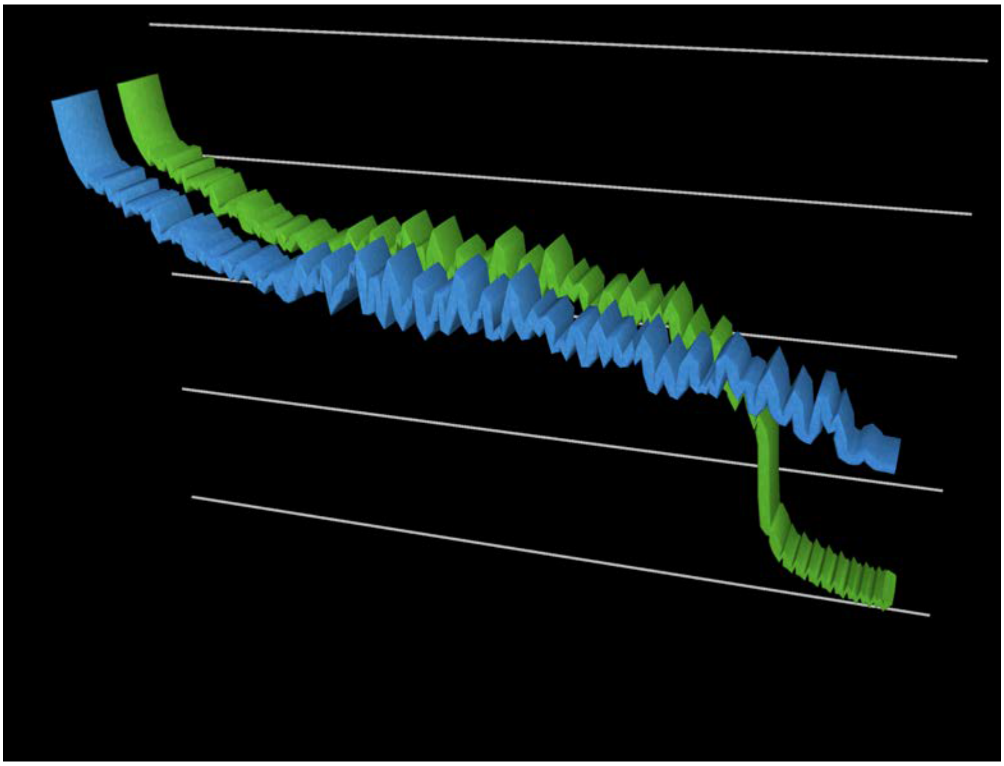


# on media playback format's role for the emotional impact and the appreciation of music in everyday life

a master's thesis


**Meng Wu**


supervisors:

Prof. Dr. Stefan Weinzierl

Dr. Steffen Lepa


Audio Communication Group

Technische Universität Berlin

Master's Thesis

# On Media Playback Format's Role for the Emotional Impact and the Appreciation of Music in Everyday Life

MENG WU

[redacted]

Supervisors:
Prof. Dr. Stefan Weinzierl
Dr. Steffen Lepa

*Fachgebiet Audiokommunikation*
Institut für Sprache und Kommunikation
Fakultät I - Geistes- und Bildungswissenschaften
Technische Universität Berlin
Berlin, 2015

1. und 2. Gutachter:
Prof. Dr. Stefan Weinzierl, *Leiter*
Dr. Steffen Lepa, *wissenschaftlicher Mitarbeiter*
*Fachgebiet Audiokommunikation, TU Berlin*

Masterarbeit 2015
Fachgebiet Audiokommunikation
Institut für Sprache und Kommunikation
Fakultät I - Geistes- und Bildungswissenschaften
Technische Universität Berlin

Sekretariat EN-8
Einsteinufer 17c
D-10587 Berlin
Telephone +49 30 - 314 222 36
Fax +49 30 - 314 211 43

Cover: Spectral distribution of one music piece in CD quality (blue line) and in Youtube quality (green line). Horizontal axis: frequency in Hz (from left to right: approximately 20Hz to 22kHz, linear); Vertical axis: level in dB (upper limit 0dB, lower limit -140dB).

Typeset in LaTeX
Berlin, Germany 2015

Meng Wu
*on media playback format's role for the emotional impact and the appreciation of music in everyday life*
a master's thesis

*Das Publikum sucht Wohlleben und Zeitvertreib,*
*nicht Belehrung.*

— Arthur Schopenhauer

# ACKNOWLEDGEMENTS

This study is also a small part of the Cluster of Excellence research project "Media and Emotion" (part of the Berlin Cluster of Excellence "Languages of Emotion").

I must express my sincere gratitude to some dozens of people for this thesis.

First of all, the supervisors: many thanks to Dr. Lepa for his kind and patient help before, during and after the planning, realization and analyze, as well as the discussion of this thesis; and to Prof. Dr. Weinzierl for his prompt admission and kind support for the subject recruiting. Without their profound knowledge and experience the interpretation of this work would have been much more difficult and slowly.

Then, of course, all the 49 punctual, polite and cooperative subjects who had to personally come to the *Minilab* in spite of this crazily hot summer and put on those big headphones and patiently went through all the music pieces and the endless questionnaire.

Last but not least, my family, for their love, care and selfless support during my jobless time back in school.

*Berlin, December 2015*

# ABSTRACT


## ABSTRACT

*English*

How do music playback spatiality and audio quality affect the felt emotions when listening in everyday life? In an experiment 49 German speaking subjects aged between 20 and 30 (mean=24.18, SD=2.62) were sorted into four stimuli groups which are the four combinations of stereo vs. mono and CD quality (.wav, 44.1kHz) vs. Youtube quality (AAC, 128kbps). They heard four different musical pieces of the same stimulus type and filled in M-DAS questionnaires on felt emotions. The data were analyzed with three GLMs. The results are: Music playbacks 1) in stereo and 2) in Youtube quality could trigger stronger emotions. While the first result about spatiality is in line with the research literature, the unexpected latter finding might be explained by the assumption that the subjects are more used to this playback quality.

## ZUSAMMENFASSUNG

*Deutsch*

*Wie beeinflussen Wiedergabespatialität und -qualität die empfungene Emotionen beim alltägigen Musikhören? In einem Experiment wurden 49 deutschsprachige Probanden im Alter zwischen 20 und 30 Jahren (mean=24.18, SD=2.62) in vier Stimuligruppen einsortiert, welche die vier Kombinationen von Stereo vs. Mono und CD- (.wav, 44.1kHz) vs. Youtube (AAC, 128kbps)-Qualität sind. Sie hörten vier unterschiedliche Musikstücke des selben Stimulustyps und füllten den M-DAS Fragebogen zur Erfassung empfundener Emotionen aus. Die Daten wurden mit drei ALMs analysiert. Die Ergebnisse sind: Musikstücke 1) in Stereo und 2) in Youtube-Qualität konnten stärkere Emotionen hervorrufen. Während das Ergebnis zur Spatialität dem Forschunsstand entspricht, lässt sich das überraschende Ergebnis zur Soundqualität wohlmöglich dadurch erklären, dass die Probanden sich an dieser Art Wiedergabequalität gewöhnt haben.*

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACRONYMS

**GLM** General Linear Model

**ALM** das Allgemeine Lineare Modell

**CD** Compact Disc

**DVD** Digital Versatile Disc or Digital Video Disc

**MP3** MPEG-1 or MPEG-2 Audio Layer III

**AAC** Advanced Audio Coding

**A/D** Analog to Digital Conversion

**D/A** Digital to Analog Conversion

**HiFi** High fidelity

# Part I

## THEORY AND METHODOLOGY

The way how people listen to music has been changed a lot in the past century. Recorded and digitalized productions gave us a new sound of music: can this even change our listening aesthetic in everyday life? In this study, given the assumption that the younger generation today, whose primary source of music is streamed service, is affected by this shift of art of music playback and have a “new” listening aesthetic, an experiment was designed and carried out. In this part, the theoretical as well as methodological preparation are documented.

# 1 INTRODUCTION

## 1.1 CONCERNING MUSIC

The answer to the question "what is music" can always be found in different books like "music theory". And the answer always lies already in the introduction or preface. In the introduction of the Chinese version of the popular schoolbook *Basic Music Theory* by Li [1], music is "a collection of tones created by the people, resulting from a long-term production life and class struggle". This definition of music is surely a materialist's view, while thousands of miles away, in Germany, the German version *ABC Musik* [2] has a much more romantic explanation on where music came from: "The gift of the gods Apollo and Muse".

*Function of music*

Plato explained music as coming from imitation of the natural world. But being used in rituals, it seemed to be able to speak the "language of gods". The ability to communicate emotions, whether consciously or unconsciously, is one central function of music. This function was used to manipulate people's emotions in religious rituals, and this is maybe the "class struggle" Li was referring to. After the "long-term production life", today's music became a musicians' tool. "Musicking" [3] as a mental interaction between the musician and the listener, is a game of emotions.

For a game, the most important element is the expectation: to experience so called "safe danger" and "expected surprise" [4]. People don't hesitate to pay to experience a performer playing a piece that has been listened to for a thousand times. The different interpretation and a new interaction with a stranger in music listening represents the seeking of "safe danger" and "expected surprise".

The reason that people listen to music is because it can evoke emotions, and people value music mainly according to the emotions it can bring. The emotions that music can evoke and the several psychological mechanisms [5] behind it is not unique to music, and researchers can always take the assumption that music can trigger certain emotions without knowing exactly how the mechanisms work. The study of emotions during music listening can contribute to general cognitive science research as unlike other research instruments or stimulus, music is completely controllable to researchers as it unfolds precisely timed emotional events and can be exactly repeated in each measure [6].

## 1.2 THE "SOUND" AND THE PROBLEMATIC

With the introduction of the art of recording in the last century, people's listening form of music has ever been changed. From recording, mixing and mastering to distribution and playback, recorded music is played out loud through headphones or loudspeakers after a long chain of A/D and/or D/A signal processing and amplifying. What people listen to today is rather called *sound* [7] than music itself. As manipulated recordings and mixings make the productions more marketable, for example the *loudness war*, concerns came that compressed sound quality may cause the listeners to lose interest for it [8]. This is reasonable as different sound of the same music content could evoke different levels of emotions and this would affect how people value these sounds as productions.

An earlier explorative study (N=43) [9] showed that the playback format (mono vs. stereo) plays a clear role in the emotional impact of a music piece: the subjects who listened to stereo versions of the stimuli felt significantly stronger emotions (self-reported) than the subjects who listened to mono versions of the same stimuli. Interestingly, none of the subjects who listened to mono versions had complained about the poor playback itself; some of them just said "the music pieces were boring". Obviously, they didn't realize that they were actually talking about the sound rather than the music.

Therefore, at the beginning of this study, it is interesting to find out what kind of role do sound playback attributes like spatial format and sound quality play, as the younger generation (university students) today may have a shifted listening aesthetic because what they mainly listen to is not live music, but compressed digital playbacks which are reduced in dynamic and frequency range and presumably also compromised in other attributes. Streaming services like YouTube, Spotify and Apple Music are taking over the market, and CD is slowly becoming history for the younger generation.

For instance, let's assume, that we have the same recorded Album, and the only difference is that some people listen to it via CD, and some people listen to it via YouTube. Do they statistically perceive different levels of emotions? Is it possible, that, statistically, those who listened to CDs say they like the Album very much, while those who listened via YouTube don't like it very much? Or even perhaps the contrary?

## 1.3 FACTORS AND COVARIATES

*Factors*

The focus of this study is then set to how different spatial formats and playback qualities affect the felt emotions in music listening for the younger generations. The main factors of the research are then the spatiality and quality. Here, the spatial playback formats have been chosen as two different formats: monophonic and stereophonic (another spatial format could be surrounded playback or "live concert simulation"). According to a survey in Germany in 2012 [7], apart from radio and CD, YouTube was the main source of music listening in Germany. Thus the playback qualities in this thesis was chosen as two different qualities: CD quality (WAVE 44.1kHz, 16Bit) and YouTube quality. According to an informal research of Vogt [10], the sound quality of YouTube is constant and independent to the video quality, and is measured as a lossy AAC quality with 126 kb/s. Considering the measure was done in a streaming environment, there could be some loss/noise of the data flow that reduced the original bitrate of 128 kb/s, which is more likely to be the intended playback quality. Therefore the "YouTube quality" in this thesis was chosen as 128 kb/s AAC.

*Spatiality*

*Quality*

*Covariates*

Nevertheless, the "musical expertise" of listening to digital music today also plays a role [11] in music perception. The expertise here includes not only professional or hobby training of a music instrument, but also the experiences or hobbies related to acoustics or sound equipments. In other words, anyone who has rich experience with digitalized music playback and is hence presumably quite aware of the playback's sound quality should be regarded as one with musical expertise. The *sound* itself plays a big role in music listening today for these people, as they are more aware of the acoustic performance of the playback. For them, the sound quality could be part of their music listening aesthetic. Hence the attribute of musical expertise within subjects could be considered as a moderator variable.

*Expertise*

Like musical expertise, another attribute that relates to a subject's experience of music perception is age [12]. Diverse research with children [13] and young adults [14] as well as with elderly people [15] have shown evidence that age has significant effects on music perceptions and understanding of emotions. Nevertheless, these evidence showed that for some cases younger subjects perceive stronger positive [14] emotions while older subjects perceive less

*Age*

negative emotions [15]. Thus age could also be considered as a covariate.

*Gender*

Diverse neurological studies [16, 17, 18] as well as empirical studies [19, 20] have shown whether significant or insignificant [14] gender effect or interaction between gender and musical expertise [21] in auditive or visual [22] perception. Therefore the gender difference must be observed and could be calculated as another covariate in this study. However, the main hypothesis of this study does not concern gender differences.

*Emitter*

There are generally two ways of listening to digitalized music playback in everyday life: through headphones/earphones or through loudspeakers. Research (for example, the research of Kallinen et al [23] and Nelson et al [24]) confirms the difference of the perception of music playback through headphones and loudspeakers. Headphones were used in this study as the sole source of stimuli. When a subject listens to music through loudspeakers more often, then the headphones would be for him/her likely a seldom/new experience, and the perception of music through headphones could be different than usual. Therefore, besides the music expertise and gender, the frequency of listening to music through loudspeakers could also be considered as a covariate.

## 1.4 RESEARCH STATUS

In the recent years researchers like Cupchik [25] and Oliver et al [26, 27] examined the role of media product (film, music etc.) in the evoking of emotions within perspectives of hedonism and eudaemonism, revealing that the feelings and emotions happen concurrently by the usage of media products. On the other hand, from utilitarian's point of view, researchers like Scherer [28] and Zentner et al [29] tried to measure the aesthetic appreciation of media products. The results from all these studies have provided rational understandings and tools to measure the relationship between auditory attributes of media products and the appreciation and gratification of them.

### ON SPATIALITY

Already in 1966 had Morgan and Lindsley tried to find out the difference in preferences for stereophonic over monophonic music [30]. In their research the subjects would directly select between stereophonic and monophonic music, and the result shows directly which playback format the subjects preferred. Over the years researchers realize that music can be used as an excellent instrument for psychological experiments [6], not only because music is a precisely repeatable stream of events, but also because, more importantly, the (perhaps sole) function of music is to evoke emotions. Today, researchers inspect effects of playback spatiality targeting not on subjective selection any more, but on perceived affective expressivity of music listening. One most resent research of Lepa et al [31] showed significant effects of playback's spatiality in music perception in the sense of affective experience of listening.

In the explorative study of Pischel et al [9] the effect of spatiality (mono vs. stereo) was also observed. Nevertheless there were also another tendency in the findings of this study: the difference between mono and stereo playback format has greater influence on the evoked emotions for naïve listeners than experienced listeners.

#### ON QUALITY

The effect of different audio playback qualities also has been diversely researched as the digitalized music today is usually very compressed. Although the upper frequency limit of human hearing is believed to be somewhere around 16kHz [32], which is the reason for most music productions to cut the higher frequencies by down sampling or by other processes to maintain a smaller file size (nevertheless this kind of process cannot always really "cut" the frequencies [33] so the linearity near the 'edge' is always steeply damaged, causing a increasingly loss of energy in the upper frequency section), different researches [34, 35, 36][1] showed the fact that the difference in higher frequencies (above 20kHz) in music listening could be discriminated by the human hearing for certain excerpts. Apart from the concerns in higher frequency, reduced bitrate of digital music production puts limitations not only to frequencies, but also to acoustic attributes like dynamic ranges. In this sense researchers [37, 38, 39][2] compared popular digital audio

1 In the study of Yoshikawa et al. [34] (N=11, age 21 to 24) 2 subjects (18.2%) could discriminate the difference of 96kHz and 48kHz sample rate for both given sound materials. Ando et al. [35] (N=13, age 19 to 51) reported similar findings that some subjects were able to discriminate musical sounds with or without high (above 21kHz) frequency components. In the experiments of Pras and Guastavino [36] (N=16, mean age=30, SD=7.1) All these results were observed depending on the sound material that was given to subjects.

2 Pras and Guastavino [37] tested 13 subjects (mean age=28, SD=5.6) and the listeners could distinguish CD from MP3 up to 192 kb/s and preferred the former. In the experiment of Ruzanski [38] (N=5, age 19 to 32) with all listeners "untrained", some of them couldn't even tell the difference between CD and MP3 in 96 kb/s for some musical tracks. In the tests Olive [39] conducted (N=18, age 15 to 18) the teenager listeners preferred CD than MP3 in 128 kb/s. In addition to the above findings, in the study of Pras and Guastavino a significant effect of musical expertise and music genres was observed.

solutions like CD, DVD, MP3, AAC etc. and could mostly justify the distinguishing of different bitrates in dependency to the musical expertise of listeners and the music type that were given to them. All these research results indicate that the discrimination of higher frequency and compression of bitrate is possible and is dependent on the type of music stimulus. In other words: different playback quality could affect the affective perception of musical expressivity.

#### THE LEARNING EFFECT

As Olive referred in his study [39] on teenagers, the research of Kirk [40] and some discussion over the Internet [41, 42] on the informal research[3] which Jonathan Berger has done, as well as other internet media discussion [43, 44, 45] on the same phenomena have been proposing the idea that the younger generation prefers "bad sound over good sound" because of their preconceived playback quality. According to Kirk, this is a kind of *learning effect*[4].

### 1.5 HYPOTHESES

Generally the hypotheses of this study consist of the effect in music perception of different **a)** spatial playback (mono vs. stereo) and **b)** playback quality (CD quality vs. YouTube quality). With regards to the above mentioned studies and explores the stereo playback spatiality is hypothesized to have a positive effect on felt emotions when listening to music. As for the playback quality, although the *learning effect* being discussed already gave a clue that in the younger generation the compressed playback quality may have a positive effect in affective expressivity of music listening, the chosen playback quality in this study is not MP3 but AAC. Some subjective listening tests suggest that the AAC format is much better than MP3 format at 128 kb/s and does not have audible sizzle-aka distortions mentioned in the above studies. Hence the playback quality of CD was still hypothesized as having a positive effect on the perceived musical emotions.

As mentioned above, **c)** additional effects of musical expertise, age, gender and experience with loudspeakers on spatiality and quality

3 Jonathan Berger is a Stanford University professor of music. He conducted a six-year informal research on first-year students by giving them the same musical stimuli with different formats. He discovered that from year to year, more and more students preferred MP3. Berger has the opinion that they are used to and hence prefer the so called sizzle-aka distortion of compressed music.

4 In his study in 1956, he found out that "1) learning plays an important role in determining preferences for sound reproducing systems; 2) continued contact with a particular system produces shifts in preference for this system; and 3) the average college student prefers music and speech reproduced over a restricted frequency range rather than an unrestricted frequency range."

are expected to be observed. And their effect on the felt emotions are hypothesized as following:

**EXPERTISE:** According to the explorative study of Pischel et al [9], people with musical expertise are less affected by the different playback spatiality; and at the same time they appear to perceive weaker emotions than non-expert listeners. In other words, subjects with musical expertise appeared to "stay calm". Therefore musical expertise is hypothesized to have negative effect on felt emotions.

**AGE:** As mentioned above, subjects advancing in age perceive weaker emotions. Therefore age is hypothesized to have negative effect on perceived emotions.

**GENDER:** In the above mentioned studies, women are observed as more sensible than men on musical expressivity. If the array is defined as Sex[`0 = female, 1 = male`] (For details, refer to the data cleaning section on Page 19 in the next chapter) then gender should be hypothesized to have a negative effect on observed self-reported emotions.

**EXPERIENCE WITH LOUDSPEAKERS:** The more experience with loudspeakers the subjects have, the newer the experience with headphones in this experiment would appear to them. As generally hypothesized that newer experiences bring better concentration and therefore result in stronger felt emotions, the experiences with loudspeakers is then hypothesized to have positive effect on perceived musical emotions.

# 2 METHODS

Now that the theoretical preparation was done, practical operations like preparation of musical stimuli and questionnaire, recruiting of subjects were carried out before the experiment. After the experiment the collected data were sorted, rearranged and pre-calculated so that they could be analyzed according to different hypothetical models.

## 2.1 THE MUSICAL STIMULI

The following circumstances were taken into consideration when choosing the musical stimuli:

NUMBER: The number of musical stimuli were set to four, as well as other similar researches like [31, 9]. This could give the experiment and the statistical models enough measurement levels and test power. Too many musical stimuli could bore or exhaust subjects and lead to distorted test results.

LENGTH: The length of each musical stimulus should not be too short or too long. When too short, the felt emotions could then have a higher probability of randomness; when too long it could bore and exhaust the subjects and it would also be a waste of time. Therefore the chosen stimuli were all approximately 2 to 3 minutes long.

GENRE: Recent research about music genres [46, 47] claim that the common division of music types per genre (like Classic, Pop, Jazz, Rock etc.) may not be accurate any more in the context of music preferences[1]. Nevertheless the experiment itself should concentrate on felt emotions, so the chosen music pieces should be able to trigger strong emotions. From these considerations four instrumental film music pieces were chosen, and they are listed here in Table 1. They represent different emotions and some of these are quite complexed (for example, the Stimulus B). These pieces are all instrumental so there would be no cultural misunderstandings for lyrics.

ACOUSTICS: As the experiment involves technical acoustics attributes as factors, especially playback quality, two aspects

1 The research of Rentfrow [47] has given a new five-factor model for music preferences (Mellow, Unpretentious, Sophisticated, Intense and Contemporary) which describes mainly how these music types make listener feel, rather than the conventional genres which only describe the music type's common category.

were considered: 1) the original source of these stimuli should at least have the higher quality in the experiment, which is the CD quality; 2) the stimuli themselves should have a balanced or commonly accepted frequency distribution.

Table 1: The four stimuli

| **music pieces A, B, C and D** | **description** | **duration** |
|---|---|---|
| Joe Hisaishi<br>*No Way Out*<br>キッズ・リターン, film music, 1996 | electronic<br>desperate,<br>repeated | 2:52 |
| Tom Tykwer<br>*All Boundaries Are Conventions*<br>Cloud Atlas, film music, 2012 | orchestral<br>inspiring,<br>encouraging | 2:38 |
| Ennio Morricone<br>*Ma L'Amore no*<br>Malena, film music, 2000 | orchestral<br>happy,<br>sexual fantasy | 1:52 |
| James Howard<br>*End Credits*<br>Waterworld, film music, 1995 | orchestral<br>world,<br>thoughtful | 2:18 |

As listed in Table 1, two of them are from the 90s, and the other two from the new century. The first stimulus, *No Way Out* comes from Japan, and its composer, Joe Hisaishi, composes in a western way. This is the only stimulus that is not orchestral / live recorded, but electronically produced with synthesizer.

These four stimuli will be further referred to as stimuli A, B, C and D.

## 2.2 THE VERSIONS: FORMATS AND QUALITIES

*Imported as .wav files from CDs*

*stereo and mono versions*

The original data of the four chosen stimuli in Table 1 came from commercially produced CDs. They were imported as `.wav` files to the computer (44.1kHz, 16-Bit, 2CH). A duplicate of these files were then processed with the software tool Audacity (Version 2.1.0) to mono versions. So now, in this step, there are two files for each stimulus: one stereophonic, another monophonic. Their file format are still the same, and yet representing the same playback quality.

Due to the prevalence of this format in the current media sphere, as justified in the introduction chapter, the chosen YouTube quality is AAC, 128kbps. Each one duplicate of the stereophonic and monophonic files from the last step were processed with xACT (X Audio Compression Toolkit, Version 2.36) into AAC format (`.m4a` files).

*CD and YouTube versions*

As a result, there are four files of each stimulus, and they are labeled with initials shown in Table 2:

Table 2: Label system of stimuli versions

| | stereophonic | monophonic |
|---|---|---|
| CD-quality | SH | MH |
| YouTube-quality | SL | ML |

It is easily understandable, that the S and the M stand for stereo and mono, while the H and the L stand for high (CD) and low (YouTube) quality.

The files were then renamed with a combination of its stimulus label and its version, for example, BSH stands for Stimulus B in stereo format and in CD quality. Then, from ASH to DML, there are totally 16 stimuli files ready.

## 2.3 THE SUBJECTS

The aim of this study is to describe the playback format's impact on music listening in a certain social group and try to generalize it. And, the experiments were planned to be conducted in a room on the main campus of the Technical University of Berlin (TU Berlin), due to the age distribution on the university campus the age group was limited to 20 to 30 years.

*age range*

As the questionnaire (details in next section) consists of words that describe delicate emotions, it is important that all the subjects have a common comprehension of all these words; nevertheless the cultural difference is of no interest to this study, therefore the subjects selection was limited to native German speakers.

*cultural background*

Roughly, the subjects would come from two sources: "internal" students of the Audio Communication Group, who are obligated to participate in a total amount of three hours hearing experiments, so as to finish one of their mandatory study models; and "external" students of the TU Berlin. For the external students there was a small reward of 5 Euro, and for the internal students there was no

*sources of subjects*

monetary reward but a signature from the author to confirm that they participated in this hearing experiment.

*the "cover story"*

There were generally three ways to advertise the hearing experiment: via E-Mail lists, via posters on the campus, and via posts on Facebook pages. To accomplish an experiment in terms of a single-blind trial, by writing the content of the poster and E-Mails a *cover story* was used, which was *"this experiment is about musical emotions of different music styles"*. By doing so, the subjects would not be implied of what the experiment was really about, and, furthermore, they'd probably concentrate better on the listening and their felt emotions.

*recruiting by pre-registration*

All the interested applicants were directly asked to fill in a online registration form at their first contact with the author. In the online registration form socio demographic data of the potential subjects, for instance gender, age; and their musical expertise were collected, so that they could be sorted into the four stimuli groups SH, SL, MH and ML. This assures that the attributes of gender and musical expertise are distributed in each group as similarly as possible to other groups.

Specifically, besides the gender and age which are simple questions to ask, the inspection of musical expertise was realized through two questions (shown below, already translated into English), to which if a subject answers at least one "yes", he/she would be considered as a subject with musical expertise. These questions are:

*inspection of musical expertise*

Q1 Do you play a music instrument? (If yes, how many years?)

Q2 Do you have job/study/hobby related to sound/acoustics? (If yes, how many years?)

Applicants whose age were outside the age range of 20 to 30 were told they could not participant in this experiment, and their application could not be sent. When one applicant had successfully finished the online registration form, he/she would be then put to a certain group where there were still vacancies for subjects with the same attributes. The attribute of gender and musical expertise were considered by the distribution operation, and age was not considered. After doing so, the author makes an appointment with each subject. In each appointment only one subject participated in the experiment.

## 2.4 THE MEASUREMENTS

### 2.4.1 *Measuring felt emotions: The M-DAS*

As the function of music is to trigger emotions, the questionnaire should then focus on the emotional impact of listeners. Therefore the Modified Differential Emotion Scale (M-DES, in German M-DAS, Modifizierte Differentielle Affekte Skala [48]) which already helped deliver satisfying results in the pilot study [9], was used again in this study. It is a set of words which describe emotions each with a five-level Likert-scale. Each subject would then fill one M-DAS after listening to each musical stimulus, totaling a number of four M-DAS measurements levels[2].

Some emotions, i.e. negative emotions, which are believed to be unable to be significantly felt during media or music usage, were deleted from the Scale. According to the Table 2 in [48], these emotions are *Wut*, *Ekel*, *Angst*, *Scham* and *Schuld* (anger, disgust, fear, shame and guilt). The scale then consisted of 11 words, and are shown with their English translation in Table 3:

Table 3: M-DAS emotions used in this experiment

| German | English |
|---:|---|
| Vergnügen | enjoyment |
| Freude | cheerfulness |
| Zufriedenheit | satisfaction |
| Zuneigung | fondness/love |
| Faszination | fascination |
| Ergriffenheit | sentiment |
| Interesse | interest |
| Überraschung | surprise |
| Trauer | sadness |
| Verachtung | contempt |
| Langeweile | boredom |

Each of these 11 words would then be presented with three sub-scales (see the list of all M-DAS questions in APPENDIX 1: M-DAS).

As each subject had to fill in four M-DAS inventories, the questions were randomized each single time it was shown, so as to

2 As the subjects filled the questionnaires in German, this scale is referred to as M-DAS instead of M-DES in this thesis.

avoid repetition effect or similar effects.

### 2.4.2 *Listener experience and miscellaneous*

As discussed in the introduction chapter, apart from the musical expertise, it was also interesting to know if a subject was used to headphones or loudspeakers. Therefore the following questions were asked:

*inspection of exposure to loudspeakers*

Q1 How offen do you listen to music via headphones? (Daily / weekly / monthly / yearly / never)

Q2 How offen do you listen to music via loudspeakers? (s. above)

The listening experience questions show the listening behavior of the subjects. This is quite similar to the musical expertise questions in the registration form, but for the following reasons they have to be asked, and are only asked **after** the subjects have listened to the stimuli and finished the M-DAS questionnaire:

- These questions can imply the true purpose of the experiment and violate the term of a single-blind trial.
- It is possible, that some people do not have any experience with music instrument or acoustics, yet are still very experienced with digital music listening and aware of the playback format and quality. So these questions could help correcting the musical expertise questions by the data analyzing.

All the questions in the registration form were asked again here in the formal experiment, after the M-DAS: gender, age, and musical expertise. This was not only to save some effort combining the data of the registration form to the main stream, but can also ensure that any error that could have been made during the organizing of the subjects could be corrected now, as the whole experiment was anonymous. This was useful as for example, a subject registered before his 23rd birthday but came to the experiment after it.

The first page of the questionnaire was for the author to fill in, where the author selects the group of this subject (SH, SL, MH or ML), so that this attribute could also be combined into the main data set. This action was not seen by the subjects.

Every subject was also asked to write some "comments" at the end, if they wanted. This was but not mandatory.

## 2.5 HARDWARE AND SOFTWARE

On each experiment appointment only one subject was treated. All subjects finished their experiment in the same room, with the same equipment and at the same loudness. The experiment was carried out in the *Minilab*, which is the tiny media lab room H2001D in the *Hauptgebäude* of the TU Berlin. The room is approximately 1m × 2m in size, with a desk and a chair, and an overhead light. The walls are furnished with soundproof material. The subjects finished the questionnaire inside H2001D, while the author played music from outside the room. Below is a listing of used hardware and software:

### FOR INSIDE THE ROOM:

- Macbook Pro with Retina 13", early 2015, Mac OS X Yosemite, to fill in the online questionnaire;
- Limesurvey system as survey software operated online from the server of the Audio Communication Group;
- Terratec AUREON XFIRE8.0 HD USB external sound card, connected via USB to the Surface Pro and via cable to the Audio Technica ATH-M50x.
- Headphones: Audio Technica ATH-M50x;
- An instruction on the wall, which the subjects are requested to read before the experiment (see APPENDIX 2: PRINTED INSTRUCTION IN THE LAB).

### FOR OUTSIDE THE ROOM:

- Microsoft Surface Pro 64G, Windows 8.1 64bit, to play the stimuli, connected via USB to the Terratec AUREON inside the room;
- Winamp for Windows 5.666 Full.

## 2.6 EXPERIMENTAL PROCEDURE

To demonstrate the detailed procedure of an experiment with one subject, the operations are listed below, from the first contact until the finishing of the experiment:

a. The subject saw the information via E-Mail/Poster/Facebook and wrote the author an E-Mail wishing to participate;

b. The author sent him/her the address of the online registration form;

c. After the subject successfully (age in the desired range) finished the registration form, the author assigned him/her to a certain stimuli group (SH, SL, MH or ML) and then make an appointment with the subject via E-Mail;

d. The subject came to the *Minilab* on appointed time;

e. The subject waited outside the room while the author opened the questionnaire and filled in the first page where the stimuli group was set, then invite the subject to come inside the room and to sit down before the computer, read the instructions and put on the headphones;

f. The author closed the door and played the first stimulus from outside, and the subject filled in the first M-DAS after listening to it;

g. The subject finished the first M-DAS and told the author to play the next one. The last step and this step were repeated until all four M-DAS were finished. Then the subject continued filling in other questionnaires. They could ask the author if they did not understand a question;

h. All questionnaires were finished. The author gave the subject reward/signature.

## 2.7 DATA ANALYSIS

All data were processed and analyzed with the statistic software IBM® SPSS® Statistics (Version 23).

### 2.7.1 *Data cleaning*

To use the collected raw data more efficiently, they had to be defined, recomputed and modified with the following data cleaning procedure:

VERSION represents the stimuli groups SH (1), SL (2), MH (3) and ML (4). At the beginning of each experiment procedure, as the author selects the version on the first page of the questionnaire, a value of `1` to `4` was given to this string.

SPATIALITY was set to `0` for mono (if `version = 3` or `4`) and `1` for stereo (if `version = 1` or `2`).

QUALITY was set to `0` for YouTube quality (if `version = 2` or `4`) and `1` for CD quality (if `version = 1` or `3`).

M-DAS SCALES were defined by simply taking the mean value of every three sub-scales for each scale (refer to APPENDIX 1: M-DAS), for each measurement level. There were four stimuli, which means every subject answered $33 \times 4$ M-DAS questions, so the results were 11 M-DAS scales for every stimuli group A, B, C and D.

EXPERTISE consisted of two questions mentioned above on Page 14, and was set to `0` (no expertise) as default and subjected to be true (`expertise = 1`) if answer to any of the both questions was "yes".

GENDER had only two values `1` (for female) and `2` (for male) in the Limesurvey system. To realize a binary and logical computation in SPSS, each value was deducted by 1 resulting `0` = female and `1` = male.

AGE was a simple scale value and could be directly used in the data analysis in the next steps.

EMMITER was directed in the Limesurvey system as "daily, weekly, monthly, yearly, never" from `1` to `5`. This direction had to be revised in the data cleaning, as a united direction of frequency and value was desired. So it was redefined as "never" `= 1`, "yearly" `= 2`, $\cdots$, "daily" `= 5`.

### 2.7.2 *Data analysis*

Initially, descriptive statistics were calculated to describe the obtained sample's composition. Then scale variables were calculated for the sub-scales of M-DAS and their reliability was checked by calculating Cronbach's Alpha.

Finally, repeated measurements MANOVA was used in order to test the central hypotheses.

Part II

# RESULTS AND DISCUSSION

49 subjects from the TU Berlin finished the experiment. Results from the data analysis justified almost all the theoretical hypotheses, except that YouTube quality evoked stronger emotions than the CD quality. Detailed documentation of the data analysis results and some discussions over the conclusion are included in this part. At last, some criticisms and methodological outlooks are made at the end of this thesis.

# 3 RESULTS

## 3.1 SAMPLE STATISTICS

In total, 49 subjects from the TU Berlin came to the *Minilab* and finished the experiment. 28 of them were male (57.1%) and their age were all between 20 and 30 (mean = 24.18, SD = 2.62). As stated in the previous part, the author managed to distribute them as evenly as possible to the four stimuli groups according to their gender and musical expertise. The distribution of gender and musical expertise[1] is shown in Figure 1:

Figure 1: Subjects distribution according to gender and musical expertise: sorted by versions heard.


Generally, 33 recruited subjects (67.3%) were with musical expertise. It is also obvious to see, that recruiting male subjects with musical expertise was easier than other groups on the TU Berlin campus.

1 Here, subjects with musical expertise are called expert for short, and others called laymen.

## 3.2 RELIABILITY TEST

Scale variables for the 11 emotions (a comparative translation of these emotions into English are listed in Table 3 on Page 15) of the M-DAS were initially calculated by taking the mean value of their sub-scales. Then, Cronbach's Alpha, the the arithmetic mean for Cronbach's Alpha over four stimuli and a standard deviation were calculated. Results are listed in Table 4.

Table 4: Cronbach's Alpha of all 11 emotions for the four stimuli A, B, C and D.

| Emotion | A | B | C | D | Mean | SD |
|---|---|---|---|---|---|---|
| Vergnügen | 0.73 | 0.81 | 0.93 | 0.83 | 0.82 | 0.06 |
| Freude | 0.90 | 0.91 | 0.94 | 0.84 | 0.90 | 0.03 |
| Zufriedenheit | 0.81 | 0.84 | 0.82 | 0.80 | 0.82 | 0.01 |
| Zuneigung | 0.87 | 0.85 | 0.87 | 0.88 | 0.87 | 0.01 |
| Faszination | 0.80 | 0.81 | 0.84 | 0.93 | 0.84 | 0.05 |
| Ergriffenheit | 0.71 | 0.85 | 0.80 | 0.87 | 0.81 | 0.06 |
| Interesse | 0.71 | 0.77 | 0.71 | 0.87 | 0.76 | 0.06 |
| Überraschung | 0.84 | 0.85 | 0.83 | 0.87 | 0.85 | 0.01 |
| Trauer | 0.78 | 0.71 | 0.54 | 0.72 | 0.69 | 0.09 |
| Verachtung | 0.86 | 0.89 | 0.91 | 0.71 | 0.84 | 0.08 |
| Langeweile | 0.80 | 0.85 | 0.91 | 0.84 | 0.85 | 0.03 |

Marked red, the Cronbach's Alpha for the emotion Trauer in the third stimulus is lower (.54) as the threshold of .7. And the comparatively low mean of the sub-scale 'Trauer' rises further doubts of the suitability of this sub-scale for measuring musical emotions in this experiment. But on the other hand, the stimulus C was an "extremely happy" one, which means having reliability problem by measuring sadness for this stimulus shouldn't affect its reliability in other measurement levels for the stimuli A, B and D.

## 3.3 THE GENERAL LINEAR MODELS (GLM)

As the general hypothesis of this study proposes, there could be at least two GLMs to test:

1) one with spatiality and quality as factor without covariates; and

2) one advanced GLM on basis of the first one adding all the hypothesized covariates.

In the actual analysis, due to homogeneity problems of the emotions scales 'Verachtung' and 'Überraschung' and test power problems, a third GLM was used which is:

3) a final GLM based on the 2nd one, leaving out the interaction of spatiality and quality (`spatiality*quality`) and the emotions scales with homogeneity problems in the 2nd GLM (`Verachtung` and `Überraschung`).

By testing these GLMs, the sphericity assumption and homogeneity of subjects were testified prior to the calculating of the MANOVA.

A detailed list (Table 18) of these GLMs and their results as a summary is to be found at the end of this chapter (on Page 37).

### 3.3.1 *1st GLM*

To testify the basic hypotheses for **a)** spatiality and **b)** quality, an initial general linear model with four levels of repeated measures and `spatiality` and `quality` as factors was used.

MAUCHLY'S TEST OF SPHERICITY was carried out in terms of the necessary sphericity of the underlying data. The values of the Greehouse-Geisser corrections in the Mauchly's Test of sphericity for each emotion scale are listed in Table 5:

*sphericity test*

Table 5: Mauchly's Test of Sphericity for the 1st GLM

| Emotions scales | Greenhouse-Geisser |
|---|---|
| `Vergnügen` | .86 |
| `Freude` | .87 |
| `Zufriedenheit` | .82 |
| `Zuneigung` | .94 |
| `Faszination` | .90 |
| `Ergriffenheit` | .92 |
| `Interesse` | .92 |
| `Überraschung` | .92 |
| `Trauer` | .85 |
| `Verachtung` | .64 |
| `Langeweile` | .74 |

As the probability of all Mauchly's test statistics with Greenhouse-Geisser corrections were larger than .05, the null hypothesis that the conditions of all variances are equal couldn't be rejected. Therefore the sphericity assumption was validated for the 1st GLM.

*homogeneity test*

**LEVENE'S TEST FOR THE HOMOGENEITY.** The assumption of this experiment was that the subjects variances in each group were equal. To test this homogeneity of variances, Levene's Test was conducted and the results are shown in Table 6.

Table 6: Levene's Test of equality of error variances for the 1st GLM. Values <.05 are marked red.

| Emotions scales | A | B | C | D | Mean |
|---:|---|---|---|---|---|
| `Vergnügen` | .84 | .95 | .39 | .22 | .60 |
| `Freude` | .73 | .68 | .07 | .16 | .41 |
| `Zufriedenheit` | .99 | .75 | .40 | .76 | .73 |
| `Zuneigung` | .49 | .09 | .32 | .97 | .47 |
| `Faszination` | .31 | .87 | .70 | .19 | .52 |
| `Ergriffenheit` | .86 | .29 | .80 | .99 | .73 |
| `Interesse` | .83 | .63 | .81 | .67 | .74 |
| `Überraschung` | .72 | .34 | .22 | .20 | .37 |
| `Trauer` | .41 | .40 | .53 | .28 | .40 |
| `Verachtung` | .19 | .00 | .00 | .00 | 04 |
| `Langeweile` | .44 | .00 | .00 | .00 | .11 |

For the emotions scales `Verachtung` and `Langeweile` for the 2nd, 3rd and 4th stimuli (marked red) the null hypothesis of equal variances was rejected. This concludes that for these emotions and stimuli there was a significant difference between the variances in the subjects, and that the homogeneity could be questionable for these combinations.

*MANOVA*

**THE MANOVA** was conducted for all emotions scales. Table 7 shows the multivariate results according to Pillai's Trace for `stimuli`, `spatiality` and `quality` and the interaction between `spatiality` and `quality`.

*multivariate effects*

**THE MULTIVARIATE TESTS** for between subjects effects in Table 7 indicated that for the stimuli and spatiality, the null hypothesis that there was no significant difference could be rejected, and this was

Table 7: Pillai's Trace for the multivariate effects in the MANOVA of the 1st GLM. df = 11.

| | value | F | $p$ | $\eta_P^2$ |
|---|---|---|---|---|
| `stimuli` | .99 | 554.25 | .00 | .99 |
| `spatiality` | .41 | 2.25 | .03 | .41 |
| `quality` | .28 | 1.24 | .29 | .28 |
| `spatiality*quality` | .22 | .94 | .51 | .22 |

not rejected for the quality and the interaction between spatiality and quality. Continuing from these results, univariate effects for the spatiality were inspected. Results are shown in Table 8.

Table 8: Univariate effects for the 1st GLM for the spatiality.

| | $p$ | $\eta_P^2$ |
|---|---|---|
| `Vergnügen` | .73 | .00 |
| `Freude` | .59 | .00 |
| `Zufriedenheit` | .90 | .00 |
| `Zuneigung` | .32 | .02 |
| `Faszination` | .64 | .00 |
| `Ergriffenheit` | .51 | .01 |
| `Interesse` | .36 | .01 |
| `Überraschung` | .73 | .00 |
| `Trauer` | .01 | .13 |
| `Verachtung` | .01 | .12 |
| `Langeweile` | .00 | .23 |

Results from Table 8 indicate significant differences for three emotions: `Trauer` ($p$ = .01), `Verachtung` ($p$ = .01) and `Langeweile` ($p$ = .00). But, as there was reliability problem with the emotion 'Trauer' and the homogeneity for the last two emotions could be questionable, these results could not be trusted.

### 3.3.2 *2nd GLM*

The 1st GLM didn't deliver trustable results. One possible reason was, that the effects of the necessary covariates (expertise, gender, age and exposure to loudspeaker) were not included. In the 2nd GLM these covariates were calculated together with the factors spatiality and quality.

*sphericity test*

MAUCHLY'S TEST OF SPHERICITY was firstly conducted. The values of the Greehouse-Geisser corrections in the Mauchly's Test of sphericity for each emotion scale are listed in Table 9:

Table 9: Mauchly's Test of Sphericity for the 2nd GLM

| Emotions scales | Greenhouse-Geisser |
|---|---|
| Vergnügen | .86 |
| Freude | .85 |
| Zufriedenheit | .81 |
| Zuneigung | .95 |
| Faszination | .92 |
| Ergriffenheit | .93 |
| Interesse | .93 |
| Überraschung | .92 |
| Trauer | .82 |
| Verachtung | .65 |
| Langeweile | .75 |

Sphericity assumption was validated for the 2nd GLM.

*homogeneity test*

LEVENE'S TEST FOR THE HOMOGENEITY was conducted and the results are shown in Table 10.

For the emotions Überraschung and Verachtung for some stimuli the null hypothesis of equal variances was rejected. This concludes that for these emotions and stimuli there is a significant difference between the variances in the subjects, and that the homogeneity could be questionable for these combinations. Unlike in the 1st GLM, the emotion scale 'Langeweile' appeared to have no homogeneity problem in the 2nd GLM.

*MANOVA*

THE MANOVA was conducted for all emotions scales. Table 11 shows the multivariate results according to Pillai's Trace for stimuli, spatiality and quality and the interaction between spatiality and quality, together with all the covariates.

*multivariate effects*

THE MULTIVARIATE TESTS for between subjects effects in Table 11 indicated that for the stimuli and spatiality, the null hypothesis that there was no significant difference could be rejected,

Table 10: Levene's Test of equality of error variances for the 2nd GLM. Values <.05 are marked red.

| Emotions scales | A | B | C | D | Mean |
|---|---|---|---|---|---|
| Vergnügen | .81 | .98 | .34 | .24 | .59 |
| Freude | .57 | .81 | .23 | .11 | .43 |
| Zufriedenheit | .77 | .64 | .50 | .57 | .62 |
| Zuneigung | .27 | .74 | .46 | .63 | .52 |
| Faszination | .21 | .63 | .46 | .15 | .36 |
| Ergriffenheit | .09 | .37 | .79 | .97 | .55 |
| Interesse | .77 | .83 | .63 | .26 | .62 |
| Überraschung | .43 | .44 | .03 | .24 | .28 |
| Trauer | .28 | .57 | .49 | .34 | .42 |
| Verachtung | .50 | .00 | .09 | .00 | .14 |
| Langeweile | .96 | .11 | .06 | .13 | .31 |

Table 11: Pillai's Trace for the multivariate effects in the MANOVA of the 2nd GLM. df = 11.

| | value | F | $p$ | $\eta_P^2$ |
|---|---|---|---|---|
| stimuli | .65 | 5.36 | .00 | .65 |
| expertise | .19 | .67 | .75 | .19 |
| gender | .36 | 1.58 | .15 | .36 |
| age | .10 | .34 | .96 | .10 |
| emitter | .26 | .99 | .47 | .26 |
| spatiality | .48 | 2.61 | .01 | .48 |
| quality | .31 | 1.29 | .27 | .31 |
| spatiality*quality | .20 | .74 | .68 | .20 |

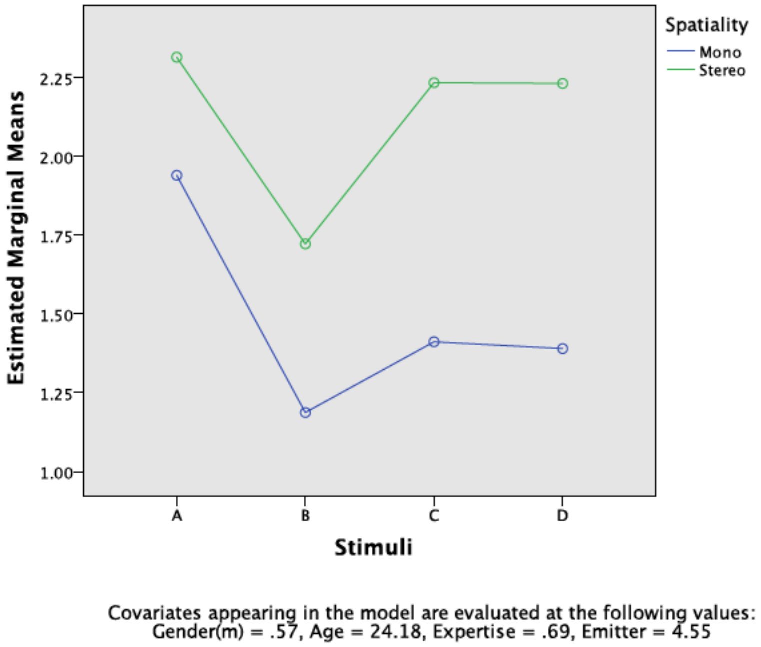
and this was not rejected for the quality, the covariates and the interaction between spatiality and quality. Continuing from these results, univariate effects for the spatiality were inspected. Results are shown in Table 12.

Again, the results from Table 12 indicate significant differences for three emotions: `Trauer` ($p$ = .01), `Verachtung` ($p$ = .00) and `Langeweile` ($p$ = .00). And this time, apart from the 'Trauer' that had reliability problem and the 'Verachtung' that had homogeneity problem, the emotion scale 'Langeweile' indicated significant between subjects effect on playback's spatiality. A visualization of this effect is shown in Figure 2.

Table 12: Univariate effects for the 2nd GLM for the spatiality.

| | $p$ | $\eta_p^2$ |
|---|---|---|
| Vergnügen | .93 | .00 |
| Freude | .77 | .00 |
| Zufriedenheit | .92 | .00 |
| Zuneigung | .29 | .02 |
| Faszination | .72 | .00 |
| Ergriffenheit | .34 | .02 |
| Interesse | .45 | .01 |
| Überraschung | .66 | .00 |
| Trauer | .01 | .13 |
| Verachtung | .00 | .15 |
| Langeweile | .00 | .32 |

Figure 2: 2nd GLM: Spatiality's between subjects effect on the emotion *Langeweile*.




This concluded that on the scale 'Langeweile', subjects who listened to the music pieces in stereo playback format felt significantly stronger expressivity. This conclusion justifies the hypothesis **a)**, that stereo playback format has positive between subjects effect on felt emotion. For the hypothesis **b)**, which argues that the CD playback quality, in contrast to the YouTube quality, has positive between subjects effect on felt emotion, although MANOVA (Table 11) didn't deliver interesting overall result ($p$ = .27 for `quality`), some significant effects could be observed in the univariate tests. The reason why the overall effect of `quality` was not significant could be due to the fact that the effects of the emotions scales had different directions (for example, positive effect was observed in the emotion scale 'Langeweile' while negative effects were observed for all the other emotions scales which had significant effects). This fact would then be further testified in the next GLM.

### 3.3.3 *3rd GLM*

By designing this GLM, following considerations have been made:

THE EMOTION SCALE 'TRAUER' had reliability problem with Stimulus C. This music piece was a "happy" one, and it might have been the reason why measuring *sadness* when listening to this song with three different sub-scales (see Table 19 on Page 57) returned unreliable results. In spite of this, this emotion scale proved to have some significant effects in the 1st and the 2nd GLM. Therefore, this emotion scale was kept in the 3rd GLM in the interest of test power.

THE INTERACTION `spatiality*quality` was removed from the model design as it didn't show any significant effect in the previous models and removing it could save some test power for more interesting effects.

THE EMOTIONS SCALES 'ÜBERRASCHUNG' AND 'VERACHTUNG' were removed in the 3rd GLM due to homogeneity problem in the previous GLMs. As the design of the 3rd GLM is quite similar to the 2nd GLM, removing them could save some test power without violating the test's validity.

Results are listed as follows:

*sphericity test*

MAUCHLY'S TEST OF SPHERICITY. The values of the Greehouse-Geisser corrections in the Mauchly's Test of sphericity for each emotion scale are listed in Table 13. Sphericity assumption was validated for the 3rd GLM.

Table 13: Mauchly's Test of Sphericity for the 3rd GLM

| Emotions scales | Greenhouse-Geisser |
|---|---|
| Vergnügen | .88 |
| Freude | .83 |
| Zufriedenheit | .80 |
| Zuneigung | .96 |
| Faszination | .92 |
| Ergriffenheit | .92 |
| Interesse | .93 |
| Trauer | .85 |
| Langeweile | .75 |

*homogeneity test*

LEVENE'S TEST FOR THE HOMOGENEITY was conducted and the results are shown in Table 14. And the assumption of homogeneity was validated for all emotions scales.

Table 14: Levene's Test of equality of error variances for the 3rd GLM.

| Emotions scales | A | B | C | D | Mean |
|---|---|---|---|---|---|
| Vergnügen | .78 | .92 | .44 | .30 | .61 |
| Freude | .57 | .93 | .49 | .11 | .52 |
| Zufriedenheit | .74 | .61 | .70 | .58 | .65 |
| Zuneigung | .33 | .74 | .47 | .67 | .55 |
| Faszination | .21 | .61 | .45 | .32 | .40 |
| Ergriffenheit | .09 | .38 | .78 | .97 | .55 |
| Interesse | .79 | .87 | .74 | .26 | .66 |
| Trauer | .31 | .82 | .49 | .35 | .49 |
| Langeweile | .98 | .13 | .09 | .13 | .33 |

*MANOVA*

THE MANOVA was conducted for all emotions scales. Table 15 shows the multivariate results according to Pillai's Trace for `stimuli`, `spatiality` and `quality`, together with all the covariates.

Table 15: Pillai's Trace for the multivariate effects in the MANOVA of the 2nd GLM. df = 11.

| | value | F | $p$ | $\eta_p^2$ |
|---|---|---|---|---|
| stimuli | .61 | 6.07 | .00 | .61 |
| expertise | .19 | .90 | .53 | .19 |
| gender | .30 | 1.66 | .13 | .30 |
| age | .09 | .37 | .93 | .09 |
| emitter | .24 | 1.21 | .31 | .24 |
| spatiality | .43 | 2.93 | .01 | .43 |
| quality | .31 | 1.74 | .11 | .31 |

*multivariate effects*

THE MULTIVARIATE TESTS for between subjects effects in Table 15 indicated that for the `stimuli` and `spatiality`, the null hypothesis that there was no significant difference could be rejected, and this was not rejected for the quality and the covariates. Continuing from these results, univariate effects for the spatiality were inspected. Results are shown in Table 16.

Table 16: Univariate effects for the 3rd GLM for the spatiality.

| | $p$ | $\eta_p^2$ |
|---|---|---|
| Vergnügen | .83 | .00 |
| Freude | .66 | .00 |
| Zufriedenheit | .78 | .00 |
| Zuneigung | .25 | .03 |
| Faszination | .71 | .00 |
| Ergriffenheit | .34 | .02 |
| Interesse | .39 | .01 |
| Trauer | .01 | .14 |
| Langeweile | .00 | .31 |

Results from Table 16 indicate significant differences for `Trauer` ($p$ = .01) and `Langeweile` ($p$ = .00). Apart from the 'Trauer' that had reliability problem, playback's spatiality indicated significant between subjects effect on the emotion scale 'Langeweile'. For a visualization of this effect, refer to the Figure 2 in the 2nd GLM, as the "direction" of the effect stayed the same: on the emotion scale 'Langeweile', playback's spatiality format had significant positive effect.

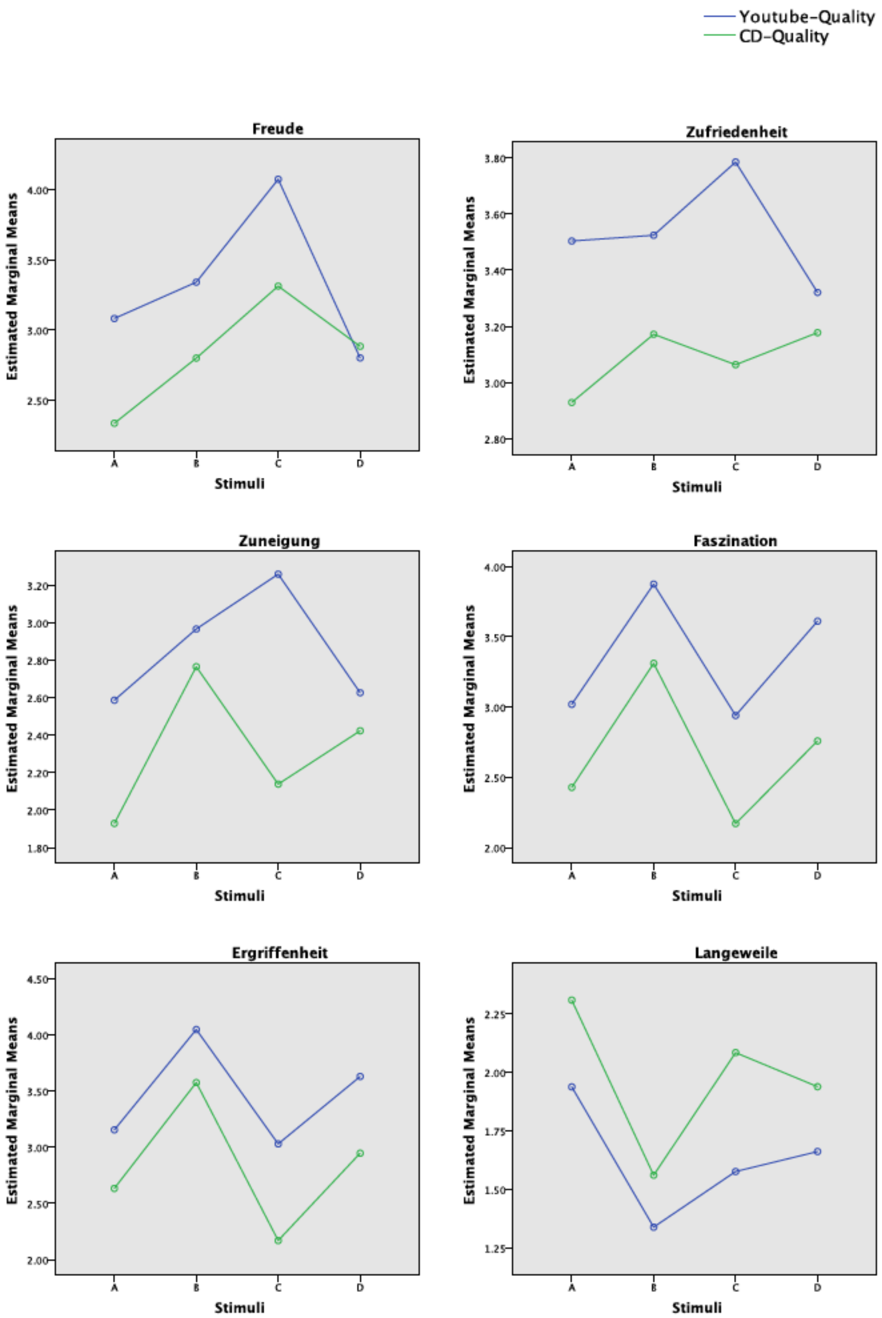
Again, this observation justified the hypothesis **a)**, that stereo playback format has positive between subjects effect on felt emotion. And for hypothesis **b)**, although MANOVA didn't deliver significant overall effect ($p = .11$), the univariate effects for the `quality` should be inspected. These effects are listed in Table 17.

Table 17: Univariate effects for the 3rd GLM for the quality.

| | $p$ | $\eta_p^2$ |
|---|---|---|
| `Vergnügen` | .08 | .07 |
| `Freude` | .00 | .19 |
| `Zufriedenheit` | .00 | .16 |
| `Zuneigung` | .01 | .14 |
| `Faszination` | .00 | .22 |
| `Ergriffenheit` | .00 | .23 |
| `Interesse` | .11 | .06 |
| `Trauer` | .67 | .00 |
| `Langeweile` | .02 | .11 |

The `quality` had significant univariate effects on six emotions scales except `Vergnügen`, `Interesse` and `Trauer`. For the emotion scale 'Langeweile', playback quality had, according to the hypothesis **b)**, positive univariate effect, which means for this emotion scale the CD quality could trigger a higher value than the YouTube quality. For the other five emotions scales, the playback quality had a negative effect. A visualization of these effect can be seen in Figure 3. Despite the result from the MANOVA, these results could still be trusted, as different 'directions' were observed. And the conclusion could be that for most of the felt emotions, playback's quality had negative effects: YouTube quality could trigger stronger emotions in most cases.

Figure 3: 3rd GLM: Quality's univariate effect estimations for the emotions scales Freude, Zufriedenheit, Zuneigung, Faszination, Ergriffenheit and Langeweile.



### 3.4 Parameter estimates of the 3rd GLM

The estimated multinomial logistic regression coefficients for the 3rd GLM were observed. As for all the emotions scales, given the assumption that while all other variables in the model are held constant:

**expertise** appeared to have a negative between subjects effect on most emotions scales. This means, subjects with musical expertise tend to feel less powerful emotions on these scales.

**age** didn't show significant between subjects effect on most emotions scales. This is probably because the age range in the subjects set was not quite wide (mean = 24.18, SD = 2.62). On the emotion scale 'Faszination' for Stimulus A ($p$ = .01) and the emotion scale 'Langeweile' for Stimulus D ($p$ = .01) age showed significant negative effects: in these situations older subjects felt less "fascinated" or "bored".

**gender** had negative between subjects effect on almost all emotions scales. In other words: male subjects tend to feel less emotions than female subjects.

**emitter** , which stands for experience with loudspeakers, had positive effects on most emotions scales which appeared to have significant interaction with this covariate. Conclusion is then: Subjects with more exposure to loudspeakers in everyday life felt stronger emotions while they listened to the music pieces with headphones in this experiment.

The hypothesis **c)** was then justified for all of its components. As a summary, three GLMs with brief details are listed in Table 18.

Table 18: Summary of three GLMs: factors, covariates, scales and effects

| **Models** | **Factors** | **Covariates** | **Scales** | **Effect** | **Note** |
|---|---|---|---|---|---|
| 1st GLM | `spatiality`<br>`quality`<br>`spatiality*quality` | - | all emotions | - | 'Trauer' (reliability problem);<br>'Verachtung' (homogeneity problem);<br>'Langeweile' (homogeneity problem). |
| 2nd GLM | `spatiality`<br>`quality`<br>`spatiality*quality` | `expertise`<br>`gender`<br>`age`<br>`emitter` | all emotions | Spatiality (+) | 'Trauer' (+) (reliability problem);<br>'Langeweile' (+);<br>'Überraschung' (homogeneity problem);<br>'Verachtung' (homogeneity problem). |
| 3rd GLM | `spatiality`<br>`quality` | `expertise`<br>`gender`<br>`age`<br>`emitter` | all except<br>`Überraschung`<br>`Verachtung` | Spatiality (+)<br>Quality (-)<br>*Expertise (-)*<br>*Gender (-)*<br>*Age (-)*<br>*Emitter (+)* | For Spatiality: 'Trauer' (+) (reliability problem),<br>'Langeweile' (+);<br>For Quality: most emotions (-), 'Langeweile' (+). |

(+) *positive effect*; (-) *negative effect*.

# 4 DISCUSSION AND OUTLOOK

## 4.1 DISCUSSION

In this experiment there were two main findings: First, the stereophonic playback format could evoke higher value of affection on the emotion scale 'Langeweile' than monophonic playback quality. Second, the playback quality used by YouTube (AAC, 128 kb/s) appeared to have better musical expressivity than CD playback quality on the majority of emotions scales. Besides these main findings, the roles of the independent subject attributes as covariates were also clarified, which are: musical expertise, gender, age and exposure to loudspeakers when listening to music in everyday life.

### 4.1.1 *Spatiality and the perception of music*

Previous researches on audio playback's spatiality have already confirmed the advantage of multi-channel playback, whether these researches focused on the spatiality quality itself [49, 30] or on the felt emotions when listening to music [31, 9]. The spatiality quality, according to Toole's definition in [49], consists of "sound images", "continuity" and "width" of the "sound stage", "impression of distance" and "abnormal effects". When focused on felt emotions, these attributes, however, are also limited by the evoking ability of the musical excerpts. In this study, significant between subjects effect of spatiality was found on the emotions scales 'Trauer' and 'Langeweile', which are both negative emotions. This could be concluded that the "more spatial" playback quality acts like a kind of catalyst which enhances felt emotions, whether these are positive or negative emotions.

The result, that the playback spatiality did not show significant between subjects effect on most (positive) emotions, could be explained by the fact that the majority of recruited subjects in this experiment listened to music in everyday life with quite "mixed" manners. Figure 4 shows how the subjects listened to music by representing their answers to the following questions (these questions were asked and answered after the listening test, and the last two questions were already mentioned in the methodology chapter, of which the last one was the question for the covariate `emitter`):

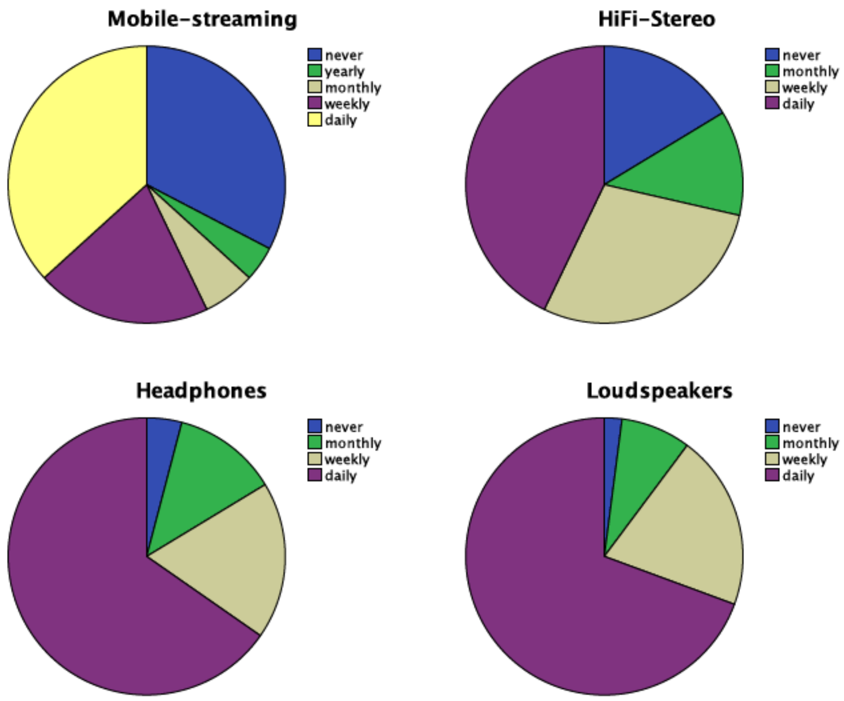
**HOW OFTEN** do you listen to music per mobil-streaming device?

**HOW OFTEN** do you listen to music per domestic HiFi-stereo device?

**HOW OFTEN** do you listen to music per headphones?

**HOW OFTEN** do you listen to music per loudspeakers?[1]

Figure 4: Subjects' frequency of listening behavior: per mobile-streaming or HiFi-stereo; per headphones or loudspeakers.




The two lower pie diagrams show strong similarity in the most frequency usage of headphones and loudspeakers. In fact, the dataset shows that 22 subjects (44.9%) had answered "daily" to both of the last two questions. By asking these questions, the definition of "loudspeakers" was not clarified. It could be the loudspeakers of a home theater system, or the loudspeakers of a laptop, or even the loudspeaker of a mobile phone; Nowadays single-bodied "capsule" loudspeakers connected wirelessly via bluetooth are also quite popular in younger generation, which although could be equipped with multiple speakers inside but in comparison to the size of the listening room, the sound they produce appear to be a "point sound

1 In Figure 4 the legend of the first pie diagram is different to those of the other pies, because in the 2nd, 3rd and 4th questions no subject had selected "yearly" as an answer. Therefore there are five elements in the first pie diagram while there are only four in the other three pie diagrams.

source" and rather mono than stereo. The "sound images", "continuity" and "width" of the "sound stage", as well as "impression of distance" conveyed by these diverse systems could be very different, nevertheless the room or environment to use these loudspeakers could be various in size and acoustical attributes. Adding the headphones as another source, these subjects are believed to be statistically quite "well adjusted" to different spatial playback environments and could focus better on the music itself in the daily changing maneuver[2]. In these subjects, it could be quite possible that playback's spatiality doesn't have any significant effect on felt emotions at all.

Two of the four musical excerpts in this study were also used in the earlier explorative study conducted by the author and a colleague [9]. The procedure, equipment and loudness, as well as lab environment of these two studies are quite similar to each other. The fact, that the effects of spatiality found in the study in 2011 was more significant than that in the study in 2015 could also be explained by the fact that during these years the fast development of playback devices has raised various listening experiences in everyday life among the younger generation, resulting them to be better "well adjusted" to different spatial playback qualities today. However, some subjects, who listened to ML versions in this study, commented like "the playback quality was too poor and I think that had affected my enjoyment". As already mentioned in the introduction chapter (second paragraph of 1.2 on Page 4), in the study in 2011 no subject talked about the playback quality, even given the fact that half of that sample set consisted of students of the Audio Communication Group who are believed to be experts of audio playback. These phenomena reveal altogether the fact that the young people today, as compared to 2011, are not only better adjustable to different playback qualities, but also more aware of the playback quality itself. In fact, these facts say one simple truth: the young people today have better "musical expertise" than a few years ago.

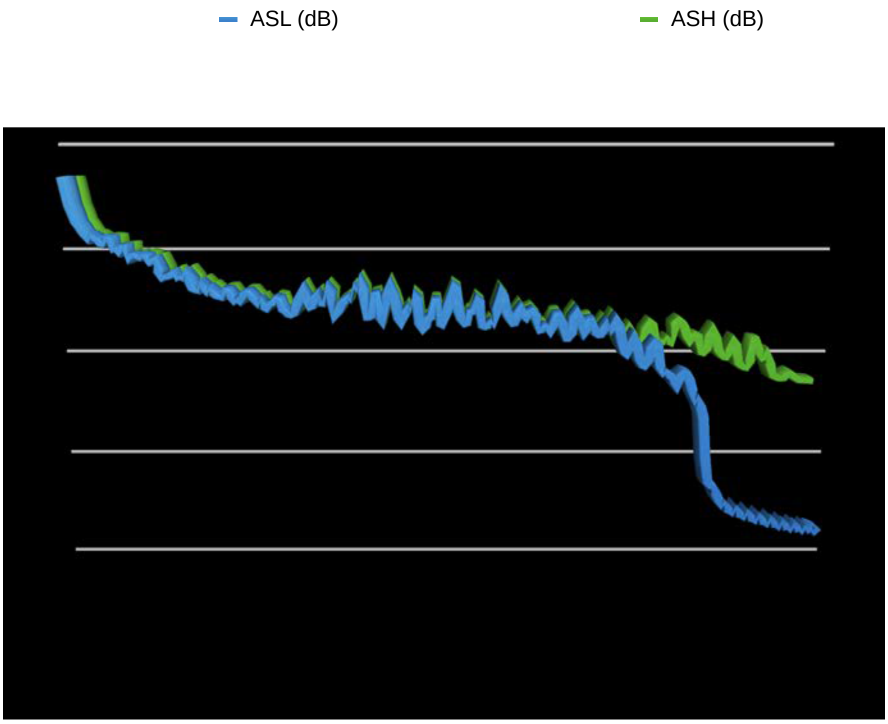
### 4.1.2 *Why streaming quality beats lossless CD quality*

Continued from the above discussion on spatiality: although the subjects' answer "daily" to the first two questions which are represented by the two upper pie diagrams in Figure 4 also seem quite similar to each other, only 8 subjects (16.3%) had selected

2 As a subject asked the author one day after the experiment about the real purpose of the experiment, the author told him he listened to the mono versions; then he expressed his strong feelings for the musical excerpts and was quite sure the mono playback didn't affect his felt emotions, as he "could focus on the music itself, and enjoyed it very much".

"daily" to both questions, according to the dataset. At least this fact means the subjects' listening behavior of the streaming/HiFi variant is not so "mixed" as the headphones/loudspeakers variant. In other words, compared to the various usage of devices, the subjects seem to "stick to" the same music source in everyday life: either streamed music, or local music (e.g. CD). This could explain why they are statistically not so "well adjusted" between different playback qualities, as quality's effect on quite a few emotions scales, which were mostly positive emotions (see Figure 3), was observed in this experiment.

Toole [49] defined sound quality as consisted of "clarity", "softness", "fullness", "brightness", "pleasantness" and "fidelity". These subjective attributes were also evaluated by the author and a fellow sound engineer in a blind test with the same stimuli in SH and in SL, with the above components as scales. Both the author and the sound engineer could surely tell the difference of the versions, and the key to manage that was on the scale "softness". The SL version sounded obviously "softer" and "warmer", while on the other scales the difference between them was quite inaudible. Taking the first stimulus for example, the frequency distribution was analyzed with Audacity and shown in Figure 5. On this 3-Dimensional figure, with both distributions in overlay displayed, one can clearly see that the difference between both versions in the frequency domain lies above approximately 16kHz. As already discussed in the introduction chapter, although these frequencies lie beyond human audible range, they could still make a difference in music playback [34, 35, 36]. In this case, the YouTube quality, which lack of higher frequencies as compared to the CD quality, sounded "softer" and "warmer".




69% of all subjects had "musical expertise" and in this sense are believed to be quite aware of the sound quality, but they were not asked to evaluate the sound quality like in Toole's experiment - they were asked to report their felt emotions. They concentrated on their own perceptions during the test, and the M-DAS score represented how they enjoyed the *sound* [11]. And the result was, they evaluated the stimuli in YouTube quality as better enjoyable than the CD quality. As the author seldom listens to CD in the recent years, a subjective impression is that the CD quality sounded "sharp in higher frequencies" and it was somehow even "distracting". Although technically, the CD quality reproduces a better linearity in frequency domain, the point of listening to music is for the enjoyment. That the listening aesthetic has been changed, is presumably the *learning effect* [40].

Although live music performance is believed to be one of the most "buoyant part of the music economy" [50], music from YouTube or similar sources is free. The younger generation today, as represented by the subjects of this experiment, use mediated music in everyday life and a live orchestral performance could be a new experience for them [51]. The majority of the musical excerpts used in this experiment are actual orchestral recordings, in which higher pitch noises such as the noise of strings of violins, strumming of guitar, paper noises when turning the pages of sheet music etc. are quite actively audible in the CD quality versions, while these are not so "obvious" in the YouTube quality versions, thanks to the "high-cut" attribute of the sound quality. On one hand, in line with Kolb's theory in "You Call This Fun? Reactions of Young First-time Attendees to a Classical Concert" [51], the reproduction of such a live performance could also affect the enjoyment of music in younger generations who seldom visit live orchestral concerts. On the other hand, younger listeners are believed to be more sensitive for higher frequencies than older listener [32], thus the high pitch noises in a precisely reproduced live performance could be even more distracting for them.

Therefore, the conclusion of this finding could be that the subjects are more used to the YouTube quality or similar qualities, as music playback in these qualities are more familiar to them, and could meet their expectations for a "better sound". The hypothesis **b)** was not justified.

### 4.1.3 *Roles of musical expertise, gender, age and emitter usage*

*on expertise, gender and age*

The hypothesis **c)** was justified for all of its components. The fact, that the majority of randomly recruited subjects in this study were with musical expertise [11] as defined in this study, could be related to the technical background of this university campus and the nature of this experiment itself: when advertised as a "hearing experiment", people who enjoy "listening" would be more attracted to attend. As neurological [21] studies have shown physical evidence of musicianship, empirical studies [14, 12] found out trained listeners are more accurate on realizing their felt emotions. This could explain why the subjects with musical expertise in this study appeared to be more "emotionally steady" when listening to music. Same goes for the age and gender, as older subjects are likely more experienced in feeling emotions they were less "fascinated" or "bored", and as female subjects could feel more powerful emotions when listening to music than male subjects.

*understanding on emitter*

The experience in everyday life also applies for the covariate `emitter`, which stands for the exposure to loudspeakers, as it stands for how "fresh" the headphones experience could be for the subjects. But as already discussed with Figure 4, almost half of all subjects (44.9%) use both headphones and loudspeakers every day, and for these subjects, the headphone is not quite a new experience. The positive effect of `emitter` could be explained by how "serious" they listen to music in everyday life. People listen to music when cooking, washing, studying, sporting etc. This could be the reason why they are exposed to music with both loudspeakers and headphones on the same day [47]. In this sense, "music from loudspeakers" could be heard in a supermarket, in a restaurant, in an airport etc. and this listening is not serious appreciation. Although music in public areas also function as trigger of emotions [52, 53] the listeners do not listen seriously as in the experiment, where they concentrate on their felt emotions: And for *them*, listening to music through headphones or even through loudspeakers and concentrating on their felt emotions is a relatively new experience.

## 4.2 CONCLUSION AND OUTLOOK

### 4.2.1 *Personality, social connotation and "sound preference"*

*music preference*

As the development of technology enables the mediatization of music as product today, the younger generation *learn* [40] the "media grammar" [54] as they were born into this mediatized environment. In light of the research of Rentfrow et al [46] on the correlation between personality and music preference, people develop their music preferences with unconscious consideration on aspects of the social connotation of the music products. Young people listen to certain music types not only because of the music itself, but also because other young people are listening to them. In this sense music not only communicates emotion *within* people, but also works as a sort of social identification *between* people.

*"sound" preference*

Continued from this discussion of preference for music types (whether distinguishing music conventionally per genres, or in sense of how music make listener feel [47][3]), another dimension of music preference is yet to be uncovered: the playback format. Here, playback format stands for the combination of all technical circumstances, as *sound* [11], especially of spatiality and quality. Questions like "Is playback format also used as reference of social identification as well as music genre preferences in the media grammar of the younger generation today? For instance, do CDs only belong to "certain" people? Can we be true friends, if you use high-end playback devices and I use YouTube? Do people use HiFi equipment for pure musical enjoyment or for self distinguishing from the crowd?" are yet to be answered.

3 Refer to the footnote on Page 11

### 4.2.2 *Criticism and methodological outlook*

*measurement levels*

The subjects could only hear the excerpts once, before they finish each M-DAS questionnaire. Some of them suggested that it would be nice to hear it again by filling out the questionnaire as music was too short, compared to the time to finish the 33 questions. One subject wrote *"I almost forgot how I felt after answering some questions"*. However, these suggestions were not taken, as measurement level should be equal on each subject.

*precise procedures*

Although it was clearly instructed before the experiment (see APPENDIX 2: PRINTED INSTRUCTION IN THE LAB) that they should start answering questions *after* listening, some subjects already started answering halfway in the stimuli. The author made such assumption because the time taken for these subjects to finish each M-DAS was extremely short: almost as long as the excerpts[4] (one subject even finished as music was still playing, and the author reminded him to read the instruction again); it was impossible that they finish the M-DAS in seconds. However, it is still not clear if this could cause distortion to the results. Anyway, in future tests this instruction has to be orally emphasized.

*loudness*

One or two subjects have complained about the playback volume. Although each subject had heard the excerpts with the same loudness, it could be true that for certain people the average volume is too loud. Unfortunately the volume could not be adjusted for each subject, as loudness is believed to be a dominant part of the music perception [55, 56]. So how to calibrate the equivalent loudness for each subject is a question yet to be explored.

*listening behavior*

The definition of "music per streaming" or "music per HiFi" was not clear enough. A handful of subjects have asked about the definition of these terms before answering them. These subjects should have then understood the questions correctly. How did the other subjects silently interpret these questions, is unknown.

*sensation seeking*

In future experiments, it would be interesting to inspect subjects' *openness to new experiences* as well. On one hand, since the music pieces as stimuli were supposed to be unfamiliar to all subjects, it was a new experience for all of them to hear these music pieces; and this concerns the overall experience of the experiment. On the other hand, concerning the music itself, even if the subjects listen to familiar music pieces, their personality trait also plays a role in the

4 The author was aware of the time taken for each M-DAS finishing because in the experimental procedure (2.6, on Page 18), music was manually played upon request of the subject after each measurement.

perception. Continued from the argument that headphones would be a new experience for some subjects who used loudspeakers more often, another dimension that described each subject's openness to new experience as calibrating covariate could be introduced, and is called Sensation Seeking [57]: *"the need for varied, novel and complex sensations and experiences and the willingness to take physical and social risks for the sake of such experience."* According to this theory high sensation seekers tend to have a high level of stimulation [58]. The findings of Litle et al [59] (based on music genres) indicated that high sensation seekers "tolerate and like high intensity and/or complexity in music". The introduction of Sensation Seeking should help to better understand the subjects' enjoyment of music.

# Part III

# BIBLIOGRAPHY AND APPENDIX

[1] Li, Chongguang (Beijing, 1962): *Basic Music Theory*, vol. 10. People's Music Press of China.

[2] Ziegenrücker, Wieland (Leipzig, 1997): *ABC Musik: allgemeine Musiklehre ; 446 Lehr- und Lernsätze*. Breitkopf und Härtel.

[3] Small, C. (2011): *Musicking: The Meanings of Performing and Listening*. Music Culture. Wesleyan University Press. URL https://books.google.de/books?id=7vS8yQwvuGcC.

[4] Huizinga, J. (Boston, 1949): *Homo Ludens:*. Homo Ludens: A Study of Play-element in Culture. Routledge & K. Paul. URL https://books.google.nl/books?id=ALeXRMGU1CsC.

[5] Juslin, Patrik N and Daniel Västfjäll (2008): "Emotional responses to music: The need to consider underlying mechanisms." In: *Behavioral and brain sciences*, **31**(05), pp. 559–575.

[6] Huron, David Brian (Cambridge and London, 2006): *Sweet anticipation: Music and the psychology of expectation*. MIT Press.

[7] Lepa, Steffen and Anne-Kathrin Hoklas (2015): "How Do People Really Listen to Music Today? Conventionalities and Major Turnovers in German Audio Repertoires." In: *Information, Communication and Society*, pp. online–first. doi:10.1080/1369118X.2015.1037327. URL http://www.tandfonline.com/doi/full/10.1080/1369118X.2015.1037327.

[8] Vickers, Earl (2011): "The Loudness War: Do Louder, Hypercompressed Recordings Sell Better?" In: *J. Audio Eng. Soc*, **59**(5), pp. 346–351. URL http://www.aes.org/e-lib/browse.cfm?elib=15934.

[9] Pischel, Claudia and Meng Wu (2011): "Playback format's impact on evoked emotions in music reception: mono vs. stereo." In: *Unpublished term paper*.

[10] Vogt, Nick (2015): "YouTube Audio Quality Bitrate Used For 360p, 480p, 720p, 1080p, 1440p, 2160p." In: *h3xed*. URL http://www.h3xed.com/web-and-internet/youtube-audio-quality-bitrate-240p-360p-480p-720p-1080p.

[11] Lepa, Steffen; Anne-Kathrin Hoklas; Hauke Egermann; and Stefan Weinzierl (2015): "Sound, materiality and embodiment challenges for the concept of 'musical expertise'in the age of

digital mediatization." In: *Convergence: The International Journal of Research into New Media Technologies*.

[12] Castro, São Luís and César F Lima (2014): "Age and Musical Expertise Influence Emotion Recognition in Music." In: *Music Perception: An Interdisciplinary Journal*, **32**(2), pp. 125–142.

[13] Pons, Francisco; Joanne Lawson; Paul L Harris; and Marc De Rosnay (2003): "Individual differences in children's emotion understanding: Effects of age and language." In: *Scandinavian Journal of Psychology*, **44**(4), pp. 347–353.

[14] Robazza, Claudio; Cristina Macaluso; and Valentina D'urso (1994): "Emotional reactions to music by gender, age, and expertise." In: *Perceptual and Motor skills*, **79**(2), pp. 939–944.

[15] Laukka, Petri and Patrik N Juslin (2007): "Similar patterns of age-related differences in emotion recognition from speech and music." In: *Motivation and Emotion*, **31**(3), pp. 182–191.

[16] Evers, Stefan; Jörn Dannert; Daniel Rödding; Günther Rötter; and E-Bernd Ringelstein (1999): "The cerebral haemodynamics of music perception." In: *Brain*, **122**(1), pp. 75–85.

[17] Kölsch, Stefan; et al. (2003): "Children processing music: electric brain responses reveal musical competence and gender differences." In: *Journal of Cognitive Neuroscience*, **15**(5), pp. 683–693.

[18] Kölsch, Stefan; Burkhard Maess; Tobias Grossmann; and Angela D Friederici (2003): "Electric brain responses reveal gender differences in music processing." In: *NeuroReport*, **14**(5), pp. 709–713.

[19] LeBlanc, Albert; Young Chang Jin; Lelouda Stamou; and Jan McCrary (1999): "Effect of age, country, and gender on music listening preferences." In: *Bulletin of the Council for Research in Music Education*, pp. 72–76.

[20] Karageorghis, Costas I; Kevin M Drew; and Peter C Terry (1996): "Effects of pretest stimulative and sedative music on grip strength." In: *Perceptual and motor skills*, **83**(3f), pp. 1347–1352.

[21] Lee, Dennis J; Yi Chen; and Gottfried Schlaug (2003): "Corpus callosum: musician and gender effects." In: *Neuroreport*, **14**(2), pp. 205–209.

[22] Sabatinelli, Dean; Tobias Flaisch; Margaret M Bradley; Jeffrey R Fitzsimmons; and Peter J Lang (2004): "Affective picture perception: gender differences in visual cortex?" In: *Neuroreport*, **15**(7), pp. 1109–1112.

[23] Kallinen, Kari and Niklas Ravaja (2007): "Comparing speakers versus headphones in listening to news from a computer–individual differences and psychophysiological responses." In: *Computers in Human Behavior*, **23**(1), pp. 303–317.

[24] Nelson, Thomas M and Thomy H Nilsson (1990): "Comparing headphone and speaker effects on simulated driving." In: *Accident Analysis & Prevention*, **22**(6), pp. 523–529.

[25] Cupchik, Gerald C (2011): "The Role of Feeling in the Entertainment= Emotion Formula." In: *Journal of Media Psychology*, **23**(1), pp. 6–11.

[26] Oliver, Mary Beth and Anne Bartsch (2010): "Appreciation as audience response: Exploring entertainment gratifications beyond hedonism." In: *Human Communication Research*, **36**(1), pp. 53–81.

[27] Oliver, Mary Beth and Anne Bartsch (2011): "Appreciation of Entertainment." In: *Journal of media psychology*, **23**(1), pp. 29–33.

[28] Scherer, Klaus R (2004): "Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them?" In: *Journal of new music research*, **33**(3), pp. 239–251.

[29] Zentner, Marcel; Didier Grandjean; and Klaus R Scherer (2008): "Emotions evoked by the sound of music: characterization, classification, and measurement." In: *Emotion*, **8**(4), p. 494.

[30] Morgan, Barbara J. and Ogden R. Lindsley (1966): "Operant Preference for Stereophonic over Monophonic Music." In: *Journal of Music Therapy*, **3**(4), pp. 135–143. doi:10.1093/jmt/3.4.135. URL http://jmt.oxfordjournals.org/content/3/4/135.abstract.

[31] Lepa, Steffen; Stefan Weinzierl; Hans-Joachim Maempel; and Elena Ungeheuer (2014): "Emotional Impact of Different Forms of Spatialization in Everyday Mediatized Music Listening: Placebo or Technology Effects?" In: *Audio Engineering Society Convention 136*.

[32] Möser, Michael (2012): *Technische Akustik*. Springer-Verlag Berlin Heidelberg.

[33] Weinzierl, Stefan (Berlin, Heidelberg, 2008): *Handbuch der Audiotechnik*. Springer Science & Business Media.

[34] Yoshikawa, Shokichiro; et al. (1995): "Sound-quality evaluation of 96-kHz sampling digital audio." In: *Audio Engineering Society Convention 99*.

[35] Ando, Akio; Kimio Hamasaki; Toshiyuki Nisiguchi; and Kazuho Ono (2004): "Perceptual Discrimination of Very High Frequency Components in Musical Sound Recorded with a Newly Developed Wide Frequency Range Microphone." In: *Audio Engineering Society Convention 117*.

[36] Pras, Amandine and Catherine Guastavino (2010): "Sampling Rate Discrimination: 44.1 kHz vs. 88.2 kHz." In: *Audio Engineering Society Convention 128*.

[37] Pras, Amandine; Rachel Zimmerman; Daniel Levitin; and Catherine Guastavino (2009): "Subjective evaluation of mp3 compression for different musical genres." In: *Audio Engineering Society Convention 127*.

[38] Ruzanski, Evan P (2006): "Effects of MP3 encoding on the sounds of music." In: *Potentials, IEEE*, **25**(2), pp. 43–45.

[39] Olive, Sean (2011): "Some New Evidence That Teenagers May Prefer Accurate Sound Reproduction." In: *Audio Engineering Society Convention 131*.

[40] Kirk, R. (1956): "Learning, a major factor influencing preferences for high-fidelity reproducing systems." In: *Audio, IRE Transactions on*, **AU-4**(5), pp. 133–136. doi:10.1109/TAU.1956.1165648.

[41] Plambeck, Joseph (2010): "In Mobile Age, Sound Quality Steps Back." In: *New York Times*. URL http://www.nytimes.com/2010/05/10/business/media/10audio.html.

[42] Dougherty, Dale (2009): "The Sizzling Sound of Music." In: *O'Reilly Radar*. URL http://radar.oreilly.com/2009/03/the-sizzling-sound-of-music.html.

[43] Capps, Robert (2009): "The Good Enough Revolution: When Cheap and Simple Is Just Fine." In: *Wired Magazine*. URL http://archive.wired.com/gadgets/miscellaneous/magazine/17-09/ff goodenough?currentPage=all.

[44] Edgecliffe-Johnson, Andrew (2010): "Could a pair of headphones save the music business?" In: *FT Magazine*. URL http://www.ft.com/cms/s/2/5e37eb34-74e0-11df-aed7-00144feabdc0.html.

[45] Music, NPR (2009): "The Loudness Wars: Why Music Sounds Worse." In: *the decade in music: 'oos*. URL http://www.npr.org/2009/12/31/122114058/the-loudness-wars-why-music-sounds-worse.

[46] Rentfrow, Peter J and Samuel D Gosling (2003): "The do re mi's of everyday life: the structure and personality correlates of music preferences." In: *Journal of personality and social psychology*, **84**(6), p. 1236.

[47] Rentfrow, Peter J; Lewis R Goldberg; and Daniel J Levitin (2011): "The structure of musical preferences: a five-factor model." In: *Journal of personality and social psychology*, **100**(6), p. 1139.

[48] Renaud, Dagmar and Dagmar Unz (2006): "Die M-DAS - eine modifizierte Version der Differentiellen Affekt Skala zur Erfassung von Emotionen bei der Mediennutzung." In: *Zeitschrift für Medienpsychologie*, **18**(2), pp. 70–75.

[49] Toole, Floyd E (1983): "Subjective measurements of loudspeakers: A comparison of stereo and mono listening." In: *Audio Engineering Society Convention 74*. Audio Engineering Society.

[50] Frith, Simon (2007): "Live music matters." In: *Scottish music review*, **1**(1).

[51] Kolb, Bonita M (2000): "You call this fun?" In: *Reactions of young, first-time attendees to a classical concert. In: Weissman D (ed.) Music Industry Issues and Studies*, **1**(1), pp. 13–28.

[52] Milliman, Ronald E (1982): "Using background music to affect the behavior of supermarket shoppers." In: *The journal of Marketing*, pp. 86–91.

[53] Milliman, Ronald E (1986): "The influence of background music on the behavior of restaurant patrons." In: *Journal of consumer research*, pp. 286–289.

[54] Gumpert, Gary and Robert Cathcart (1985): "Media grammars, generations, and media gaps." In: *Critical Studies in Media Communication*, **2**(1), pp. 23–35.

[55] Schubert, Emery (2004): "Modeling perceived emotion with continuous musical features." In: *Music perception*, **21**(4), pp. 561–585.

[56] Kellaris, James J; Susan Powell Mantel; and Moses B Altsech (1996): "Decibel, Disposition, and Duration: The Impact of Musical Loudness and Internal States on Time Perceptions." In: *Advances in Consumer Research*, **23**, pp. 498–503.

[57] Zuckerman, Marvin (2014): *Sensation Seeking (Psychology Revivals): Beyond the Optimal Level of Arousal*. Psychology Revivals. Taylor & Francis London and New York.

[58] Zuckerman, Marvin (1984): "Sensation seeking: A comparative approach to a human trait." In: *Behavioral and brain sciences*, **7**(03), pp. 413–434.

[59] Litle, Patrick and Marvin Zuckerman (1986): "Sensation seeking and music preferences." In: *Personality and individual differences*, **7**(4), pp. 575–578.

# A

# APPENDIX 1: M-DAS

## MODIFIZIERTE DIFFERENZIELLE AFFEKTE SKALA

- The M-DAS used in this experiment was a combination of 11 selected scales according to *Tabelle 2* in [48] and each scale has three sub-scales;
- Each subject fills out four M-DAS questionnaires with 33 sub-scales and the questions were randomized in each measurement.
- Following is a list of these 11 scales and their 33 sub-scales:

Table 19: M-DAS scales and sub-scales.

| Scales | Sub-scales |
|---:|---|
| Vergnügen | amüsiert - erheitert - vergnügt |
| Freude | fröhlich - glücklich - Freude |
| Zufriedenheit | ausgeglichen - wohl - zufrieden |
| Zuneigung | liebevoll hingezogen - verliebt - Zuneigung |
| Faszination | beeindruckt - fasziniert - gebannt |
| Ergriffenheit | bewegt - ergriffen - gebannt |
| Interesse | aufmerksam - konzentriert - wach |
| Überraschung | Überraschung - erstaunt - verblüfft |
| Trauer | niedergeschlagen - entmutigt - Trauer |
| Verachtung | Verachtung - Geringschätzung - Spott |
| Langeweile | unbeteiligt - Langeweile - angeödet |

# B

# APPENDIX 2: PRINTED INSTRUCTION IN THE LAB

Every subject was asked to read this instruction carefully before the start:

**Liebe/r Versuchsteilnehmer/in,**

während des folgenden Experiments werden Dir vier Musikstücke vorgespielt.

Nach jedem dieser Musikstücke wirst Du gebeten, zu vorgegebenen emotionsbeschreibenden Begriffen anzugeben, wie stark Du diese Gefühle beim Hören der Musik empfunden hast.

Bitte gib hierbei wirklich nur die Emotionen an, die Du **tatsächlich empfunden** hast, und nicht die Emotionen, die das Musikstück vermitteln soll.

Im Anschluss werden noch einige allgemeine Informationen abgefragt. Falls Du hierzu Fragen hast, wende Dich einfach an den Versuchsleiter.

Vielen Dank, dass Du an meinem Experiment teilnimmst!

Meng

Meng Wu
*on media playback format's role for the emotional impact and the appreciation of music in everyday life*
a master's thesis