in western countries [1] are more than the 79%. Therefore, we have a good data diversity from the countries point of view but not if we consider the cultural background.

The 52.6% of candidates took the test at home, in their bedroom, living room or kitchen. The second biggest portion of candidates (37%) took the test in their office or workplace. Other locations mentioned were recording studio, taxi, cafe, airport, shopping mall, train, university, library, outside, train station. The 13.3% of participants indicated that the place they were in was somehow noisy. The 39.3% wrote they were alone. The 11.1% of the candidates specified to be in a silent place while the 3% of them reported to be in a relaxed state.

Pre-Screening of Data

Most of the candidates that were familiar with the stimuli were also part of the pilot, so their answers were not analyzed. As table 6.1 displays, among the 135 participants left only one candidate was familiar with the stimuli he listened. This was with MOD, for which the tester said it reminded him of Iosonouncane, the composer of TNC. Anyhow, the imagery that he/she described was not really affected the way we expected. In fact, we expected that being familiar with the song listened would foster the candidate by moving toward the direction of an imagery related to the title of the song or the occasion when they first heard it for instance. This is not the case for this candidate that wrote "I see man on a desk working on this track and ancient men growling to traditional chants".

6.2 Visual Mental Imagery

Amidst the 135 participants, 98 reported experiencing VMI, corresponding to the 72.6%. A chi squared test showed that the result was significantly different from chance ($\chi^2 = 27.6$; df = 1; $p<0.002$). As can be seen in table 6.1, the song where participants reported more imagery was DBD with the 80.6% of listeners experiencing it. In the following we will give some examples of the verbal descriptions given by participants.

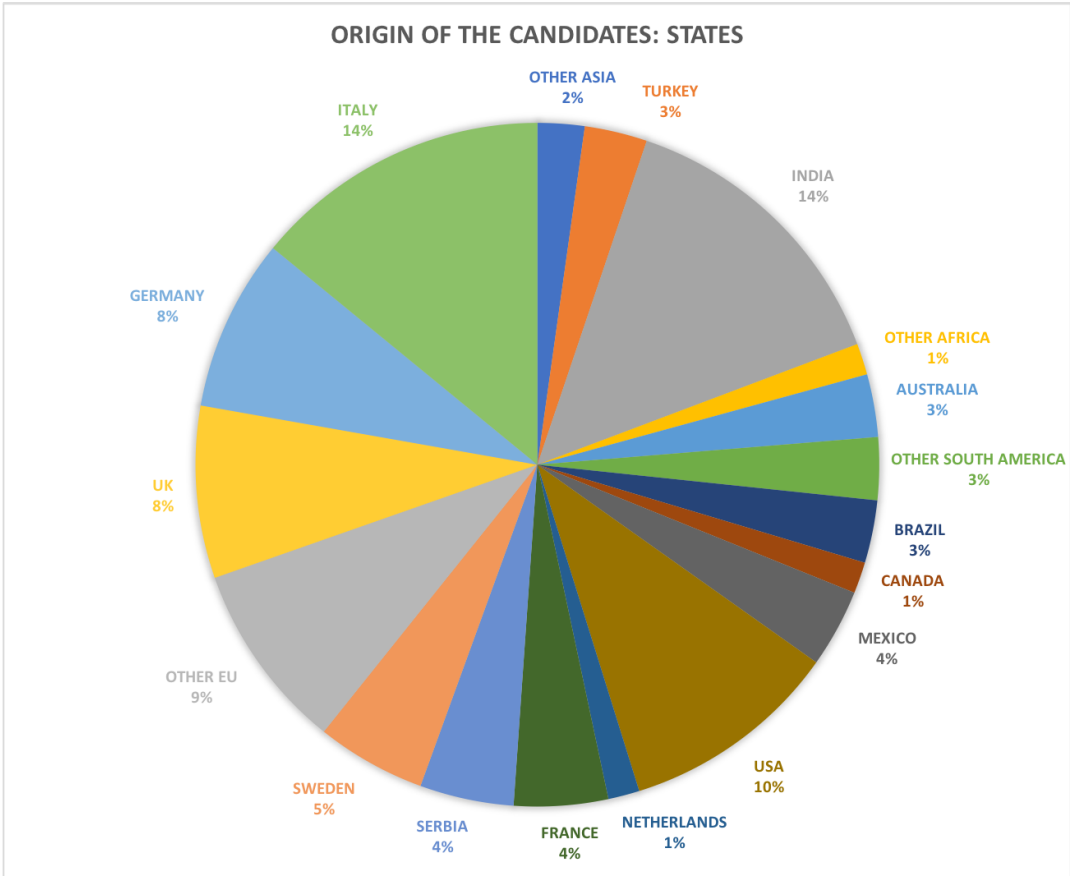

Figure 6.1: Country of origin of the candidates, this is not necessarily where they made the test.

Verbal Descriptions of VMI

As explained in the previous chapter, the 98 participants that experience VMI were asked to describe what they saw in the most detailed way. Every description was carefully read and interpreted, the analysis consisted in counting the nouns that were recurring in more than one description, synonyms and words with similar meaning were counted as different words. In the following paragraphs will be reported the seven most used words for every song together with some significant descriptions. All the descriptions and the word counts can be found in Appendix D.

Deep Blue Day From the 29 participants that experienced VMI in DBD, the words that were most used in their descriptions are the following. Number of occurrences in parenthesis, elements in common with TST are highlighted in italic.

Waves(n 7) Beach(n 5) Grass(n 4) Sky(n 4) Breeze(n 4) *Green*(n 4) Sea(n 4)

The resulting imagery is essentially positive and filled with natural elements. The colour that was predominantly used is green but indirectly we also have shades of blue, from the sea and the sky. Among the descriptions there are some clear examples of how the image created into the listener evolves with the music and has its own consistent coherence. The second and the fourth also show a pronounced cinematographic approach.

> "*The initial orchestration was* light *and* airy *and made me think of* Spring, *of* Summer. *The bass line created a* slow, relaxed feeling *and as the melody started I immediately started to visualize a* tropical beach *scene. Definitely a* lazy summer *day on the* beach *vibe. The melody seemed to evolve a bit over time and got more of a* country *and* western twang, *but it did not disrupt the summer feeling.*"

> "*A grand entrance scene by a* beautiful young woman *(*fairy*?). The start builds suspense, quickly builds into a high near-crescendo (when i can see the camera revealing her face) and then the subsequent rise and fall of the tempo and volume leads me to visualise her* walking *across a* large frozen lake. *Don't know why I visualised a frozen lake - could have been a* verdant green landscape - *but I visualised a* frozen white lake"

> *"I started of at a* beach *at* sunrise, *watching the* waves *from a distance as the* dark sky *turned* orange *as the sun rose from the horizon. In the 2nd half of the piece, I was, unexpectedly, transported to the* mountains. *I sat with a* cup

of tea *and watched the* snow-clad *mountains at a distance. A* cool, *gentle* breeze *was blowing across my face. It felt like it was mid*-afternoon. *The overwhelming emotion was one of* serenity."

"I immediately pictured a high school prom dance, *based on the style and instrumentation. The sound was muffled at the beginning, as if the "camera" was outside the dance, but then we gradually get closer and zoom in on two teens dancing to the slow number. The lights are bluish and we see flecks of disco-ball color. The boy's hair is gelled and styled carefully and the girl has perfectly-set curls that took her hours to get right. She has sequins on her dress. The ambient, distorted sound on top of the more acoustic feel made me think that we are seeing this as a memory, or that something bad is about to happen...will zombies or aliens break into the dance, and the camera follow them in slow motion as they crash the party? Or, maybe this is a memory resurfacing years later, when the girl is an old woman with Alzheimer's and has a rare lucid moment listening upon hearing "their song". Anyway, there is something bizarre or dark that taints the otherwise mellow, nostalgic feel."*

However, not every participant had imageries in line with the main stream. These are some examples:

"A busy urban street, walking along down it with lots of lights and cars rushing and excitement, but with internal cheerful peace."

"A restaurant and a soothing music in the background. Some of the notes of the music were irritating though."

"Unfolding evolving helix."

Test Song The words that were mostly adopted by the 21 participants who experienced VMI listening to TST are the following. Number of occurrences in parenthesis, elements in common with DBD are highlighted in italic.

Green(n 3) Blue(n 3) Space(n 3) Stars(n 3) Floating(n 3) Relaxed(n 3) Clouds(n 3)

Among these words, only Green seems to be in common with DBD but in fact Stars, Blue, Relaxed and Clouds were used more than once to describe also DBD (see Appendix D). The general imagery is positive also in this case, even if different shades are depicted in the descriptions. This can be a first confirmation that the two songs foster similar imageries. Here are some of the most relevant descriptions for TST:

> "Floating, flying, *looking down on* green fields, dreaming, clouds, relaxing, loving, laughing, feeling, tender, *yes. I didn't really 'see' a specific scene. The music evoked these feelings in me and I had a sense that I was possibly flying, looking down on a landscape but nothing more detailed than that.*"

> *"I would associate this music with an impression of* floating in space*. I saw a kind of* night sky *with* stars*, a bit like a "galaxy-type" visual pattern. The stars would appear and disappear and they would be of different colours with different level of focus, some would be blurry while others would be sharp. The general feeling would be quite* positive*."*

> *"I closed my eyes, felt that I am* floating *on* cool blue water-waves*, filled with* serene calmness*. Quietly proceeding ahead, I am all alone, felt, music should not stop at all. Soothing, my soul."*

> "Forest *or natural landscape,* lake *or* river *in distance, people discovering the area. People* marching *to a slow beat. Seemed to be some sort of* funeral march*. And it was* rainy*. But people were* not sad*. It was some sort of gathering of people."*

Like in DBD, the imagery is never truly negative, natural elements are present but there is a bigger focus on feelings and sensations. Below you can find some descriptions that differed the most from the main image.

> *"I see a cloud in front of my eyes, like being in front of a rain cloud, some pink neon flashes and everything happens in a black space."*

> *"Deserted and broken street at Christmas, Untidy room that's left, Something not quite happy just happened, but everyone accepted the situation, Broken tape record."*

> *"Blue/gray smoke diffusing in the dark, over an oriental city."*

Tanca The terms mostly used by the 28 listeners that reported VMI for TNC are the following. Number of occurrences in parenthesis, elements in common with MOD are highlighted in bold.

Battle(n 8) **People**(n 8) Fire(n 7) Warriors(n 4) **Ritual**(n 4) Dance(n 4) **Tribal**(n 4)

Overall, descriptions are mostly negative with at least one element that constitutes a threat or creates some sort of discomfort.

"For the first thirty seconds or so, nothing. Then, when some sounds resembling the clanking of chains *began to sound, I saw an image of chained* slaves *being made to constantly do something - like rowing a massive* ship*, or feeding* coal *into a furnace. At first it seemed like this sound was only coming from one direction,and then another set of such sounds kicked in from the opposite direction and that was when this scene took a solid form in my mind. I also pictured a few* armored soldiers *monitoring these workers."*

"Initially, scanning across a desert plain *with* rocky cliffs *in the background, as someone rode on a* horse*. As the noise built, it turned to a* battle *scene, with* swords *clashing and* cannons/gunshots*. Latterly it morphed into a picture of a* prison *with cell bars being rattled and* prisoners *shouting riotously."*

"I saw a football stadium *with* herds *of* supporters/hooligans *slowly but quite* chaotically walking around*. It is an* apocalyptic *scene and some bassi profundi and a band are stimulating them by creating a trance-like atmosphere with ominous pedal notes and a constant but driving rhythm. The scene is* dark *and* threatening*."*

"A group of heavy *set* Mongolian men *at a* camp *of sorts in a* large field *of* sweeping yellow-green grass*. There were also* horses *with shiny* discs *of a silver-cover* metal *attached to their* saddles*. Their movement made the tambourine-like sound in the background of the piece of music."*

As anticipated most descriptions are grouped by the presence of an unpleasant/negative element. Here follow three descriptions that move away from the scene generally imagined.

"Disco arena."

"I thought of a brawl in a pub, with glass bottles breaking on the floor."

"I saw in my mind a battle between monsters in the space."

Modified Song The preferred words that the 20 listeners that experienced VMI used to describe MOD are the following. Number of occurrences in parenthesis, elements in common with TNC are highlighted in bold.

People(n 4) Black(n 4) **Ritual**(n 3) Fog(n 3) Mountains(n 3) Drums(n 3) **Tribal**(n 2)

The majority of scenes described people playing drums or involved in a ritual of some sort or dances. The most significant are those that clearly show a turning point when the character of the music changes, the following four descriptions are an example of this.

> *"During the song I saw* flowers *and beautiful* spring *colours come to my mind. I felt such a rebirth of* nature *which comes with spring. Then at the end the song has become more* heavy *and* gloomy*, as a* storm *has arrived to* disturb *this beauty of nature. So I could see a real scene of a living nature."*

> *"I saw* forests *in my mind. Lots of* trees *and different* shrubbery*. Some* animals *like* hares *dashing around and a* spirit *of some sort moving around the trees. When the tune changed I saw the* sea*. It was* rough *and* stormy*. Awe* inspiring*. It was alive. It was as if there was a* strong force *that moved with the* sea*."*

> *"It was an individual telling a story of a person who lost something and then comes over it. The intro part felt like the beginning of setting of the story. The tempo then began visualizing the story of his* troubles *and then the verse section prior to outro felt like he was overcoming his troubles."*

> *"Probably a* wood*, with* sun rays *that filters through the* trees' branches*. Then someone arrives, someone* heavy*,* tall*, almost* scary*."*

The imageries start with a positive acceptation and then they evolve in something more serious, a threat or a demonstration of strength. Here follow three descriptions that were mostly different from the VMI generally perceived.

> *"A night club with disco lights and fog, and people dancing."*

> *"flashes of white dots and blueish fog in front of an all-black background"*

> *"Made me feel judged by a supreme court; the sentence was given, and I was found guilty."*

6.3 Descriptive Elements Selection

Differently from the descriptive task (performed only by the 98 participants reporting VMI), all 135 candidates had to select 12 descriptive elements -from the 47 listed on table 5.1- to associate with the song listened. Table 6.2 shows how many words were totally selected for each song, together with the number of participants that listened to the song and the average number of words selected per participants. It is interesting to notice how MOD is the song with more elements

	DBD	TST	TNC	MOD	**Total**
Descriptive elements selected	273	278	250	230	**1031**
Number of participants	36	35	36	28	**135**
Words per participant	7.6	7.9	6.9	8.2	**7.6**

Table 6.2: Number of the selected descriptive elements for each song and average of words selected per participant. Every candidate could select a total of 12 words.

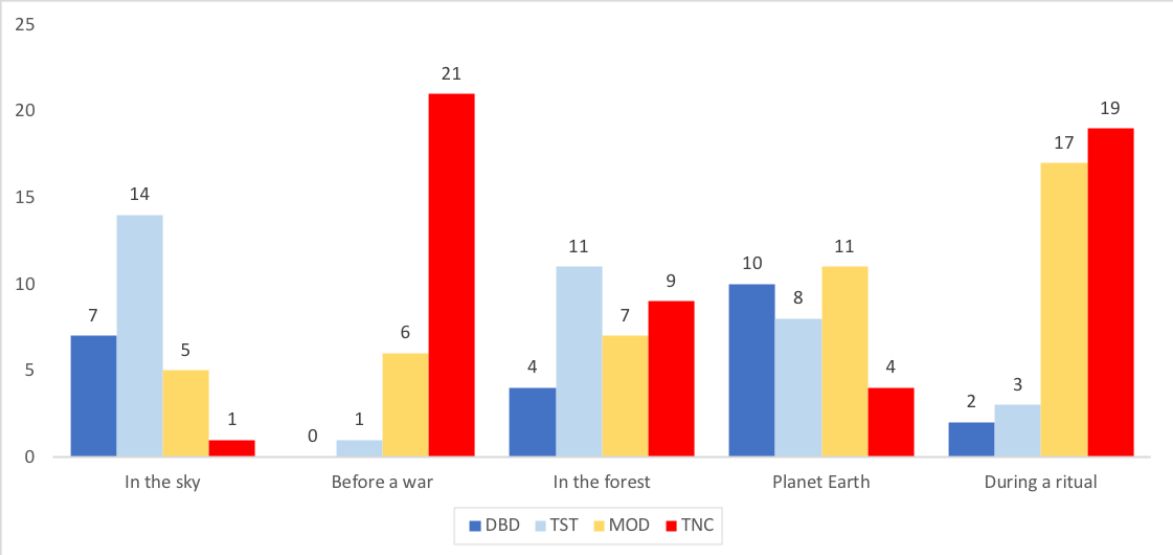

Figure 6.2: The histogram shows the first 5 elements describing **Setting**. These were most frequently selected for each song. It is interesting to notice that 'During a ritual' was selected very often for TNC and MOD. 'Planet Earth' was frequently selected, especially for TST, DBD and MOD. 'Before a war' was mostly chosen for TNC, 'In the sky' was selected especially for TST.

selected per participant, with an average of 8.2 elements selected. MOD is followed by TST with 7.9 elements selected per participant, this is peculiar if we consider that TST is the song that least fostered VMI in listeners (roughly 21 over 35).

Figures 6.5 6.4 6.3 6.2 6.7 6.6 6.9 6.8 6.10 6.11 show the results of the descriptive elements selection. As done previously they are semantically divided into five categories: *setting*, *weather conditions*, *colour and consistency*, *time of the day* and *miscellaneous*. In the same category, plots are in ascending order counting the times they were selected overall.

The most selected words for the single songs are reported in table 6.3. An element is considered *most selected* when it was chosen by at least *one participant out of four*, therefore the 25% of participants that listened to that stimuli. The number in parenthesis next to the element corresponds to the times it was selected. From table 6.3 we can see how 8 out of 16 most selected elements for TST are in common

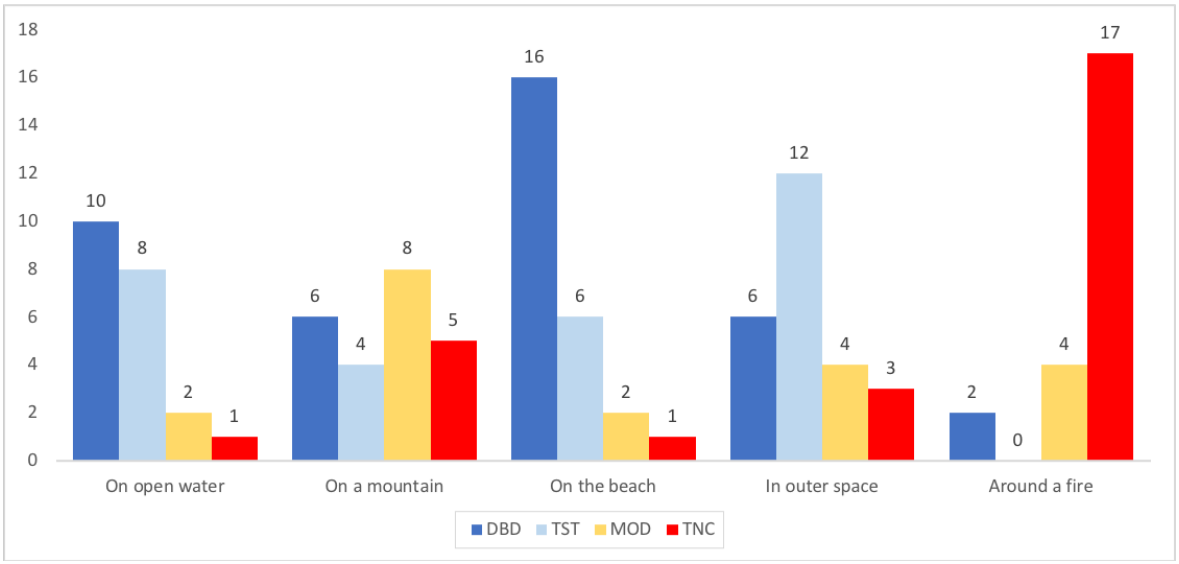

Figure 6.3: The plot shows the second 5 **Setting** elements that were selected for each song. Compared to the previous plot, the y-axis has a different scale. 'Around a fire' was chiefly chosen for TNC, 'In outer space' was selected especially for TST, while 'On the beach' was preferred for DBD.

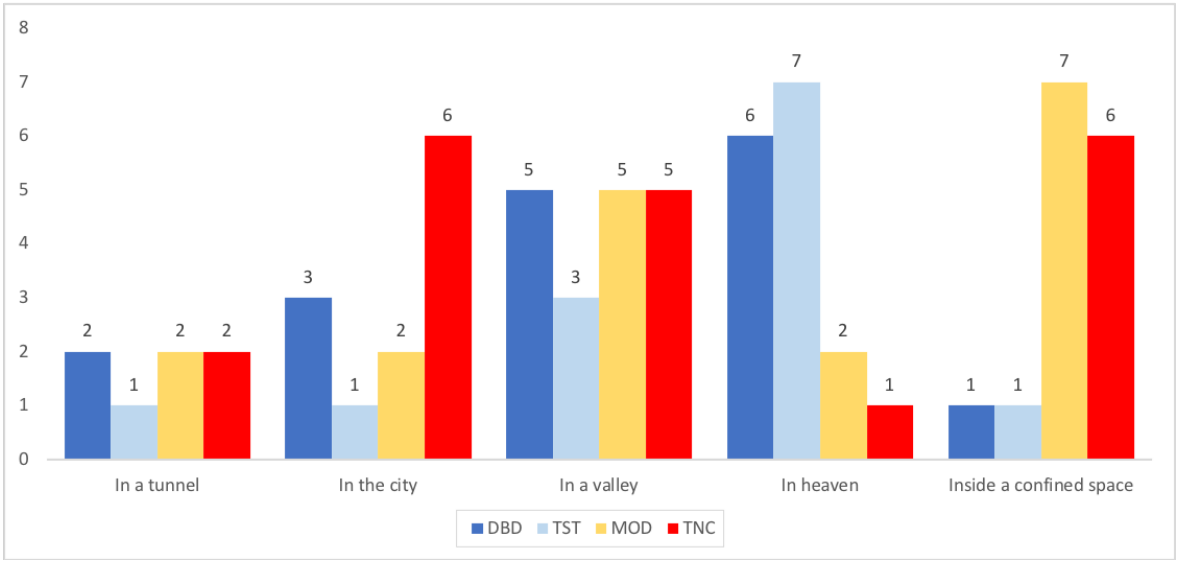

Figure 6.4: The diagram displays the third 5 elements describing **Setting** that were selected for each song. Compared to the previous plot, the y-axis has a different scale. Note that 'Inside a confined space' was selected more for TNC and MOD, while 'In Heaven' was chose more for DBD and TST. It is also interesting to see how 'In the city' was mostly selected for TNC.

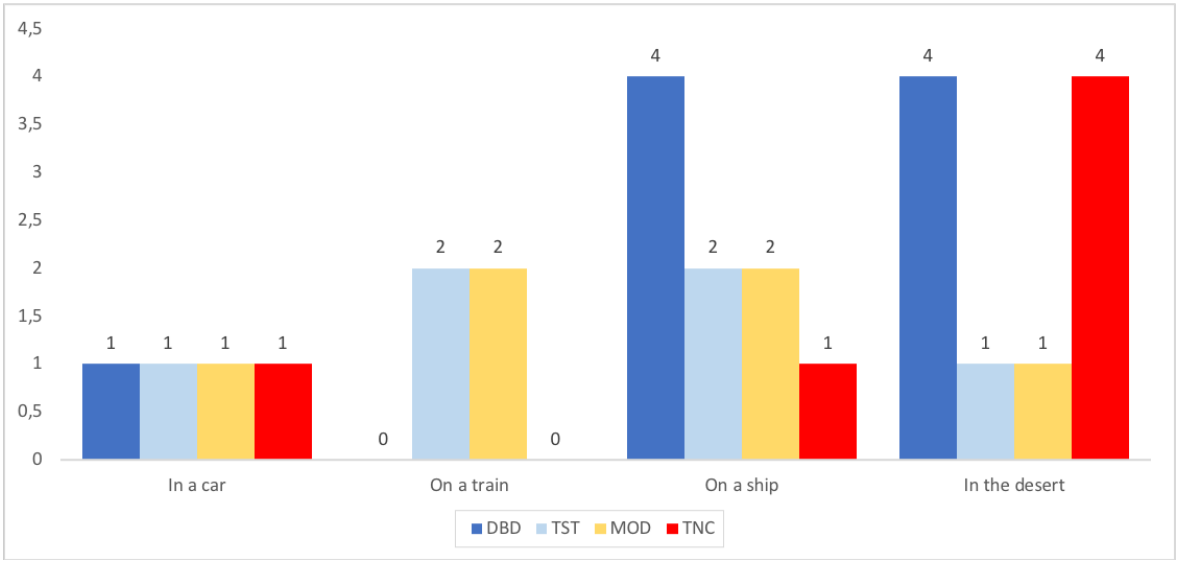

Figure 6.5: The plot shows the last 4 elements describing **Setting**, these were least selected. Compared to the previous plot, the y-axis has a different scale.

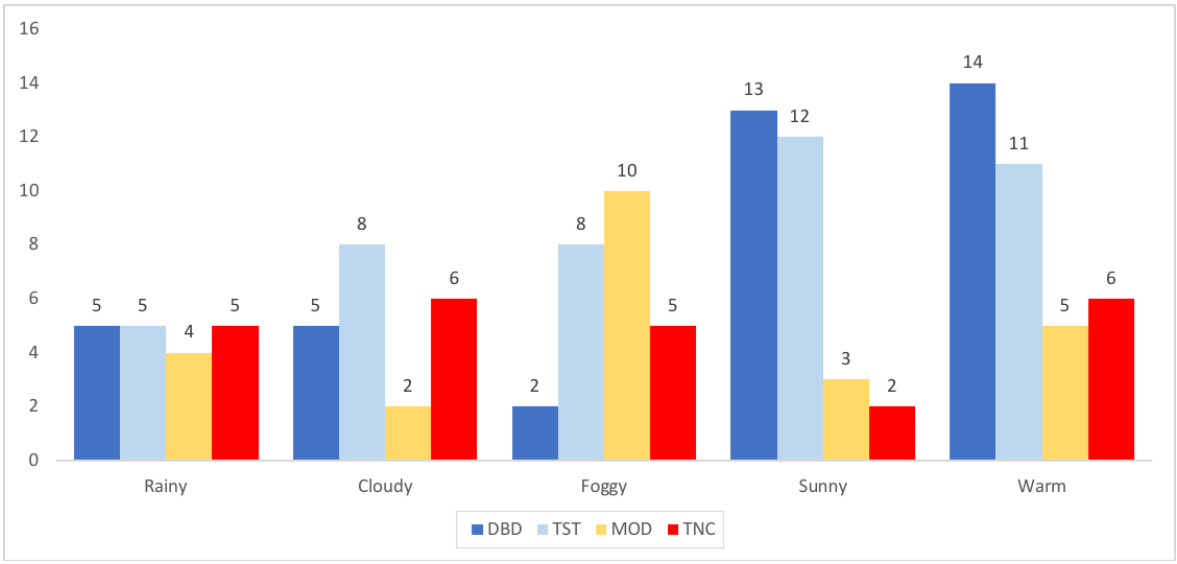

Figure 6.6: The histogram displays the element describing the **Weather Conditions** most selected by the candidates for the song they listened. Notice that 'Warm' and 'Sunny' were mainly selected for DBD and TST. 'Foggy' instead was chiefly chosen for MOD and TST. 'Cloudy' was frequently selected, but mostly for TST.

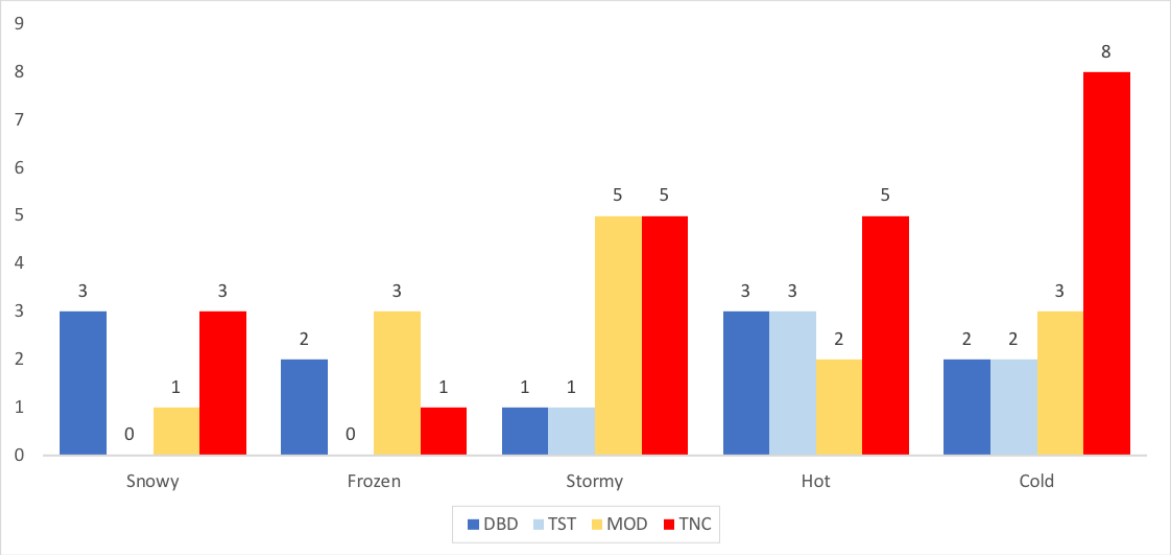

Figure 6.7: The diagram shows the **Weather Conditions** that were least selected by participants for the song they listened. Compared to the previous plot, the y-axis has a different scale. It is worth noticing that TNC was associated above all with 'Cold'.

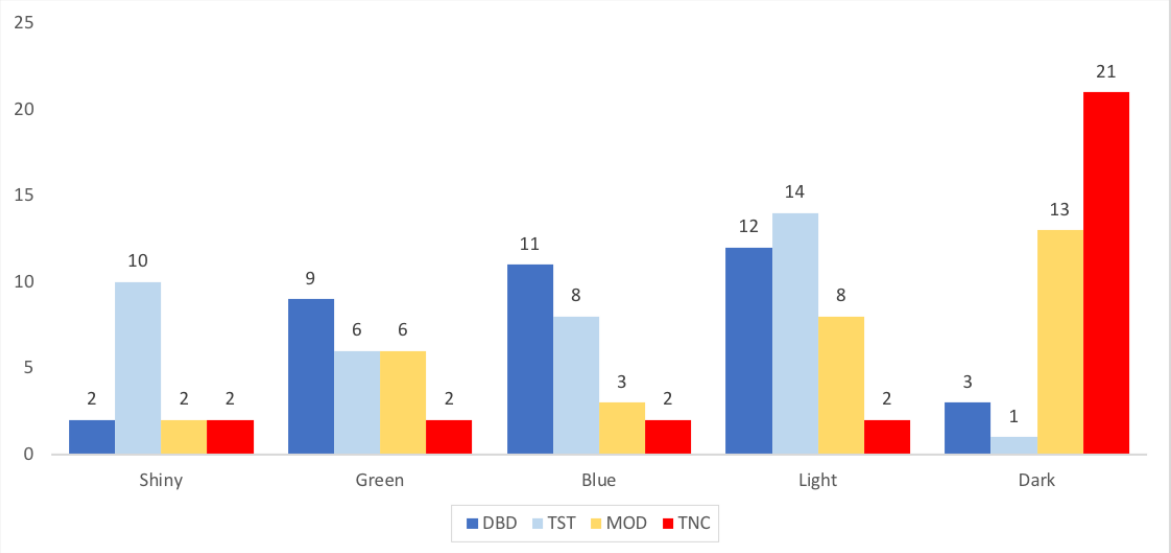

Figure 6.8: The plot shows the **Colour and Consistency** elements that were most selected. 'Dark' was selected above all for TNC and MOD. 'Light' instead was selected for DBD and TST but interestingly also for MOD. 'Blue' was also selected for DBD and TST, while 'Green' was frequently chosen for DBD, TST and MOD. 'Shiny' was both associated with TST.

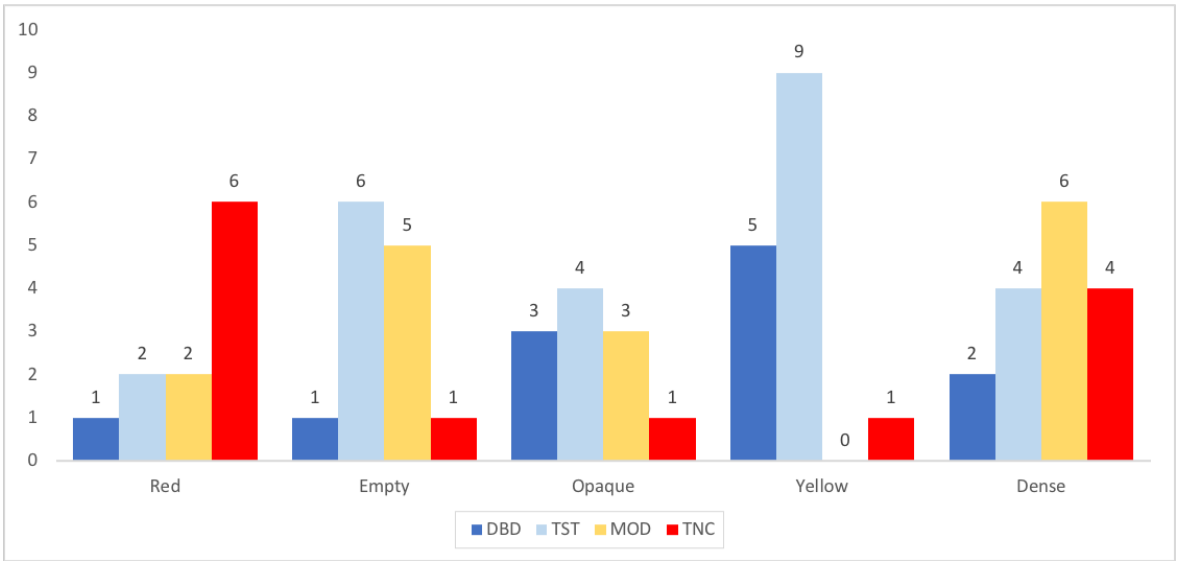

Figure 6.9: The histogram displays the **Colour and Consistency** elements that were least selected. 'Yellow' was connected with TST, 'Red' is the colour mostly associated with TNC.

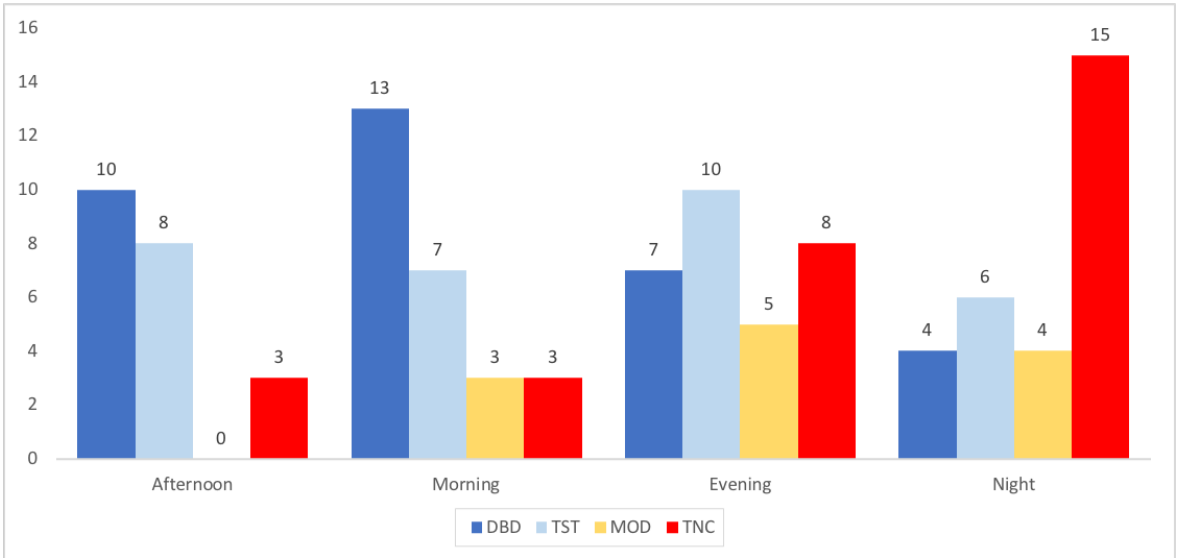

Figure 6.10: The plot shows the elements that could be used to specify the **Time of the Day** that was associated with the stimuli listened. 'Night' was associated chiefly with TNC, while 'Morning' and 'Afternoon' were more selected for DBD. 'Evening' was selected frequently, but mostly for TST.

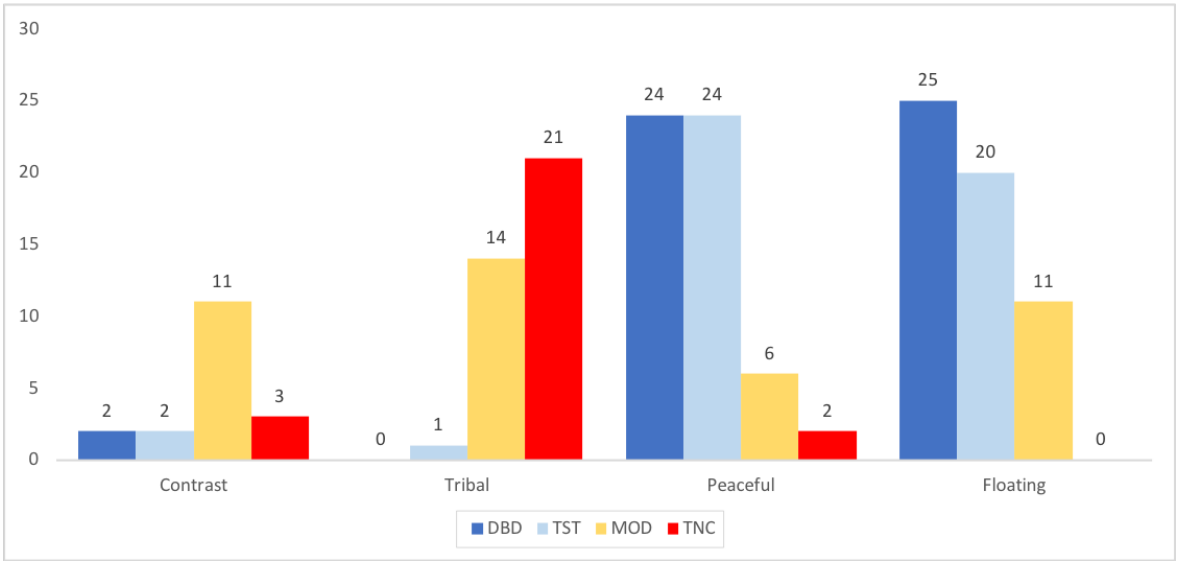

Figure 6.11: The diagram shows the descriptive elements that did not fall in any of the above categories, therefore called **Miscellaneous**. As expected 'Floating' and 'Peaceful' were chiefly selected for DBD and TST, while 'Tribal' was chosen primarily for TNC and MOD. It is interesting to notice how 'Contrast' was selected for MOD above all.

with DBD (the ones in Italics). Among these 8 elements, 4 are also in common with MOD. Furthermore, it is particularly interesting to notice that MOD has also 4 most selected elements in common with TNC (the ones in Bold), thus confirming that the imagery resulting for MOD actually features elements from both DBD and TNC.

6.4 Sonic Features

Similarly to the pilot test, not many candidates were able to clearly connect specific sonic features to the visual elements described. Nevertheless, the data gathered are displayed in tables 6.4, 6.5, 6.6, 6.7. Since we posed an open question to avoid any bias from our side, the answers were very diverse thus difficult to code and schematize while preserving the same precise meaning. When the field 'Reported Sonic Feature' is left blank it means that no sonic feature of the instrument was specified by the participant.

Most Selected Descriptive Elements

Song	DBD(36 listeners)	TST(35 listeners)	MOD(28 listeners)	TNC(36 listeners)
Setting	On the beach(16) *Planet Earth(10)* *On open water(10)*	In the sky(14) In outer space(12) **In the forest(11)** *Planet Earth(8)* *On open water(8)*	**During a ritual(17)** *Planet Earth(11)* On a mountain(8) **In the forest(7)** Inside a confined space(7)	Before a war(21) **During a ritual(19)** Aroung a fire(17) **In the forest(9)**
Weather conditions	*Sunny(13)* *Warm(14)*	*Sunny(12)* *Warm(11)* Cloudy(8) Foggy(8)	Foggy(10)	Cold(8)
Colour and consistency	*Light(12)* *Blue(11)* Green(9)	*Light(14)* Shiny(10) Yellow(9) *Blue(8)*	**Dark(13)** *Light(13)*	**Dark(21)**
Time of the day	Morning(13) Afternoon(10)	Evening(10)	Evening(5)	Night(15)
Miscellaneous	*Floating(25)* *Peaceful(20)*	*Floating(24)* *Peaceful(20)*	**Tribal(14)** Contrast(11) *Floating(11)*	**Tribal(21)**

Table 6.3: Most Selected Descriptive Elements, namely those selected by 1 out of 4 participants (25%). The words shared with DBD are in *italic*, with TNC in **bold**. The words underlined are in common for MOD and TST. The number in parenthesis next to the element corresponds to the times it was selected.

DBD		
Instruments Used	**Reported Sonic Features**	**Visual Elements**
Bass		Calm, Warm
Drums		Footsteps
Rhythm Guitar Bass Line		Hawaiian music, Train
Rhythm Guitar		Warm, Cozy
Rhythm Guitar Bass Drums	Muffled	Outer Space
	Rhythm	Walikng
Rhythm Guitar Bass Steel Guitar	Change from F to G	Hawaiian music
Steel Guitar	Slide with bottleneck	Western American music, Hawaiian music, Beach, Summer
		Desert, Serenity, Peace, Chilling at a fireplace
Steel Guitar Synthesizers	Scale played and chord	Morning, Beach, Drizzle, Grass, Hills
Synthesizers	Modulated sound intensity	Waves
	Cloudy hovering pad	Floating, Calm, No worries, Relaxed
	High pitch	Peaceful
		Outer Space, Floating, Dark, In the sky, Urban scene
	Heavy Echo	Shiny
	Initial fade in	Suspence
	Smooth	Floating

Table 6.4: Sonic Features connected to visual elements for DBD. When the field 'Reported Sonic Feature' is left blank it means that no sonic feature of the instrument was specified by the participant.

TST		
Instruments	**Sonic Features**	**Visual Elements**
Bass	Quiet	Peaceful
		Floaty
Drums	Muffled	Bad weather
		March, Ritual, Train, Car, Road trip, Rain, Storm, Cold, Floating
Lead Guitar	High pitch, Long notes	Floaty, Space
Rhythm Guitar Bass Line	Brightness	Afternoon, Cafe del Mar
Rhythm Guitar Drums Synthesizer	Reverb	Space, Clouds
Rhythm Guitar Synthesizer	Reverb, Glockenspiel-like sound	Space movies
Rhythm Guitar	Distortion	Dense, Foggy
Synthesizers	Atmosphere	Planet Earth documentaries
	High Pitch, Glockenspiel-like sound	Buddhism, Mountain, Planet Earth, Sunlight
	Metallic sound glockenspiel-like	Beach, Stars
	Hammond Organ-like sound	Yellow, Foggy, Night, Afternoon

Table 6.5: Sonic Features connected to visual elements for TST. When the field 'Reported Sonic Feature' is left blank it means that no sonic feature of the instrument was specified by the participant.

MOD		
Instruments	**Sonic Features**	**Visual Elements**
	Sound at the beginning	Pure, Positive
	Slow and repetitive structure	Calm and soothing
Drone Voices	Low frequency from middle of the song	Dark, Calm
		Dream, Dark, Tribal
Drums	Repetitive, Simple	Phone Alarm
		Rocks, Deep diving
Drums muffled, Lead Guitars, Synthesizers	Reverb	Dream
Drums, Synthesizers	Deep muffled sound	Floating
Lead Guitars	Long notes	Shiny
Synthesizer	High pitch Glockenspiel-like sound	Sea, Breeze, Fog, Green, Forest
	String-like sound	Blue, Open space, Foggy
	Reverb	Sky, Light
Synthesizers, Lead Guitars	Pentatonics	Mountain, Floating
Synthesizers, Timpani	From the middle	Contrast
Bass Synthesizer	Minor second interval from the middle	Tribal
Timpani, Percussion	Triplet drum beats	Tribal
		People sitting together
Timpani, Drone Voices		Deep, Tribal
Timpani	Loud staccato from the middle	Steps, Tribal
Side-chain compression	From first to second half	Contrast

Table 6.6: Sonic Features connected to visual elements for MOD. When the field 'Reported Sonic Feature' is left blank it means that no sonic feature of the instrument was specified by the participant.

TNC		
Instruments	**Sonic Features**	**Visual Elements**
	Rhythm	Chaos
Drone Voices	Mongolian-like chants, Similar to didgeridoo	Tribal, Valley, Ceremonies
Drone Voices, Timpani		Tribal, Forest
	Repetitiveness	Nature
High pitched voices		Increase tension
High pitched voices, Tambourine	Increased loudness	Battle intensifying
Shouts		War, City, Dense, Soldiers, Ritual, Forest
Shouts, Tamburine, Timpani		Battle
Synthesizer	Discordant harmonies	Darkness
	Side-chain compression	Tension
Tambourine	Metallic sound	City, Pub fight, Drama, Battle, War
	Increased frequency in the middle	Battle
Tambourine, Timpani	Increased frequency in the middle	Tribal
Timpani	Loud from the middle	Fight
		Gunshots, Rain, Warrior marching

Table 6.7: Sonic Features connected to visual elements for TNC. When the field 'Reported Sonic Feature' is left blank it means that no sonic feature of the instrument was specified by the participant.

Chapter 7

Discussion and Future Research

The aim of this study was to answer the two main questions:

1. Can instrumental music foster Visual Mental Imagery?
2. Is it possible to reconnect sonic elements of a piece of music to specific features of the VMI?

An online survey was created to inquire if candidates experienced VMI at all when listening to instrumental music and to ask questions on the nature of the imagery. Different imageries were depicted and together with a task that demanded the selection of descriptive elements from a fixed list, we were able to gather different sets of data that could be analyzed.

Can instrumental music foster VMI?

This query was investigated through a self-report Yes/No question. The results of the survey clearly show that instrumental music can foster VMI: 72.6% of the participants experienced it not randomly. Furthermore, the remaining 27.4% – that declared they did not experience VMI– still were able to fulfill the descriptive elements selection task. Therefore, it is reasonable to assume that even if the music did not foster VMI for them, they created some sort of reference imagery in their mind or at least an opinion on how it should have been. These associations seem to be strong enough to make them decide whether an externally proposed descriptive element is coherent or not with the music they listened.

One could argue that the descriptive elements could influence the participants by giving them suggestions for an hypothetical imagery that otherwise they would have never had. This argument is appropriate, as Juslin [20] explains: "[a] special feature of the imagery mechanism is that the listener is very much able to influence the emotions induced by the music. [...] in general a listener may conjure

up, manipulate, and dismiss images at will". Nevertheless, the bias due to the participants reading the available descriptive elements does not explain why and how they would chose some of them more than others. Hence, it is legitimate to conclude that instrumental music fosters an imagery that for some listeners is more explicit and vivid, while for others remains implicit and subconscious. However, when comparing the music with external descriptive elements, the implicit image is partially revealed to the extents that are relevant with these elements.

Is it possible to reconnect sonic elements of a piece of music to specific features of the VMI?

The inquiry was developed through two different tasks: -describing the imagery visualized, if any. -selecting descriptive elements from a fixed list. From the descriptions given (found in Appendix D) the following statements can be drawn.

DBD and TST fostered a similar imagery even though not exactly the same. DBD shows to remind cheerfulness and carefree feelings, together with suggestive nature like the beach, the sea, green fields. This is in line with Osborne's findings [34], according to which some imageries with nature scenes (e.g. sky, sun, ocean) and out-of-body experiences (e.g. floating above the Earth) were more frequent than others. Juslin et al. [20] suggested that the reason for this is connected to the peculiar music selected by Osborne, described as "spacey, synthesized electronic music with simple structure, some free form, and much repetition" [34].

On one hand, TST brings to mind circumstances that are similar to DBD. On the other hand, it also evokes situations that can not properly be defined happy although never truly sad (e.g. the fourth TST quotation in Section 6.2). The scenes depicted for TST present settings with natural elements as well as sensations related to space (e.g. floating, flying) and images of stars and galaxies. Therefore, we can affirm that even if TST was composed along the same lines of DBD, the differences in timbre of the sounds used and the slightly diverse harmony might have generated a little variance in the perceived image.

Regarding MOD, a good number of descriptions clearly showed a turning point when the TNC elements started. While the VMI associated to TNC is full of human and artificial elements, MOD's imaginary tends to be mostly natural in the first half of the song, resembling TST. In the second half –when TNC elements start– the scenes described until that moment become darker and more serious. Therefore, a good number of descriptions are characterized by the approaching of a threat, or something intimidating like a storm or a wild animal. Close to MOD's end, only TNC elements are left, the scenery depicted often evolved with the addition of a tribe performing a ritual, or people stomping their feet, but never leaving the natural setting. Thus, we can assume that even though MOD's progression brings it to be very similar to TNC, the resulting imagery still preserves elements fostered by the first half of the song. It is legitimate to affirm that the imagery fostered by

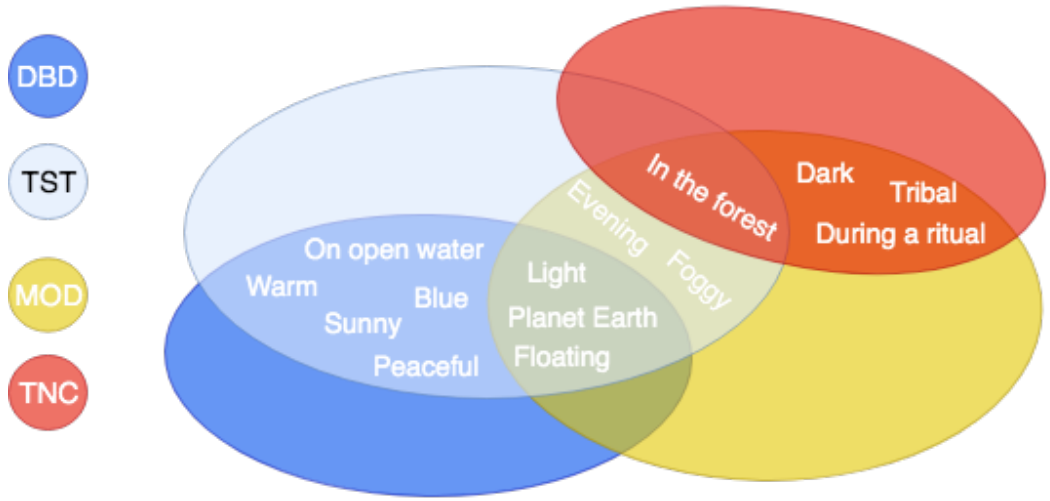

Figure 7.1: The Venn diagram shows the descriptive elements that were the most selected for two o more songs in the descriptive elements selection task.

MOD features elements belonging to both TST imagery and TNC imagery.

These statements are also confirmed by the descriptive elements selection task. Since the elements were fixed, it was easier to confront the data and similarities are evident. The Venn diagram shown in Figure 7.1 displays the descriptive elements that were most selected for more than one song. As expected, for MOD were selected elements that were also selected for TNC and DBD, corroborating the hypothesis that mixing sonic features from songs with different imagery leads to a song which imagery comprehend elements from both the imageries of the initial songs.

Curiously, the element 'Night' was associated with TNC but not with MOD (see Figure 6.10). This might be a further confirmation of what was suggested above, namely that even if MOD ends with only TNC elements, the resulting imagery still features elements from TST. It is interesting to notice that, even though TST and MOD were the stimuli that fostered VMI for the least percentage of participants (see Table 6.1), they were the first two for which there were selected more elements in the descriptive elements selection task (see Table 6.2). This could mean that they both communicate a definite imagery, however it is not as clearly defined as for the two other songs. Nonetheless, when compared with a pre-made list of elements, it becomes easier to define their imagery. One possible explanation for this could be in the shortness of time put into their creation. As a matter of fact, TST and MOD were recorded from scratch and mixed in two days. The work was developed in a home environment not acoustically treated, and not in a recording studio. All these factors might have affected the final results by making it less accurate, therefore its

imagery less explicit.

To summarize, we could affirm that since MOD fosters and imagery that features elements from both TST and TNC, there is a connection between the sonic elements of a song and the visual features of its imagery. Future research might address the question 'what sonic element should be used to foster a specific visual feature?'.

Can we associate a sonic element to a visual feature?

We tried to tackle this question by asking testers to connect sonic features of the song they listened, to visual elements they visualized. The query was an open question and we left an open field to answer, unfortunately this led to blurry answers (see Appendix D). As displayed in the tables of Section 6.4, the same sensations or descriptive elements are cited by different candidates for different sonic elements or musical instrument. Nonetheless, there are still interesting details to notice.

One participant associated the discordant harmonies of the synthesizers in TNC with darkness. As Peretz [37] explains, the mechanism underlying the feeling of unpleasantness when hearing a dissonant sound is innate in humans and strictly related to the perception of screams. Therefore, connected to something alarming like an emergency or a threat. Thus, Peretz concludes affirming that "dissonances are to music what screams are to vocal communication" [37]. Future research could explore to what extent a dissonant sound is perceived unpleasant.

Another interesting fact to notice is how for MOD, different participants associated different elements with the word 'tribal'. Being more specific, mainly percussion and timpani, but also the second interval played by the bass synthesizer line. All the descriptions of the sonic element connected to 'tribal' and 'ritual' specified that they started in the middle of the song, emphasizing the turning point depicted in several VMI. Therefore, this can be considered a further confirmation that adding TNC elements to TST implicate a transformation in the resulting VMI.

In the previous Section (7), we reported Juslin's assumption that spacey synthesizers and electronic music fostered nature scenes and out of body experiences. This is partly confirmed by the imageries reported for DBD and TST. In fact they both features synthesizers of which sounds were designed to be spacey. However, if the imagery evoked by TST features both natural elements and out of body experiences, the one fostered by DBD seems to be more focused on the natural elements but not much on the out of body experiences.

One could say that this task was a more subject to the participants experience. On one hand this is a reasonable argument, on the other the task construction could have been more precise and effective. Furthermore, some candidates expressed their inability to distinguish and recognize the instruments playing. If a similar question was asked during a in-person-experiment, candidates would be able to

show directly to the experimenters what sonic feature they are referring to. Thus bypassing the difficulties in recognizing the instrument or being able to describe the feature. Further, researches should try to address this problem in order to find more solid links between audio and visual.

More Future Research

As expressed in the previous chapter, the sampling was not uniform across cultures. The choice to create the survey in English suggests that language might be the reason why most participants were from western countries. Nevertheless, we had great affluence from India (14%) and overall the described imageries were not far from those described by participants from western countries. Further, work should take into consideration targeted collaborations with universities or researchers from non-western countries to reduce this gap.

In a future research vividness of the VMI could be tested through specific questionnaires like [3] and [28].

For shortness of time we were not able to analyze all the gathered data. In the future more content analysis should be tackled in order to look for correlation between VMI and the factors discussed in Chapter 3. Furthermore a greater number of musical features could be explored by creating more stimuli with a greater musical diversity.

Chapter 8

Conclusion

The aim of this thesis was to shed more light on the still obscure topic of Visual Mental Imagery. This was investigated, through an online survey on whether instrumental music fosters VMI, and if it is possible to establish a relation between musical features and visual elements of the imagery.

Two musical stimuli were created with the purpose of verifying that mixing sonic elements belonging to two songs –that evoke different imageries– would foster a third imagery that presents visual details from both the imageries of the original songs. In total four musical stimuli where used in the questionnaire, the two that we created and two commercial recordings.

The survey was online from April 15th, answers were collected until of May 10th. Data clearly show that VMI can be fostered from instrumental music, it can be more explicit for some of the listeners and less evident for some others. Nonetheless –even when it was not experienced– participants were still able to complete the descriptive elements selection task.

According to the analysis of the descriptions and the elements-selection task, it is evident that MOD features visual elements from both TST and TNC. Therefore, it is safe to conclude that there is a connection between the combination of specific sonic features and the imagery that they foster. Unfortunately, we were not able to precisely connect all the sonic features that we used to their respective visual counterpart. Nevertheless, acknowledging that specific sonic features carry –more or less explicitly– their own imagery, will help future researches in the field to actually focus on cataloguing musical features of interest.

Bibliography

[1] URL: \url{http://worldpopulationreview.com/countries/western-countries/}.

[2] Birgitte Andersen and Marion Frenz. *The impact of music downloads and p2p file-sharing on the purchase of music: a study for industry in Canada*. Industry Canada, 2007.

[3] Jackie Andrade et al. "Assessing vividness of mental imagery: the plymouth sensory imagery questionnaire". In: *British Journal of Psychology* 105.4 (2014), pp. 547–563.

[4] Felicia Rodica Balteș and Andrei C Miu. "Emotions during live music performance: Links with individual differences in empathy, visual imagery, and mood." In: *Psychomusicology: Music, Mind, and Brain* 24.1 (2014), p. 58.

[5] Emmanuel Bigand and Bénédicte Poulin-Charronnat. "Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training". In: *Cognition* 100.1 (2006), pp. 100–130.

[6] Lars Ole Bonde. "Music as metaphor and analogy: A literature essay". In: *Nordic Journal of Music Therapy* 16.1 (2007), pp. 73–81.

[7] Helen L Bonny and Louis M Savary. *Music and your mind: Listening with a new consciousness.* Harper & Row, 1973.

[8] Claudia Bullerjahn and Markus Güldenring. "An empirical investigation of effects of film music using qualitative content analysis." In: *Psychomusicology: A Journal of Research in Music Cognition* 13.1-2 (1994), p. 99.

[9] David Byrne. *How music works*. Three Rivers Press, 2017.

[10] Annabel J Cohen. "Music as a source of emotion in film". In: *Music and emotion: Theory and research* (2001), pp. 249–272.

[11] Sally Jo Cunningham, David Bainbridge, and Dana McKay. "Finding new music: a diary study of everyday encounter with novel songs". In: (2007).

[12] Mark H Davis. "Measuring individual differences in empathy: Evidence for a multidimensional approach." In: *Journal of personality and social psychology* 44.1 (1983), p. 113.

[13] Peter Gregory Dunn, Boris de Ruyter, and Don G Bouwhuis. "Toward a better understanding of the relation between music preference, listening behavior, and personality". In: *Psychology of Music* 40.4 (2012), pp. 411–428.

[14] Isabelle Frohne-Hagemann et al. *Guided imagery & music (GIM) and music imagery methods for individual and group therapy*. Jessica Kingsley Publishers, 2015.

[15] Susan Hallam and Vanessa Prince. "Conceptions of musical ability". In: *Research Studies in Music Education* 20.1 (2003), pp. 2–22.

[16] Patrick G Hunter, E Glenn Schellenberg, and Andrew T Griffith. "Misery loves company: Mood-congruent emotional responding to music." In: *Emotion* 11.5 (2011), p. 1068.

[17] James Jarvis. "Guided imagery and music (GIM) as a primary psychotherapeutic approach". In: *Music Therapy Perspectives* 5.1 (1988), pp. 69–72.

[18] Carl Gustav Jung. "The Collected Works of CG Jung; Vol. 9: The Archetypes And The Collective Unconscious". In: *Trans. RFC Hull. Princeton, NJ: Princeton University Press* (1969).

[19] Henrik Jungaberle, Rolf Verres, and Fletcher DuBois. "New steps in musical meaning–The metaphoric process as an organizing principle". In: *Nordic Journal of Music Therapy* 10.1 (2001), pp. 4–16.

[20] Patrik N Juslin and Daniel Västfjäll. "Emotional responses to music: The need to consider underlying mechanisms". In: *Behavioral and brain sciences* 31.5 (2008), pp. 559–575.

[21] Mohsen Kamalzadeh, Dominikus Baur, and Torsten Möller. "A survey on music listening and management behaviours". In: (2012).

[22] James M. Keller. URL: \url{http://www.nyphil.org/~/media/pdfs/program-notes/1718/R-Strauss-Don-Quixote.ashx}.

[23] Stephen M Kosslyn, Giorgio Ganis, and William L Thompson. "Neural foundations of imagery". In: *Nature reviews neuroscience* 2.9 (2001), p. 635.

[24] Stephen Michael Kosslyn. *Image and mind*. Harvard University Press, 1980.

[25] Imre Lahdelma and Tuomas Eerola. "Single chords convey distinct emotional qualities to both naive and expert listeners". In: *Psychology of Music* 44.1 (2016), pp. 37–54.

[26] D. J. Leiner. Version Program Version 3.1.06-i. 2019. URL: \url{http://www.soscisurvey.com}.

[27] Andrew Leyshon. "The Software Slump?: digital music, the democratisation of technology, and the decline of the recording studio sector within the musical economy". In: *Environment and Planning A* 41.6 (2009), pp. 1309–1331.

[28] David F Marks. “New directions for mental imagery research.” In: (1995).

[29] Cathy H McKinney et al. “Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults.” In: *Health psychology* 16.4 (1997), p. 390.

[30] Daniel Müllensiefen et al. “The Goldsmiths Musical Sophistication Index (Gold-MSI): Technical Report and Documentation v0. 9”. In: *London: Goldsmiths, University of London* (2011).

[31] Bence Nanay. “Multimodal mental imagery”. In: *Cortex* 105 (2018), pp. 125–134.

[32] Adrian C North, David J Hargreaves, and Susan A O’Neill. “The importance of music to adolescents”. In: *British Journal of Educational Psychology* 70.2 (2000), pp. 255–272.

[33] Joy E Ollen. “A criterion-related validity test of selected indicators of musical sophistication using expert ratings”. PhD thesis. The Ohio State University, 2006.

[34] John W Osborne. “A phenomenological investigation of the musical representation of extra-musical ideas”. In: *Journal of Phenomenological Psychology* 20.2 (1989), p. 151.

[35] John W Osborne. “The mapping of thoughts, emotions, sensations, and images as responses to music.” In: *Journal of Mental Imagery* (1981).

[36] Hansjörg Pauli. “Filmmusik: ein historisch-kritischer Abriß”. In: *Musik in den Massenmedien Rundfunk und Fernsehen. Perspektiven und Materialien* (1976), pp. 91–119.

[37] Patrik N Peretz Isabelle as comment to Juslin and Daniel Västfjäll. “The need to consider underlying mechanisms: A response from dissonance, answer to Emotional responses to music: The need to consider underlying mechanisms”. In: *Behavioral and brain sciences* 31.5 (2008), pp. 559–575.

[38] Alexandra Quittner. “on Visual Imagery: A Multiple Measures Approach”. In: *Journal of Mental Imagery* 7.1 (1983), pp. 105–120.

[39] Ulf-Dietrich Reips. “Standards for Internet-based experimenting.” In: *Experimental psychology* 49.4 (2002), p. 243.

[40] Géza Révész. *Introduction to the Psychology of Music*. Longmans, Green Company, 1953.

[41] Alan Richardson. *Mental imagery*. Springer, 2013.

[42] Mark Tarrant, Adrian C North, and David J Hargreaves. “English and American adolescents’ reasons for listening to music”. In: *Psychology of Music* 28.2 (2000), pp. 166–173.

[43] Liila Taruffi and Mats B Küssner. "A Review of Music-Evoked Visual Mental Imagery: Conceptual Issues, Relation to Emotion, and Functional Outcome". In: ().

[44] Various. URL: \url{https://en.wikipedia.org/wiki/Program_music}.

[45] Jonna K Vuoskoski and Tuomas Eerola. "Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories." In: *Psychology of Aesthetics, Creativity, and the Arts* 6.3 (2012), p. 204.

[46] Richard Warr and Mark MH Goode. "Is the music industry stuck between rock and a hard place? The role of the Internet and three possible scenarios". In: *Journal of Retailing and Consumer Services* 18.2 (2011), pp. 126–131.

[47] David Watson, Lee Anna Clark, and Auke Tellegen. "Development and validation of brief measures of positive and negative affect: the PANAS scales." In: *Journal of personality and social psychology* 54.6 (1988), p. 1063.

[48] Clemens Wöllner. "Is empathy related to the perception of emotional expression in music? A multimodal time-series analysis." In: *Psychology of Aesthetics, Creativity, and the Arts* 6.3 (2012), p. 214.

[49] Yi-Hsuan Yang and Jen-Yu Liu. "Quantitative study of music listening behavior in a social and affective context". In: *IEEE Transactions on Multimedia* 15.6 (2013), pp. 1304–1315.

[50] Marcel Zentner, Didier Grandjean, and Klaus R Scherer. "Emotions evoked by the sound of music: characterization, classification, and measurement." In: *Emotion* 8.4 (2008), p. 494.

Appendix A

Pilot Survey

Here is the structure of the pilot survey that was created and conducted between December 2017 and January 2018.

Welcome!

I am a Master's student at Aalborg University doing a project for a course in Music Perception and I would be very grateful if you could help me by taking part to this survey. Please ensure that your browser has Flash installed and activated. You will be listening to 2 minute of music and then follows a few questions. The whole test should take no more than 5 minutes.
Adjust the volume of your headphones to a comfortable level by playing the test audio. When it is finished, please click "next" to start the questionnaire.
[Piano Chord Audio File]

Song

Please carefully listen to this two minute song
[Song Audio File]
Here follows a few questions about your listening experience. Your responses are anonymous and there are no wrong answers.

The Song Was this piece song familiar to you?
[Yes/No]

Mental Imagery When listening to the music, did any images appear or did the music make you visualize anything?

[Yes/No]

What did you see? Please describe as best and detailed as you can what you did see. Answer as truthfully as you can. There are no right or wrong answers.
[text box]

What words would you use to best describe a scene or setting that could fit the music you heard? Double click or drag and drop the elements below. Select the only settings, adjectives, colors that you see fit and put them in the order you prefer. You do not have to use all the elements or to fill every slot.
[descriptive elements selection]

Were there any particular sounds/instrument in the song that made you link to the words you picked? Why? If you forgot the song you can listen to it again.
[Text box]

Please briefly describe your mood and how you feel after listening to the song.
[text box]

Goodbye!

Thank you for completing this questionnaire!
We would like to thank you very much for helping us.
Your answers were transmitted, you may close the browser window or tab now.
Have a very nice day :)
M.Sc. Sound and Music Computing, Aalborg University, Copenhagen

Appendix B

Survey Setup

The survey was structured as follows.

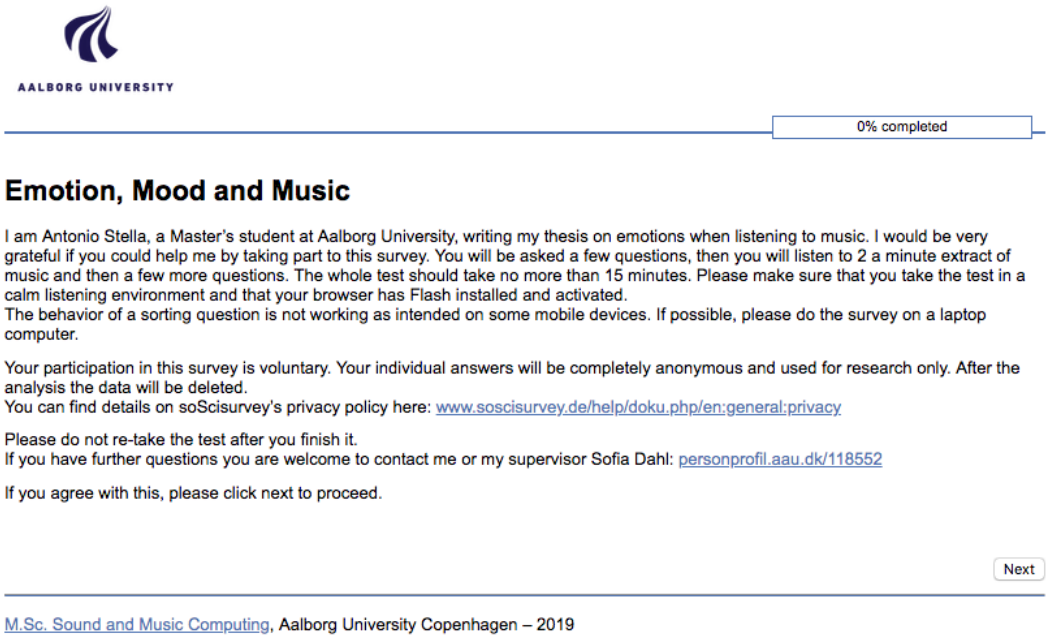

Figure B.1: Introduction

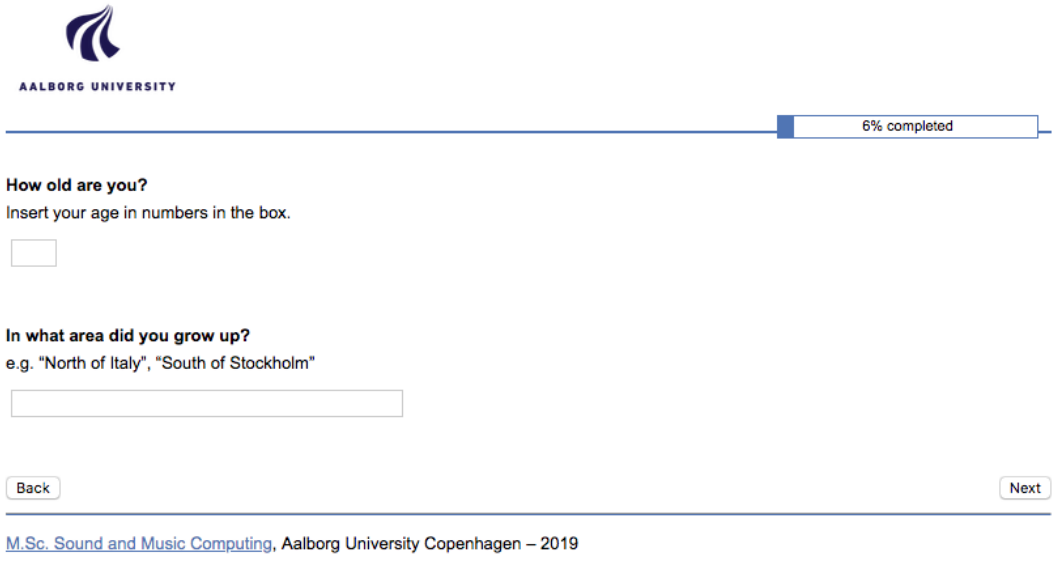

Figure B.2: Age and Origin

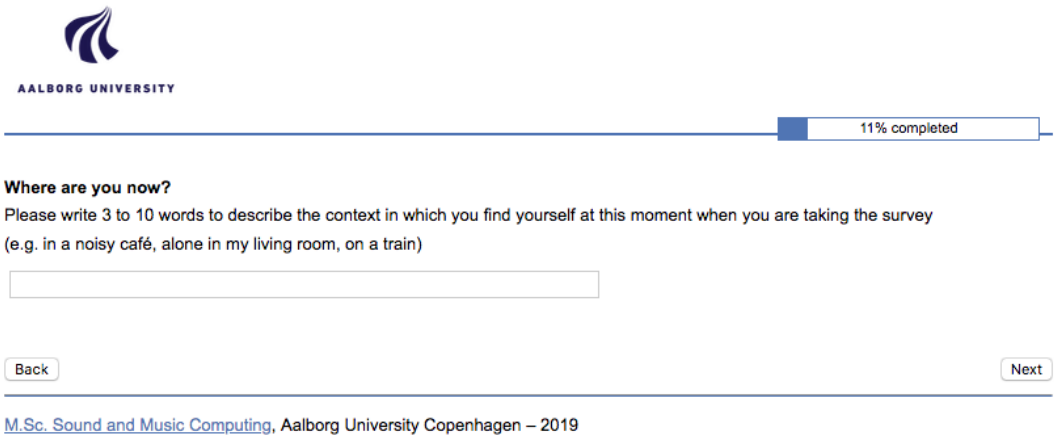

Figure B.3: Context of the Candidate

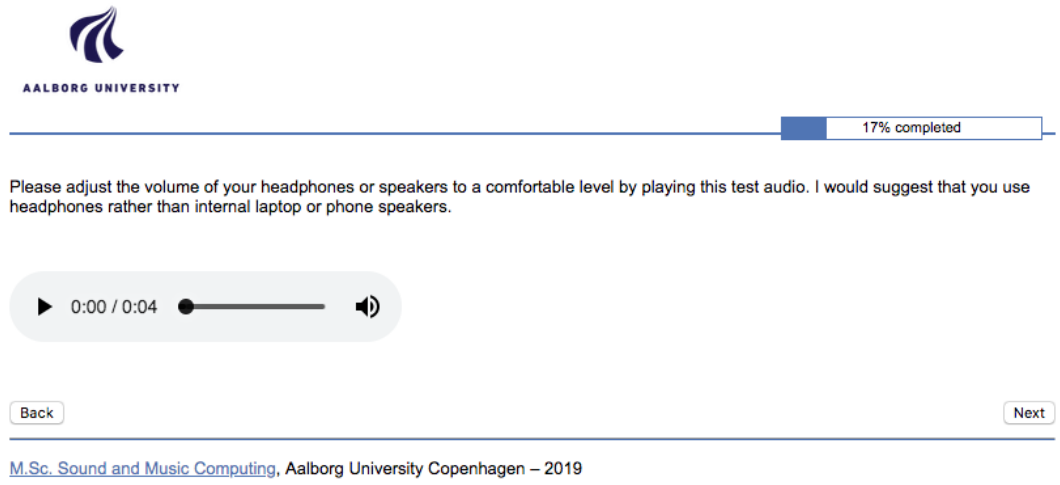

Figure B.4: Adjust Volume of your device

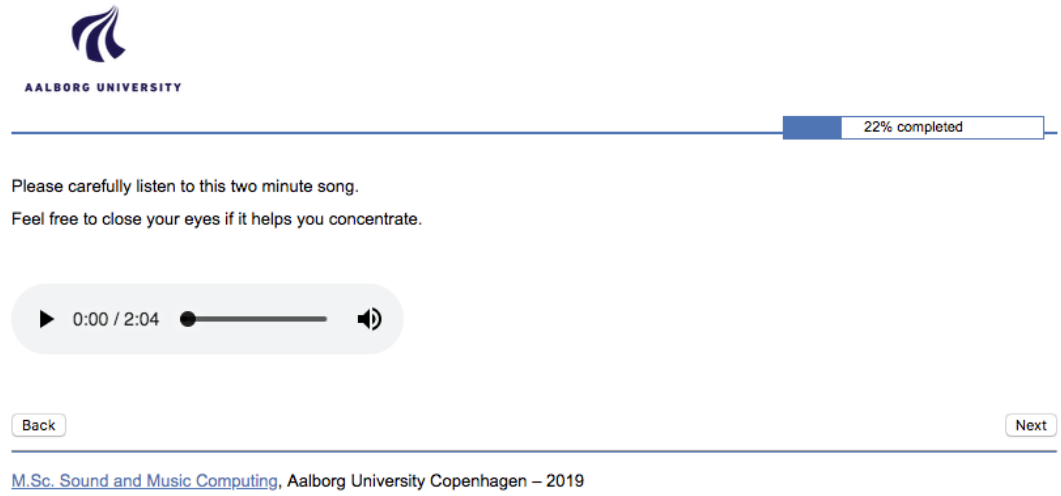

Figure B.5: Listening Test

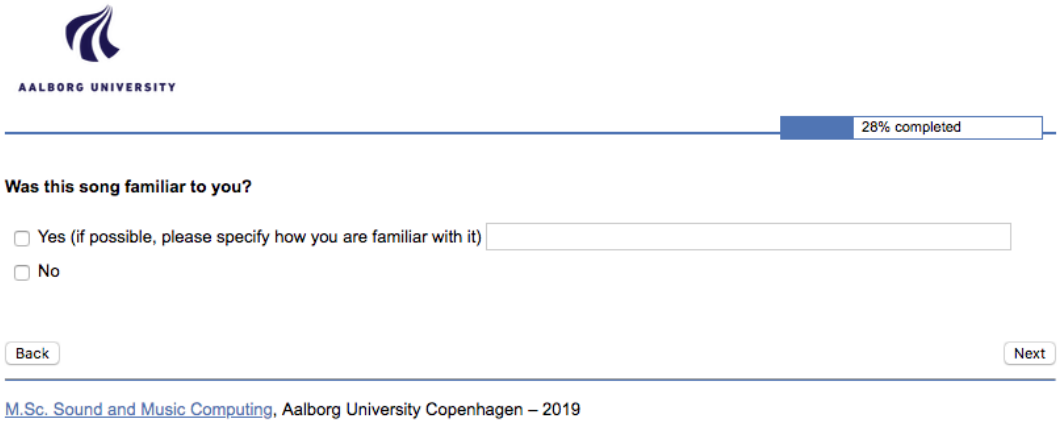

Figure B.6: Familiarity with the song

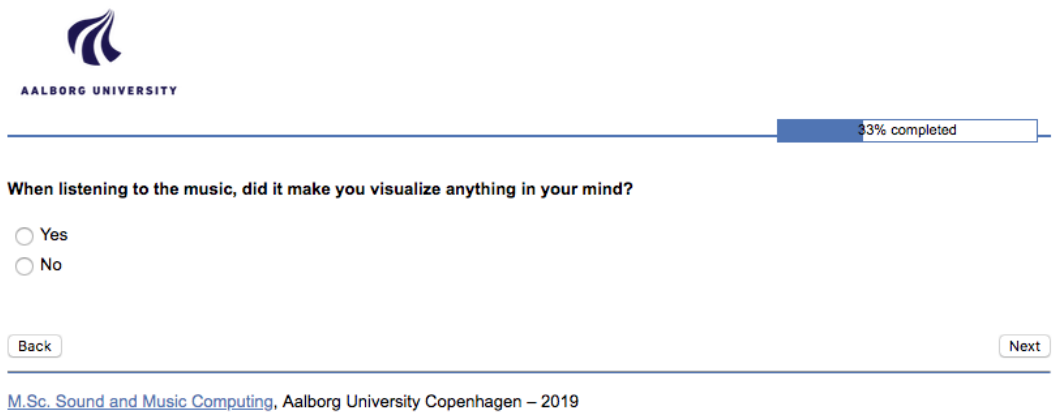

Figure B.7: Visual Mental Imagery

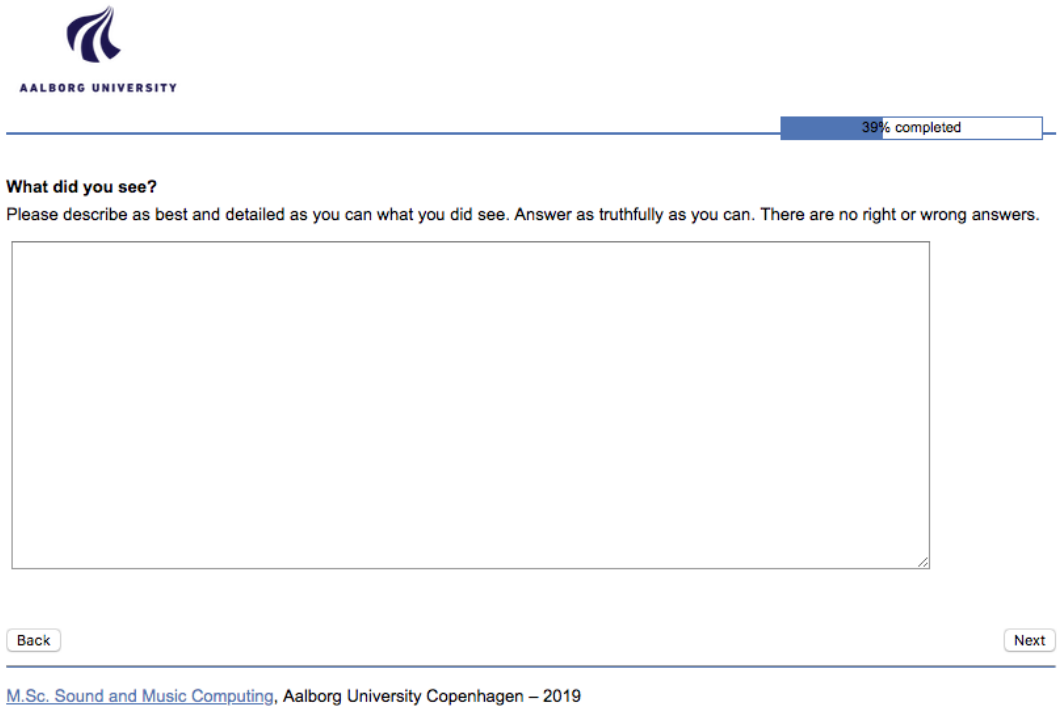

Figure B.8: Conditional Question if participants answered Yes to the question before

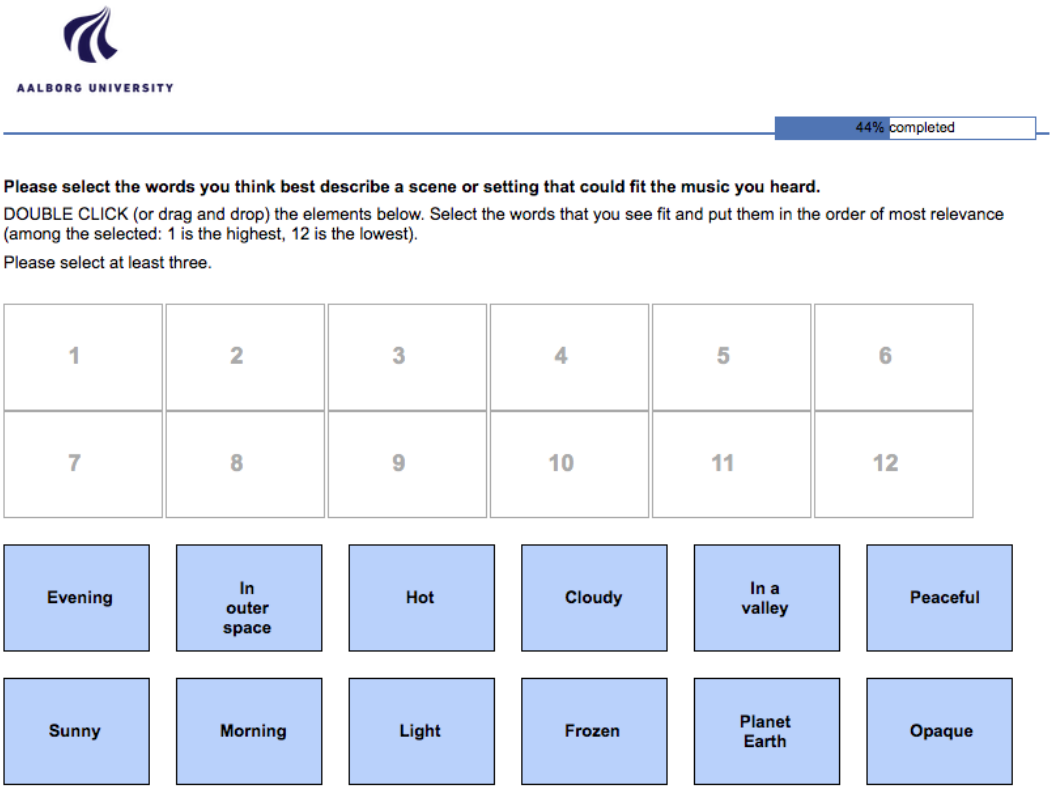

Figure B.9: Descriptive Elements Selection 1

Figure B.10: Descriptive Elements Selection 2

Figure B.11: Pilot Test

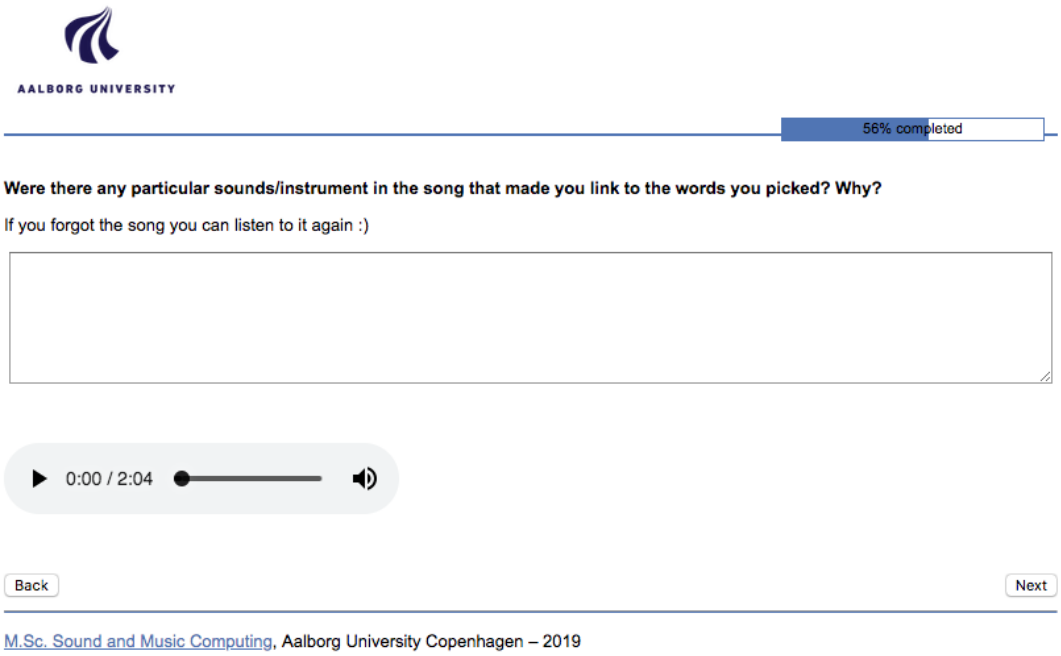

Figure B.12: Sonic Features

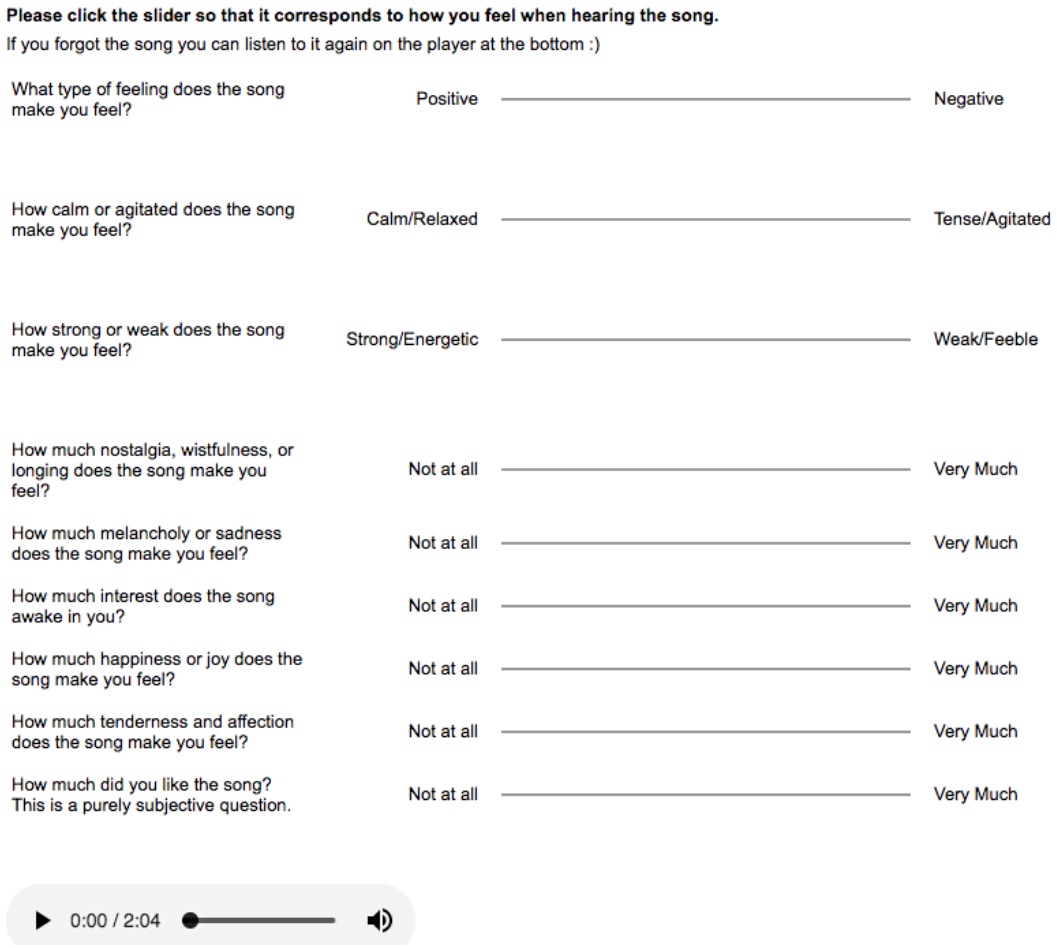

Figure B.13: Emotional State

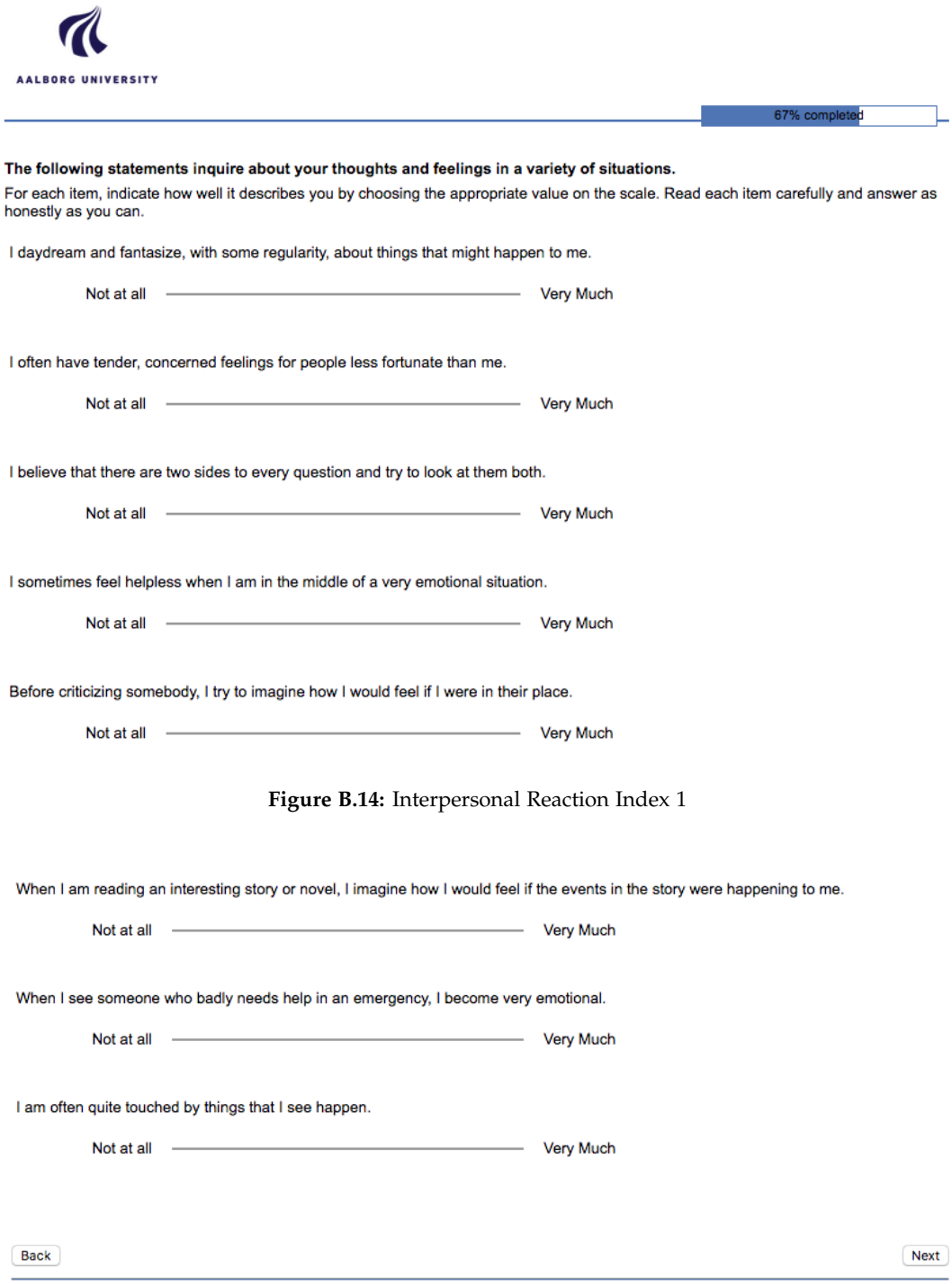

Figure B.14: Interpersonal Reaction Index 1

Figure B.15: Interpersonal Reaction Index 2

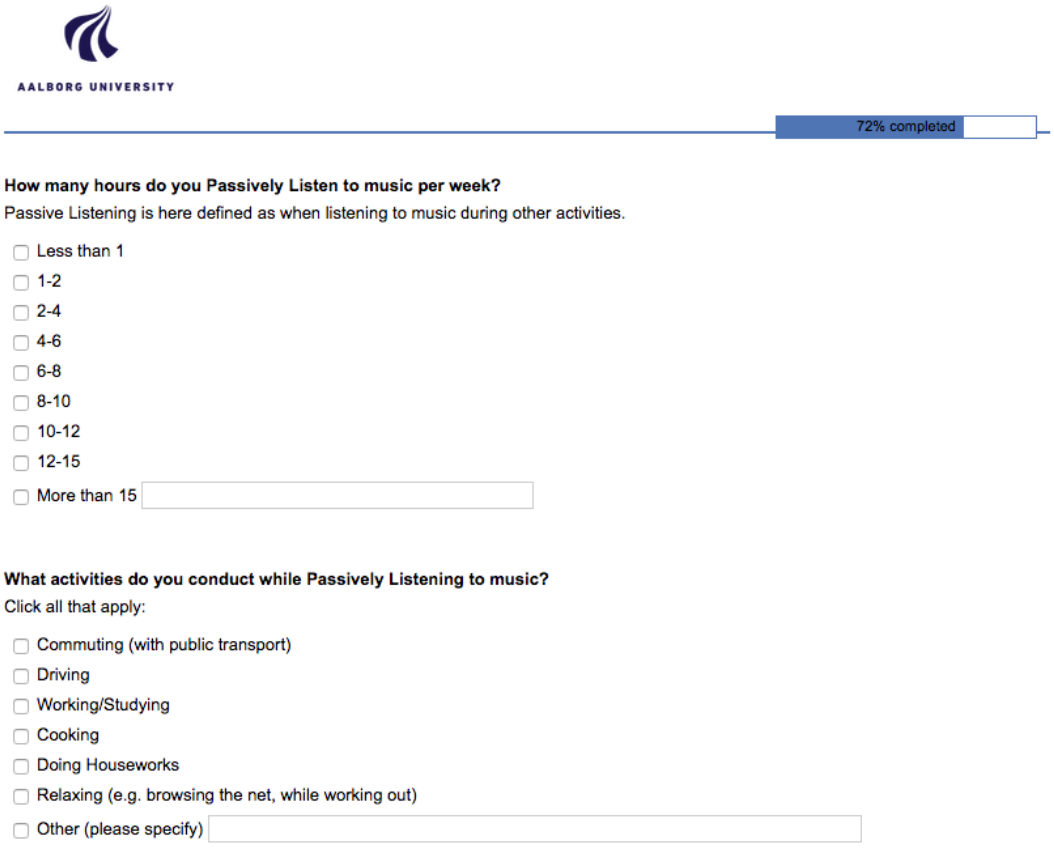

Figure B.16: Musical Behaviour 1

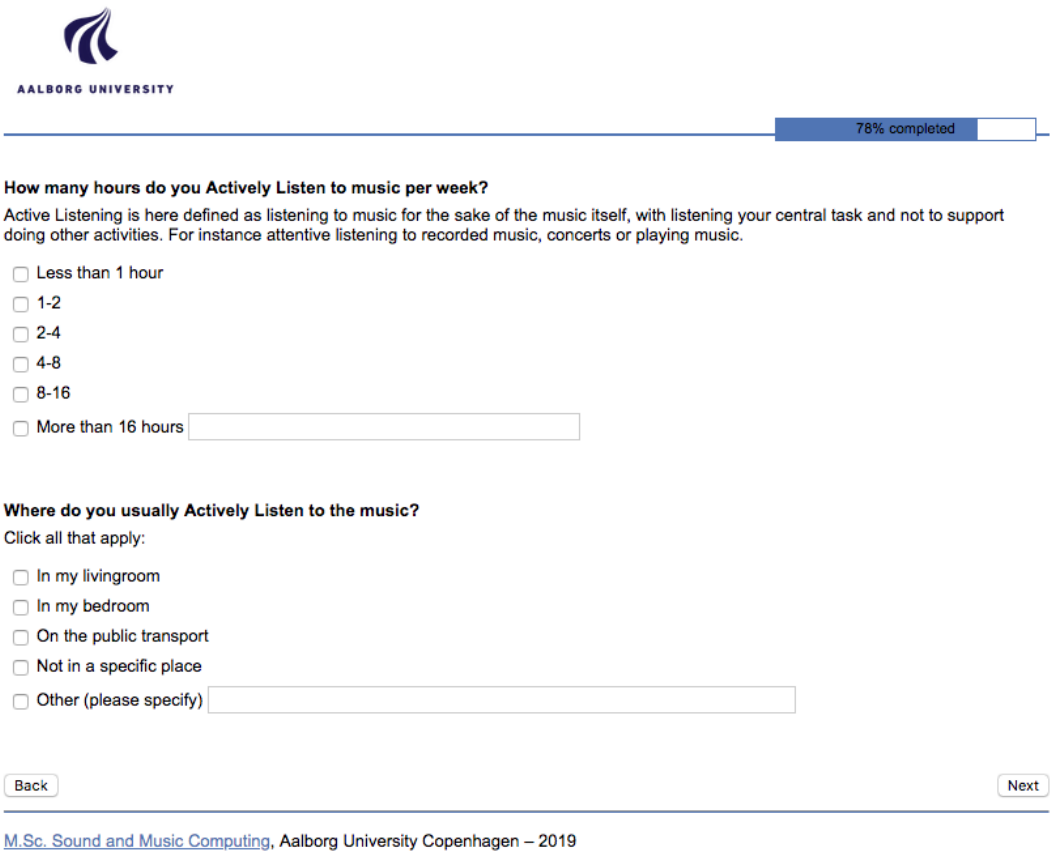

Figure B.17: Musical Behaviour 2

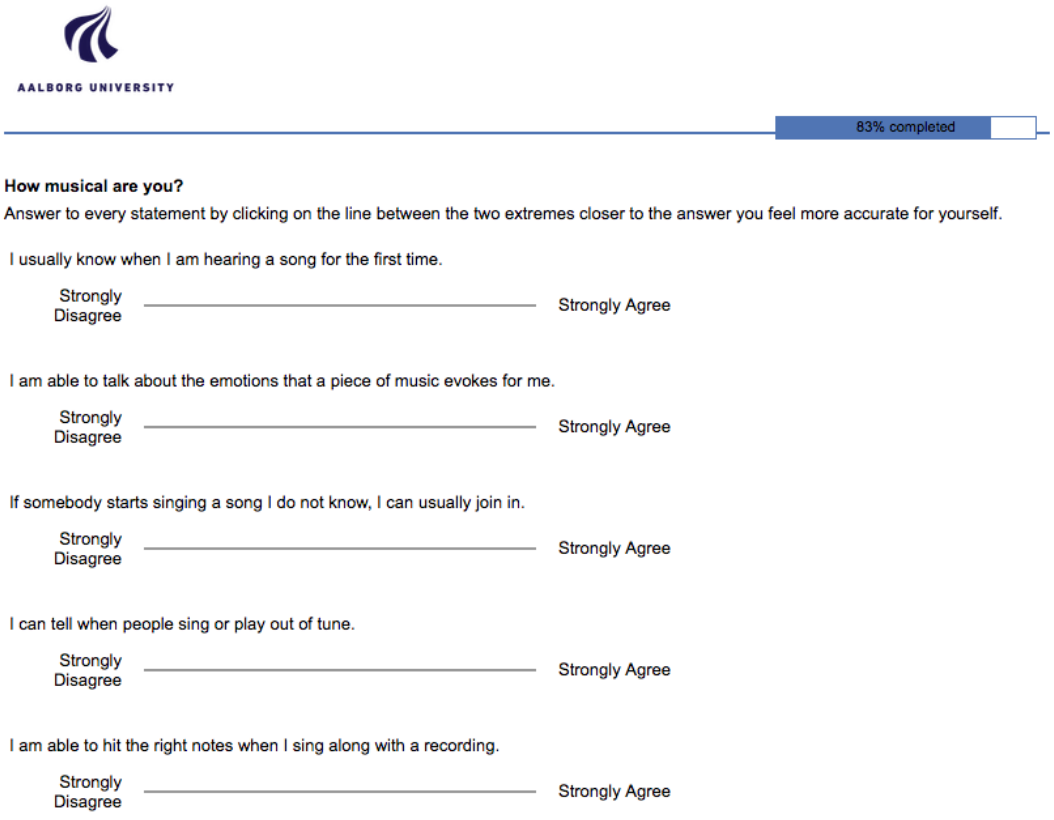

Figure B.18: Music Sophistication 1

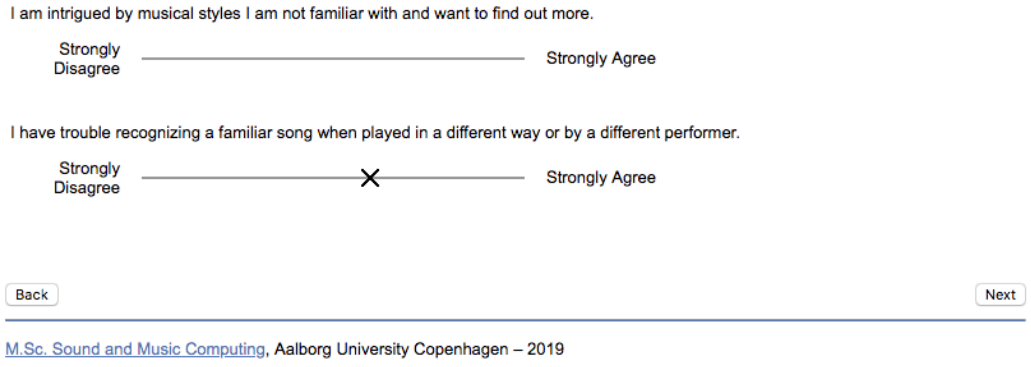

Figure B.19: Music Sophistication 2

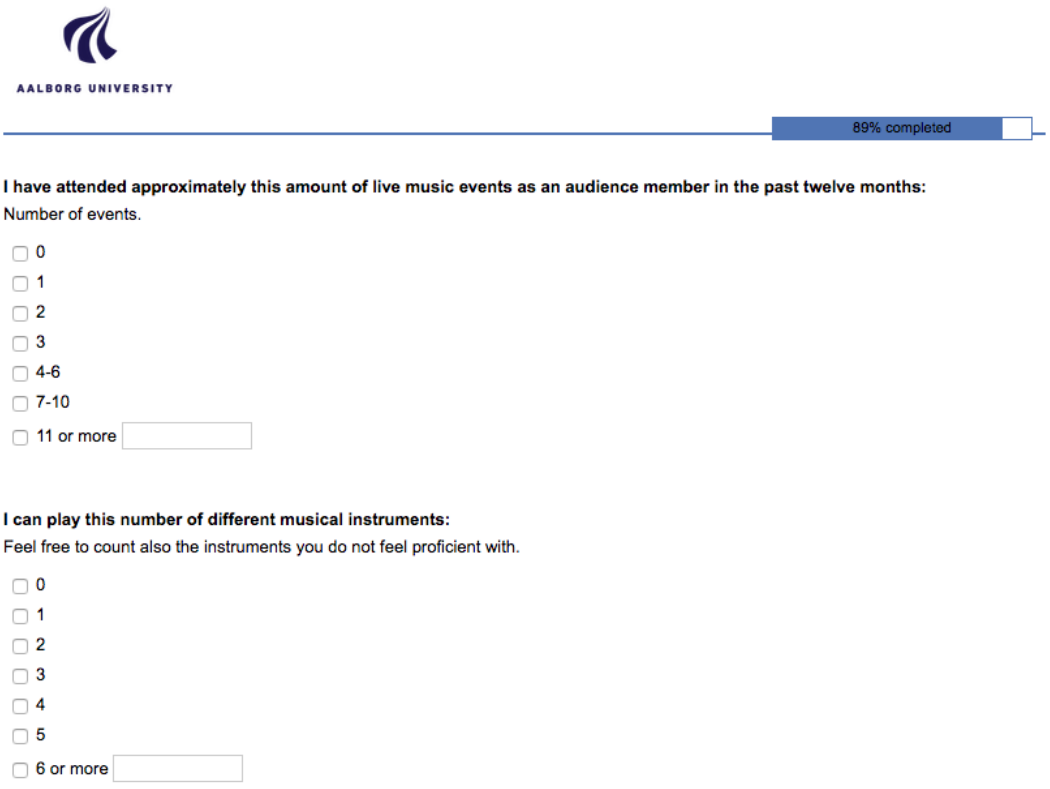

Figure B.20: Music Sophistication 3

Figure B.21: Music Sophistication 4

Figure B.22: Conclusion

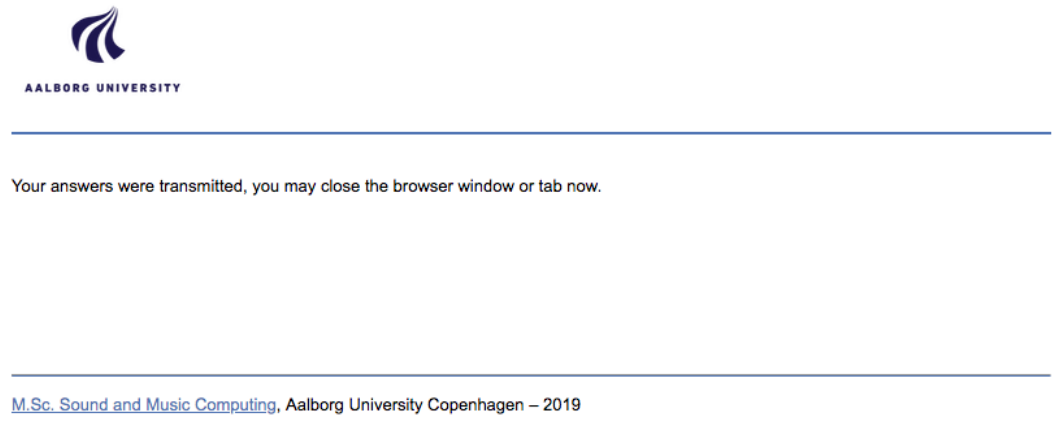

Figure B.23: Final Slide

Appendix C

Gear Used

Electric Guitar: 2007 Fender Jazzmaster American Vintage Reissue '62.

Electric Bass Guitar: 2015 Hamer Slammer.

Pedal Chain:

- tuner: Poly Tune 2 - TC Electronics
- octaver/envelope: Polyphonic Octave Generator 2 - Electro Harmonix (POG2)
- distortion: ProCo Rat 2
- delay/reverb: Avalanche Run V.2 - Earthquaker Devices
- reverb: Reverberation Machine - Death by Audio

Audio Interface: iConnectivity iConnectAUDIO 4+.

The rest of the instruments used were virtual machines native in Logic Pro X except for the synthesizers on TST that were native in Ableton Live.

Appendix D

Imagery Descriptions

Here you can find the imageries described by all the candidates together with the count of the used words. The question with the relative instructions were:

> "What did you see?
> Please describe as best and detailed as you can what you did see. Answer as truthfully as you can. There are no right or wrong answers."

D.1 Deep Blue Day

D.1.1 Descriptions

"I could see a place in the countryside, with a lot of nature around, like trees, crops, barley/wheat and grass. Mainly tons of orange, yelllow and green. This song, even though I do not remember it specifically, reminds me of some movies that are also carried out in the country side, with horses and cowboys. I could also see someone playing a guitar or a banjo."

"waves radiating out of a center light. blues, greens"

"California, late-ish evening. Hills and trees. A little bit of a breeze, enough to feel but not enough to move the trees much"

"I didn't close my eyes but so instead the feeling of the music was what gave me thoughts about things It felt like summer, like a breeze, being next to water, maybe in a warmer country and as the music shifted it felt like it was more of an evening with stars in the sky and live music playing"

"Calm waves, sunny beach by the sea"

"I see sea waves at sunset"

"Waves or circles expanding and contracting for the shimmery background sound that seemed to change in spectral energy"

"I am slowly approaching a river in a middle of forest with tall trees. There is so much peace in and outside me while I am getting close to a river. After some time I decided to move back deep in a forest away from river."

"A sea shore at sunset, with a warm breeze and palm trees. The sun was glistening gold on the water"

"I kind of "saw" a walk through a green meadow along the edge of a forest. Very GREEN :-)"

"The initial orchestration was light and airy and made me think of Spring of Summer. The bass line created a slow, relaxed feeling and as the melody started I immediately started to visualize a tropical beach scene. Definitely a lazy summer day on the beach vibe. The melody seemed to evolve a bit over time and got more of a country and western twang, but it did not disrupt the summer feeling."

"A grand entrance scene by a beautiful young woman (fairy?). The start builds suspense, quickly builds into a high near-crescendo (when i can see the camera revealing her face) and then the subsequent rise and fall of the tempo and volume leads me to visualise her walking across a large frozen lake. Don't know why I visualised a frozen lake - could have been a verdant green landscape - but I visualised a frozen white lake"

"Beach, relaxing people, drinking pina colada"

"Some of the images that crossed my mind:-A morning at the beach with a light drizzle -Happy humans -Grassy hills"

"Something wavy, like waves in the ocean or a wheat field in a windy day."

"I started of at a beach at sunrise, watching the waves from a distance as the dark sky turned orange as the sun rose from the horizon. In the 2nd half of the piece, I was, unexpectedly, transported to the mountains. I sat with a cup of tea and watched the snow-clad mountains at a distance. A cool, gentle breeze was blowing across my face. It felt like it was mid-afternoon. The overwhelming emotion was one of serenity."

"Waves"

"I can see a wide And open sky landscape with some jolly innocent white clouds. The colours are blue of the clear sky And white clouds. Some sounds evoke also light waterfalls or small fountains."

"The North America Countryside"

"A busy urban street, walking along down it with lots of lights and cars rushing and excitement, but with internal cheerful peace."

"there is no gravity. I can float. everything is floating in a dark and empty space. sometimes there is a light, a flash, like a star. everything is chill"

"a chilled environment with somewhere a guitar playing person appearing with a bottleneck or so, reminded me a bit of Hawaiian music"

"a restaurant and a soothing music at the background. some of the notes of the music were irritating though"

"An orchestra against a vast background of open field. Brown grass, a sky with white clouds."

"I immediately pictured a high school prom dance, based on the style and instrumentation. The sound was muffled at the beginning, as if the ""camera"" was outside the dance, but then we gradually get closer and zoom in on two teens dancing to the slow number. The lights are bluish and we see flecks of disco-ball color. The boy's hair is gelled and styled carefully and the girl has perfectly-set curls that took her hours to get right. She has sequins on her dress. The ambient, distorted sound on top of the more acoustic feel made me think that we are seeing this as a memory, or that something bad is about to happen... will zombies or aliens break into the dance, and the camera follow them in slow motion as they crash the party? Or, maybe this is a memory resurfacing years later, when the girl is an old woman with Alzheimer's and has a rare lucid moment listening upon hearing ""their song."" Anyway, there is something bizarre or dark that taints the otherwise mellow, nostalgic feel."

"One of the instruments in the song reminded me of Hawaii and being a part of a luau."

"Usually I imagine a person that run. The landscape all around depends on the song"

"Flowery shapes of various colours that grew from the centre and then vanished outside of my life of sight

"The mixure of sounds made me a little anxious, the contrast between the loong quiet sound and the fast rythm wasnt nice for me it were too opposite sounds. but when i focused at the fast rythm i could imagine someone jumping at a field of grass at the mountains, but still it wasnt nice, i think mostly because of the mixture of sounds."

"Unfolding evolving helix."

D.1.2 Words Count

waves(7), water(2), waterfall
beach(5), shore
grass(4), meadow
sky(4), clouds(2), stars(2)
breeze(4), wind, light drizzle
green(4), verdant, brown, blue(2), dark(2), orange(2), various colours, sea(3), ocean, river
warm(2)
trees(2), forest(2), palm trees, yellow, glistening gold, shimmery
sunset(2), sunny, sunrise, light
guitar(2), banjo, western twang, live music
summer(2)
countryside(2), country, barley/wheat crops, wheat field
relaxed(2), chilled(2), peace(2), calm, serenity, cheerful
evening(2)
Hawaiian(2), tropical
frozen white, snow
mountains(2), hills(2)

D.1.3 Sonic Features

"guitar - fireplace it made the song sound as if there was a bit of dancing going on or maybe just chilling by a fire and listening the sound throughout the song that

was more crisp felt like a summer breeze from an ocean, wether you were heading for it or being next to it"

"Guitar and something that reminds me of a banjo. Also the peaceful and repetitive melody."

"No, it was a combination of everything"

"Synth- outer space, floating, dark, in the sky. Guitar- peaceful, city, indoors"

"Yes there was one part in the song when one of the instruments went from an F to a G that gave it a distinct Hawaiian sound."

"the vibrato/resonance/harmony"

"Yes, because It's peaceful but nostalgic at the same time "

"The spacey synth sound, and the washed out piano, along with the light guitar strums. The choice of scale and chords also played some part in this."

"Piano"

"The heavily echoed sounds sounded shiny, the guitar was peaceful and the bass warm."

"Orchestration - rising, major(?) key, gave a sense of lightness, space, and warmth"

"Bass line - slow, gentle, relaxing. Guitar/Strings - memory of Polynesian music/tropical feeling. Chord progression - familiarity and comfort."

"dont know ...cant guess instruments that easily"

"I am not good at detecting which sound comes from which instruemnt. if my guess is correct it were the piano cords and the trumpet, coupled with the tak tak probbly on a congo."

"The sounds in the background particularly, but also the guitar"

"It sounded a bit like Hawaiian music, a little bit like that slide-guitar sound, but not entirely. There were a lot of overtones and the 'bass' seemed also quite high and floating rather than driving. Therefore, it made me feel like I was on a beach,

or rather like a collective-notion of what a calm beach would be like. My skin is too white to be in the sun for long, but I still have some sort of socially-embedded sense of 'beach-ness' which is more like the music sounds."

"the synth is so smooth, this makes you feel like floating"

"Hawai guitar, synth strings"

"synthesizer"

"cloudy hovering pad/synths: linked to floating, chilled environment, calm setting, no worries. rhythm guitar, melody and easy 1-5 bass: reminding me of (e.g. hawaiian) traditional music. later guitar (plus bottleneck) reminding me of both western (american) music or hawaiian"

"the background rythm remember me to outher space. the fast rythm the one with shorter intervals made me image someonse jumping, a valley, mountains and a sunny afternoon so yellow."

"Slide guitar made me think of the desert. The regular pulse of the rhythm guitars and bass reminded me of the train. The high pitched synth sounds were peaceful."

"The damp quality of the sound of the rhythm/drums. The repetition in the melody too."

"Evolving pad sound and keyboard keys which played the melody."

"The background sound is smooth"

"The synthetic sounds made me think of an urban scene."

"The initial beat reminded me of footsteps (hence a woman walking). The synthesiser(?) rising in the beginning led to suspense. There is a piano/ keyboard that I ca hear (it has discrete key-strokes, unlike what I am calling the synthesiser) lends serenity and peace (not scary)"

"All of it together, esp. the interplay of the bass, keyboard, and guitar"

"Perhaps the slow beat and the string section"

"walk = the rhythm of the song"

"Synth pads - added to the relaxed/calm feeling - floating. Bass - grounded, calmness. Strummed guitar - warm, cosy. Slide guitar - made me think of being on a beach in the summer"

"background sound that runs with different sound intensity reminded me of waves"

"steel guitar"

D.2 Test Song

D.2.1 Descriptions

"Nature in summer, lots of green leaves"

"I see the space, the stars. I'm flying, I'm free, I'm relaxed. The stars are shining on the rhythm of the music."

"I see me alone on the beach watching the ocean."

"Floating, flying, looking down on green fields, dreaming, clouds, relaxing, loving, laughing, feeling, tender, yes."

"I didn't really 'see' a specific scene. The music evoked these feelings in me and I had a sense that I was possibly flying, looking down on a landscape but nothing more detailed than that."

"Gold colours, swirling lights"

"At the beginning I was picturing vast landscapes with mountains in the background, but the scenery was changing fast. Afterwards I was picturing seeing earth from above and it turning fast. After a minute or so there were no definite pictures in my head anymore, and all I saw were the sounds itself in the form of soft blue and white textures that were moving and morphing slowly. All behind a black background. It looked a bit like blurry and overexposed pictures of no definite form. "

"I thought about the seabed, very colorful, bright and with magical creatures

"During this song I saw a field of flowers in spring. The sun was shining and the flowers were moving in the wind. The colours were very intense and there was a positive atmosphere."

"Deserted and broken street at Christmas, Untidy room that's left, Something not quite happy just happened, but everyone accepted the situation, Broken tape record"

"Blue/gray smoke diffusing in the dark, over an oriental city"

"I would associate this music with an impression of floating in space. I saw a kind of night sky with stars, a bit like a ""galaxy-type"" visual pattern. The stars would appear and disappear and they would be of different colours with different level of focus, some would be blurry while others would be sharp. The general feeling would be quite positive."

"I saw a new age band performing with silk scarves and dancing in a slow "interpretative dance" type of style. The main performer had long hair, a scruffy beard, and was shirtless."

"I see a cloud in front of my eyes, like being in front of a rain cloud, some pink neon flashes and everything happens in a black space."

"Initially felt that it was an opening for tv news, it had that anticipatory feel that something important was coming. Then after the first minute it felt more subtle and relaxing."

"I closed my eyes, felt that I am floating on cool blue water-waves, filled with serene calmness. "Quietly proceeding ahead, I am all alone, felt, music should not stop at all . Soothing,my soul ."

"Forest or natural landscape, lake or river in distance, people discovering the area. People marching to a slow beat. Seemed to be some sort of funeral march. And it was rainy. But people were not sad. It was some sort of gathering of people"

"A calm sea landscape, with a colorful sky and big black clouds. Then the image started dissolving to black and stars began to appear. Near the end, I pictured some of the instruments: the metallic sound (metallophone?) and the electric bass. The slide-guitar melody made me think of a summer/beach like scene, while the blurry effects/filters on most of the instruments added a dreamy atmosphere to the whole image."

"the ocean"

"Cafe Del Mar, 2 hours before sunset."

"Tropical island scene, near a shore, oranges of the evening"

"A forest. Lovely fresh green trees. Colourful birds on them"

"Walking around a sunny spring day in Stockholm."

D.2.2 Words Count

nature(2)
summer(2), tropical
green(3), blue(3), grey, orange, black
space(3), stars(3), flying(2)
floating(3)
relaxed(3), calm(2), serene, positive, not sad
beach(2), shore, field(2)
clouds(3), smoke, rainy
ocean(2), sea(2), seabed, waves, lake, rivers
sun, sunny, sunset
leaves, forest

D.2.3 Sonic Features

"drums"

"Melody line"

"The guitars and the effects"

"muffled sounds for the bad weather impressions"

"the guitar, the glockenspiel, the organ"

"The well-reverbed bed towards the back - the organs. Also the main synth line and the fact the drums were eq'd so they were muffled"

"Sound of water"

"The xylophone, but I can't really describe why."

"what were the words again? ;-) Hammond Sound -> yellow, foggy, afternoon, night. Snare -> cold, floating. Distortions -> dense, foggy. Bass -> floating"

"not the drums (everything else)"

"Ringing reverb-y bright sound as well as the pentatonic melody line"

"The guitar with long reverberation. The piano with long pedal. Vibraphone (if there is one there ?). I think I associated these sounds to floating in space because of the different space movies I have seen."

"The atmospheric keyboard reminded me of planet documentary soundtracks. The higher instrument with long notes is very floaty/space. The beat was more road-trip/car in the rain/storm."

"I really disliked the bass/drum-like sound. Off putting. Not the particular sound but the lack of direction"

"I'm not sure what the instrument was but it was "twangy" and reminded me of Middle Eastern music, perhaps Turkish. This song reminded me of the songs played by the band in the movie "Serendipity""

"The blurry background, the reverb, the generally echoey setup."

"The repetitive synth motif"

"not a specific instrument, but rather bright and harmonious sounds with soft attacks and long releases."

"upheaval, echoing, gradually ending. soft Drumbeats- enjoyable and instrument's rhythm was superb!"

"Airy keyboards / mallets / whatever"

"Drum beat suggested the march, the ritual. the rest, I can't say"

"The fact the the guitar and the drums sometime sounded "far away", and the sustained synths made me think of space or clouds. The metallic sounds reminded me of both the beach and the stars. The quiet bassline and the soft end are peaceful."

"Nothing specific,"

"It's ambient, with a slow, almost lazy guitar melody. But it has gentle constant beats and a driving bass progression. It sounds exactly like the kind of thing that gets played at Cafe Del Mar. But it's earlier in the day than sunset: There's too much brightness in the pitches. So, afternoon."

"Drums -> train"

"The bells made me think of Budism therefor I imagines an outer space in the mountain, earthy. And I thought of the sunlight because of the higher pitch"

"Probably the steel pan-like sound (tremolo steel pan-like sound). The harmonic progression. The melody. The sustained pad sounds. The pitched percussion. The moderate-tempo beat. The slide guitar."

"Flute"

"Drum"

D.3 Tanca

D.3.1 Descriptions

"For the first thirty seconds or so, nothing. Then, when some sounds resembling the clanking of chains began to sound, I saw an image of chained slaves being made to constantly do something - like rowing a massive ship, or feeding coal into a furnace. At first it seemed like this sound was only coming from one direction, and then another set of such sounds kicked in from the opposite direction and that was when this scene took a solid form in my mind. I also pictured a few armored soldiers monitoring these workers."

"A group of heavy set Mongolian men at a camp of sorts in a large field of sweeping yellow-green grass. There were also horses with shiny discs of a silver-cover

metal attached to their saddles. Their movement made the tambourine-like sound in the background of the piece of music."

"Soldiers preparing for the battle and encouraging each other. Scary scene with a lot of swords, armors and weapon."

"Initially, scanning across a desert plain with rocky cliffs in the background, as someone rode on a horse. As the noise built, it turned to a battle scene, with swords clashing and canons/gunshots. Latterly it morphed into a picture of a prison with cell bars being rattled and prisoners shouting riotously."

"At first it made me visualise some kind of tribe, which was then getting ready for combat of some sort, going through the jungle and sharpening their weapons."

"Darkness, clouds, fog moving quickly. thousand slaves pulling something heavy. Guards whipping"

"While i was listening to this particular music i visualized a battlefield. I felt like i was watching vikings warrior fight for something or that i was watching warrior of a looser population being executed. I know this is particularly cruel but i associated this music to a particularly scene that i saw on a TV show (called vikings)"

"I could visualise Roman warriors like in the movie Gladiator and slaves in olden times. Some sense of rhythm because of the hard work of the slaves, but dragging nevertheless. Someone belting out orders to the slaves."

"I imagined a riot, streets at night filled with angry people (apocalypse style)."

"I saw a football stadium with herds of supporters/hooligans slowly but quite chaotically walking around. It is an apocalyptic scene and some bassi profundi and a band are stimulating them by creating a trance-like atmosphere with ominous pedal notes and a constant but driving rhythm. The scene is dark and threatening."

"Warriors marching through fog, preparations for battle"

"At first I saw some warriors preparing for battle (somehow the song reminded me or Maori people). After that I also saw them fighting and dancing in a camp."

"I saw in my mind a battle between monsters in the space"

"A battle in a medieval village"

"I imagined a battle"

"I thought of a brawl in a pub, with glass bottles breaking on the floor."

"it felt as if in a war or in anxiety. It felt as if people were fighting for water or something."

"A tribe, performing in the darkness, while fires are burning a war-like ceremony/ritual. Some performers play instruments, most of them dance. Performers are male."

"Like a ritual in a forest, during the night, with fire and lots of energy!"

"I pictured the singers and drums. Also a tribal scene, a fire, people standing in a circle. It felt like a soundtrack to a film depicting a tribal scene."

"A clearing with aboriginal people celebrating around a fire."

"It seemed like a sadist tribal dance ritual in centre of the jungle before executing someone. And a a guy in background making sounds orally was disturbing. It made me want to stop."

"It started out with throat singing which automatically makes me think of the bad ass trio of female throat singers from Mongolia and World Music gigs (with them wearing skins). Then it made me think of Jews harps (often related) and the basic drum (with throat singing) made me think of being in Mongolia and seeing the old witch doctor drums and shaman tents (I have travelled through Mongolia). Once the throat singing was modulated I just thought of synths and then chimes when there was the sound of wind chimes (maybe electronic.)"

"Disco arena"

"I imagined people around a fire. Something primordial, a ritual addressed to some gods."

"Bonfires, rituals, dance, trance, night scene. Quite primitive setting in the first part, less defined towards the end."

"It makes me think about a port in the past centuries, very crowded with lot of

people talking and moving. It was a chaotic environment predominated by pirates and sailors"

"Angry vikings in a dark wooden barn near a fire."

"Ancient horde, being afraid of something, or expecting something, or preparing to do something."

"People in caves, rain and hunger, danger, and darkness , not a happy piece"

"I visualized a throat singer and it also made me think of American Indians and their chanting during a Powwow"

"Mordor and goblins."

"I saw a big group of people from central Asia playing drums and singing in an open landscape."

D.3.2 Words Count

darkness(3), night(3), dark(2)
warriors(4), soldiers(2), guards, goblins, pirates, hooligans
battle(8), war(2), apocalypse(2), combat, brawl
slaves(3), prisoners, defeated army
swords(2), armors(2), chains, weapons, metal discs, guns, cannons, whips
angry(2), fight(2), afraid, danger, energy, threat, anxiety
ritual(4), dance(4)
fire(7)
people(8)
tribal(4)
jungle(2), forest, yellow/green grass
fog(2)
Mongolian(2), primitive(3), vikings(2), American Indians, aboriginal, Mordor, roman, medieval
execution(2)
singing(2), throat singer

D.3.3 Sonic Features

"drum pattern, thin metal parts.."

"shouting people, drumbeat, weird continuous sounds by a person, sword sounds."

"in general, lo-frequency sounds, with drums, voices, and some crazy strings all over."

"Throat singing, drums, repetitive nature"

"mainly the voices imitating a didgeridoo"

"The throat signing invokes tribal images. The slow build, shouting, metallic sounds and bangs invoke the image of a battle. The climbing, pulsing tones create building tension."

"Drums, because they make me feel that something is going to happen"

"Yes, as specified, the throat singing, and the drum choice, and then further instruments/instrumentation. I am a musician and go to lots of gigs so I guess I'm quite literal and I often go to the musical structure and live performative aspects rather than picturing music as an image. I've never been very good at that unless it is given to me in words first. Okay, it also did sort of make me think of Dethlok and the opening to that."

"Loud guitar"

"Dragging rhythm, chaos"

"Guitar"

"The drums added to the feeling for a war march"

"the voice and drums - tribal + forest. the increase in intensity and metallic sounds - prepare for a battle of some sort"

"I think I tend to link metal sounds to the city. The loud sounds at the start made me think of a fight, with the sounds of glass breaking midway through linking it to a pub fight"

"no, it was the rhythm that created those meanings. maybe the percussion sounds

in the middle further fortified the tribal sense it delivered (at least to me)."

"the initial voice. the drum. the synth. the changing of semiton (create unbalance). the percussions. the several voices that come up in the second part"

"The drums mostly (and I guess the didgeredoo, but I only now realise it might be one)."

"Human voices. Glass breaking. Drums"

"The voice of a man, drumming, and some sounds like electronic"

"The vocal"

"I recognized the singing being performed as a traditional method, so I thought of Mongolian nomads on the steppes (thus "tribal" and "valley." The slightly (to western ears) discordant harmonies made me think of darkness, and the shouting men in the background made me think of war."

"The throat singing and drums."

"People screaming in the background (city, dense); sounds reminding gunshots"

"There is a sound that reminds me of swords through the whole song-which is why I chose ‘war’, and also there are human voices which made me think of some soldiers’ ritual which is connected with the forest(for me). All other words-cold,dark,.. are effect of the whole song."

"Rhythmic beat of the drums reminded me of warriors marching; The sound similar to Aboriginal didgerydoo reminded me of tribal ceremonies; Tambourine-like sound seemed like the start of the dramatic event, such as a battle. Inclusion of other instruments and voices, and increase in loudness sounded like the drama of the battlefield is intensifying."

"Rain and drums"

"complete sound is very ugly"

"The troat singing and the chimes"

"Machine sounds and vocal modulations"

"metallic sound, loud drum, constant pattern, people screaming in the background, high pitch increasing tension"

"Drums and yelling"

"The Mongolian chants. Maybe because of a TV series that I've watched before"

"The oah! sound that is constantly in the song and the kettledrums"

D.4 Modified Song

D.4.1 Descriptions

"A futuristic ritual"

"A night club with disco lights and fog, and people dancing."

"A serene environment, somewhere in the past. Everything seemed joyful."

"A woman walking on a desert ocean shore. She is powerful, content, relieved. She is wearing a draped robe that fluctuates in the wind like her hair.

"Black and Blue deep sea, I'm diving slowly moving with current. Behind me the rocks sea becoming more dark and the light penetrate the water, I see the movement of seaweed and rarely a fish run away. Until I touch the sand with feet

"During the song I saw flowers and beautiful spring colours come to my mind. I felt such a rebirth of nature which comes with spring. Then at the end the song has become more heavy and gloomy, as a storm has arrived to disturb this beauty of nature. So I could see a real scene of a living nature.

"flashes of white dots and blueish fog in front of an all-black background

"i imagined a mystic travel inside of a vein. it was as if I had entered in one of those 80's documentary with all of those hallucinogenic colours and graphics. i would travel in this tunnel of colour, while the outside was completely black. it felt as if this tunnel was suspended in an infinite empty space. then when the bass kicked in i realized I could be somehow inside a living creature, and that that

could be the sound of its heartbeat."

"I saw drums and people of tribal cultures chanting. Memories of Mongolian monks arose

"I saw forests in my mind. Lots of trees and different shrubbery. Some animals like hares dashing around and a spirit of some sort moving around the trees. When the tune changed I saw the sea. It was rough and stormy. Awe inspiring. It was alive. It was as if there was a strong force that moved with the sea.

"I see man on a desk working on this track and ancient men growling to traditional chants

"I visualised men with naked upper bodies playing drums in a monastery on a mountain.

"I visualized a group of people playing together around a fire.

"It sounded like the music to a computer game with a theme similar to Game of Thrones, Lord of the Rings or similar. Fateful, epic, big groups of people moving.

"It was an individual telling a story of a person who lost something and then comes over it. The intro part felt like the beginning of setting of the story. The tempo then began visualizing the story of his troubles and then the verse section prior to outro felt like he was overcoming his troubles

"Kind of awakening of nature... maybe because now it's sunny and it suggested me this while at the end it appeared a negative image, black, a sinister sensation

"Made me feel judged by a supreme court; the sentence was given, and I was found guilty."

"Many people singing together. Large drums."

"Rising sun, but harsh. A man dressed in sports gear running / jogging out of a cave that is perched high on mountains, like a victory run. The run was interrupted by visuals of multiple percussion instruments that started playing and I tried to visualize those."

"Rolling hills of a country side with an upbeat and positive tone however turns

dark and ominous as it crosses a more mountainous region when there is some impending event"

"Some sort of Celtic rituals in the woods, the air filled with mist."

"The earlier parts of the track had an affectation often seen in dream pop/shoegaze music. I would describe it as a foggy ethereal feeling with grainy visuals. "The latter half of the track had a more tribal repetitiveness it to it. Easy to picture a group of people stomping feet and preparing to chant in unison."

"The tambourine made me visualize a tambourine"

"tribal ritual"

"Probably a wood, with sun rays filtered by tree branches, and the arrival of someone heavy, tall, almost scary."

D.4.2 Words Count

tribal(2), ritual(3), naked man painted, group of people (4), dance
fog(3), mist
sea(2), ocean, water
black(4), blue(2), dark
mountains(3), hills
woods(2), trees, shrubs
heavy(2)
storm(2)
drums(3), percussion, tambourine
sun(2)
Mongolian monks, growling chants, monastery, Celtic
powerful, strong
rough, trouble, sinister, scary

D.4.3 Sonic Features

"The base sounds probably played a part in it, otherwise I'm not sure"

"I passi della persona che arriva,con quel forte suono costante che inizia circa a metà canzone"

"Sounds of liw pitch voiles. Sounds of drums"

"water drops"

"I think that the slow and repeating structure is soothing and calm, additionally the long deep sounds for me represent voices. Also the percussion for me evokes the image of people sitting together."

"The drums and the vocals."

"Drums and bells"

"The drumbeat from the middle to the end have it a tribal feeling. The soft high pinched notes (maybe from a piano) reminded me of fog and the green/forest setting."

"the bass creates the basic setting. At first, is rather gentle, but then it gets aggressive"

"The sound that we find during the whole song at the beginning is pure, positive and then due to drums and sounds like man voices, the sound become more deep, which can be easily related with negative feelings"

"drums, bass voice (didjeridoo-like). the fact that there is a major third, perfect fifth, and major sixth, yet a minor second (d-flat) causes distress"

"The cowbell and triplet drum beats in the end was a strong signifier of tribal repetition. The reverberating high-pitched sounds and muffled drums in the first half contributed to its dreamy sound"

"I cannot identify the instruments - but the main melodic line"

"The beginning was peaceful followed by a build up in the music sort of like an ascending direction but then it got troubled and conflicted and ended on a cold and disturbingly final sort of mood."

"tribal: the percussion and vocals; foggy: the string instrument"

"There wasn't a particular instrument, but the whole piece gave the feeling described by the words I selected."

"Synth like arches: blue open space, Drum: rocks and deep diving. Synth like chorus of voices: dream and dark"

"the drums"

"the voice and the sidechain compressor"

"Throat singing sound = Tribal. Church Organ sound = contrast to Drums. Sounded etherial = In the Sky Light"

"Tribal - In the second 'half' - The minor second interval in the bass combined with the switch to rigid and more staccato rhythmic feel. - In the first half - the use of pentatonics. Contrast - the switch from 'open' to 'dense' harmony and from legato to staccato, also the bass amplitude envelope introducing some syncopation. On a mountain and floating - during the first part of the tune - pentatonics and various keyboard pads. I enjoyed it."

"It started off happy enough, but then a kind of hoover low-frequency growl starts at 30 seconds, and it it becomes more dark, but still relatively calm..."

"Deep drumming sound with drawn out notes. Piano has the uplifting floating notion"

"Chimes are associated with sea breeze. Long notes in electric guitars make shiny tones. The repetitive simplistic drums sound like a preset phone alarm "