Perceptual Mapping of Attributes and Perceived Dissimilarities in High-Gain Guitar Amplifiers

Morten Purup Andersen

Master's Thesis

AALBORG UNIVERSITY

Aalborg University 2011

AALBORG UNIVERSITY

School of Information and Communication Technology
Department of Electronic Systems.
Engineering Psychology.
Frederik Bajers vej 7
Phone: +45 99 40 86 00
E-mail: webinfo@es.aau.dk
Web: http://sict.aau.dk

Title:
Perceptual Mapping of Attributes and Perceived Dissimilarities in High-Gain Guitar Amplifiers

Theme:
Master's thesis
Engineering Psychology

Project period:
10^{th} semester
February 1^{st} - May 31^{st} 2011

Project group:
PDP1072

Author and Candidate:
Morten Purup Andersen

Supervision:
Flemming Christensen,
Associate Professor at Acoustics
Aalborg University, Denmark

Physical Copies: 4
Number of pages: 116
Finished May 31^{st}, 2011

Abstract:

Immediately perceiving the differences between a series of products within the same category can be a difficult task. For customers buying the product, the only option seem to ask for help by more experienced customers or experiencing the products one by one. For a development team developing products within a narrow category, the team members may not share a common understanding of exactly what makes these products different and most certainly do they not share a common vocabulary for describing the characteristics differentiating these products. A perceptual mapping can provide a common understanding of a product series and has relatively easily been made by conducting listening experiments asking the listeners for dissimilarities and attributes within the range of products under investigation. Seven high-gain guitar amplifiers were recorded through a controlled signal chain only changing the specific amplifier. Multidimensional scaling of dissimilarity data gathered from 16 participants provided a two-dimensional stimulus plot and individual attribute ratings of each amplifier aided the interpretations of the dimensions. Perceptual mapping is suggested as a helpful tool for e.g. development teams developing products possessing complex perceptual attributes.

Preface

This report is written with the goal of studying the overall challenge of using perceptual mapping as a helpful tool for e.g. development teams within the field of audio and music equipment. Throughout the report this challenge is emphasised, described, and discussed. However, as a concrete example of perceptual mapping, a listening experiment is conducted and and the Conclusion in Chapter 5 solely concludes on this listening experiment and the circumstances under which it is conducted. For a rounding of the problem formulation, the reader is instead refered to the Discussion in Chapter 6.

This report is my master's thesis in support of my degree in Engineering Psychology at Aalborg University, Denmark. For as long as I remember I have been into music, psychology, and usability. Combining these three sciences in this thesis will increase user understanding of auditive characteristical attributes of different high-gain guitar amplifiers. As an amplifier customer, user, and musician; I too have found it hard to differentiate the auditive characteristics of the many products sold within the field of audio. This is my approach to increase the systematisation of auditive perception of high gain guitar amplifiers.

A lot of people have helped me directly or indirectly with this project. I would like to thank Anders Ruby for great discussions and assistance - especially during the recording of the program material, Claus Vestergaard and Peter Dissing for technical and practical support, Teodor Georgiev and 4Sound for borrowing me guitar amps, Mathias Jensen for the Gibson Les Paul, Christian Sejer Pedersen for introducing me to psychophysical methods and scaling, Christian Andersen and Johan Trettvik for bouncing off ideas with me, Poul Svante Eriksen and Florian Wickelmeier for statistical support, Søren Bech for introducing me to great litterature on MDS, Bo Stilling for developing GUIs for the tests. Also, thanks to Anthony Price, Esben Skovenborg, Lars Arknæs, and Tore Mogensen.

Department of Electronic Systems,
Aalborg University

Date: May 31st 2011

Morten Purup Andersen

Contents

Chapter 1

Introduction

1.1 Introduction

The sound of an overdriven electric guitar has appealed to many people through music for more than half a century. A major industry has grown out of this appeal in order to create overdriven guitar sounds and implementing them in music. New musical genres have demanded the need for a more overdriven guitar sound and the most overdriven guitar sounds are now created from high-gain guitar amplification. The perceived sound of a high-gain guitar amplifier is affected by many parameters which raise various complex sensations in the listener. The many parameters have an influence on the overall quality and liking of an amplifier and it is therefore of interest to evaluate these subjectively. For an inexperienced customer it might seem like an overwhelming challenge to understand what differentiates one high-gain guitar amplifier from the vast range of similar high-gain guitar amplifiers available on the market. If he has only heard one or two amplifiers, he would have no chance knowing how different these are compared to the rest of the available amplifiers? Furthermore, an even more interesting and difficult question could be in exactly what way they sound different - and how would this be described to him?

For audio professionals - or a development team - developing new and innovative audio products several issues are raised: How is it certain that everybody in the team has a common understanding of a certain description or denotation of a given sound characteristics? How do the team members describe sounds that are not necessarily unidimensional? What if a team member could just point in a geometrical space and the rest of the team would know exacly what he means?

Also, if having a common understanding of the competitors' audio products, how can they make sure that their product prototype does indeed differentiate itself from the characteristics of competing products?

If this knowledge is not readily available in the mind of the customer or development team, it should be available to them in another manner. This leads to the question: "*How can information containing a number of music professionals' perception of high-gain guitar amplifiers* *be mediated to other people who do not posses the same experience and knowledge?*"

The following sections will adress the problems related to discriminating different high-gain guitar amplifiers. After a quick definition of high-gain guitar amplification, section 1.3 adresses the issues in remembering characteristics of auditory stimuli followed by a section on how these characteristics and differences in auditory stimuli can be mediated through visualisation. The sections hereafter (1.6, 1.7, and 1.8) focus on the perception of these visualised mappings and the intuitive aspect of mediating information this way. Before a discussion of this chapter leading to the problem formulation, existing research on perceptual mapping of guitar amplifiers are outlined.

1.2 The sound of high-gain guitar amplification

A high-gain guitar amplifier is a an amplifier designed to amplify the signal from the guitar and add a vast amount of nonlinear distortion to the signal. Audio examples of a high-gain amplified guitar is heard within the first 10 seconds of songs by e.g. Silverchair [1997], Black Sabbath [1971], and Pantera [1992]. The sound of high-gain guitar amplification is often utilised in genres such as heavy metal, rock, hardcore, punk, and pop - to name a few.

1.3 Auditory memory interference

Textbooks on perceptual audio evaluation have typically not emphasised the cognitive and perceptual aspects of the listener in an ecological context and see the human cognitive factors as e.g. distinct variables such as expertise and familiarity [Kortum, 2008] and filters [Bech and Zacharov, 2006] including the listener's expectations, emotional state, previous experience with the type of stimuli. However, these models has proven sufficient in many instances [Bech and Zacharov, 2006]. Combining these sensory impressions with the listener's cognition forms the degree of liking or annoyance of the sound (affective measurements and hedonic tests) [Bech and Zacharov, 2006]. Figure 8.6 in the Appendix illustrates the so-called filter model. This is a view often taken within fields of engineering bordering perception. Within textbooks of cognitive psychology models of perceptual- and cognitive systems are also widespread, e.g. the construction-integration model suggesting that human inference processing involves the two succesive stages: Generation (i.e. production of all possible inferences) and integration (i.e. the usage of the most appropriate inference) [Eysenck and Keane, 2005]. Critiques of such views are seen in [Trettvik, 2001; Gibson, 1986] i.a. for neglecting activity as a factor in perception. An important psychological aspect of remembering sound is *interference*. Interference is defined as "*the distortion or disruption of memory which happens as a result of other information being learned or already stored in memory*" and has been found to be most likely to occur with information that is similar [Eysenck and Keane, 2005]. Products within a narrow range such as high-gain guitar amplifiers can be hard to distinguish due to their perceptual similarity. It must be assumed that even music professionals' and experienced guitarists' memory of the auditive characteristics are distorted in a given degree. Several years between exposure of certain high-gain guitar amplifier sounds are likely to occur. Other phenomenons such as the *Ranschburg effect* where poorer human memory performance is observed in sequential stimuli lists containing repeatet stimuli than sequential stimuli lists containing no repeatings [Mewaldt and Hinrichs, 1973]. Recency effects, interference, similarity-related issues, and poor memory performance is also a common issue in experiments with auditive stimuli, as proven by e.g. [Ueda, 2004; Deutsch, n.d.; 1972; 1975; Turner et al., 2001]. Furthermore physical factors are likely to influence the perception of auditive stimuli. Such factors could include room acoustics, the guitar on which the music is played, sound pressure level, distance from listener to speaker, timbre, and amount of gain.

Due to poor human memory of auditory stimuli including characteristics of high-gain guitar amplification, the information regarding these characteristics must be made readily available. In other words, the information not readily available must be mediated.

1.4 Mediating experience

Bærentsen [2000] argues that mediating experiences using technology has a long history:

> *"Sculptures, paintings, and other devices have been utilized for recording, writing, calculation and transmission of information for thousands of years. Binoculars, telescopes, and microscopes were invented hundreds of years ago. The telegraph, telephone, gramophone, radio and television are more recent inventions. Today, computers with complex programmed functions reside in lots of products used to mediate such activities."* [Bærentsen, 2000]

This leads to the question:
"*How can information containing a number of guitarists' - and music professionals' - perceived dissimilarity of high-gain guitar amplifiers be mediated to other people?*"
To present information in an *intuitive*[1] manner that is not the act of playing all guitar amplifiers sequentially or simultaniously, introducing a spatial representation of the perceptual attributes of the relevant high-gain guitar amplifiers is proposed.

1.5 Models and Maps

Exploring entire environments such as cities requires a great deal of effort and many people find themselves using maps to find their location of interest. Asking for directions can also serve as a valueable source of communicating information. In short; a map is regarded as a visual display of spatial information [Newscombe and Huttenlocher, 2000]. In psychology there are generally three views on understanding maps: Children show little interest in maps untill elementary school (Piagetians), second; Map understanding is innate (nativists), and finally; Map understanding is a transmission of invented and socially shared symbol systems (Vygotskyans) [Newscombe and Huttenlocher, 2000]. The most significant way in which these three views differ is the processing of input. Piagetians, i.e. researchers following

[1] Definitions of intuition and intuitive interfaces are described in [Bærentsen, 2000]

the first view, see input as stimuli to facilitate the growth of the child whereas nativists consider input as the catalyst triggering biologically determined programs [Newscombe and Huttenlocher, 2000] . Finally, Vygotskyans "*have popularrized the metaphor of the scaffold that supports and instrumentally aids in development*" [Newscombe and Huttenlocher, 2000].

1.6 An Ecological Approach to Perceptual Learning and Development

Instead of viewing the world egocentrically, Gibson and Pick [2000] proposed that children also have to learn to view the world as "allocentric", i.e. "*an objective, unbiased view of the layout of the world as opposed to an egocentric one*" [Gibson and Pick, 2000]. The way an object looks from where person A is positioned does not necessarily equal the way the same object looks from the view of person B. Viewing the world as allocentric does correlate with Gibson [1986]'s ecological approach to visual perception in which he introduced the term "invariants" [Gibson, 1986].

1.7 Modalities of Perception

Perceiving the world, perceiving perspectives, and spatial relations are developed from birth and visual perception is actually a prerequisite to some phenomenons in language [Gibson and Pick, 2000]. According to Gibson and Pick [2000] inter- and multimodal perception of structured stimuli is the foundation of future perception in infants. This indicates that children are born with the possibility of perceiving multimodally and later learns to seperate perceptions into modalities, i.e. we learn to differentiate stimuli instead of integrating stimuli. Multimodal perception is therefore supramodal to unimodal perception. If combining this with the statements of Stern [1995], emotion is supramodal to multimodal perception [Stern, 1995] - which again is supramodal to unimodal perception [Gibson and Pick, 2000]. Ekman [1954], who's data is often refered to in MDS litterature, was also convinced that emotion played a significant role in the degree of perceived similarity between stimuli. This theory is written in section 8.12 in the Appendix. Emotion also plays a role in perception according to the theory of *physiognomic* perception proposed by Heinz Werner in 1948 suggesting another form of amodal perception [Stern, 1995] where amodal qualities are seen as categorial emotions rather than perceptual qualities such as e.g. size, intensity, and amount. The physiognomic perception is a kind of amodal perception due to emotional experiences not being related to specific perceptual modalities [Stern, 1995].

Perceptual-emotional qualities in music have been studied by Wedin [1972] who argued that if music is capable of communicating quasi-emotional meanings, emotions in general will also be recognized in the perception of music. He found out that descriptive adjectives can be elicitated by having subjects describe music. Hereby, emotionally coloured adjectives such as "Tension/Energy", "Gaiety-Gloom" and "Solemnity-Triviality" can act as the end-points of semantic differential scales [Wedin, 1972]. It is also argued that the most important aspect of emotional states may be described in terms of power, strength, and tension (i.e. level of activation from strong to weak, tough to tender, tense to relaxed, active to passive). Furthermore, adjectives can also be more strongly associated with affective evaluation of music, which is seen in descriptors such as "pleasantness", "mirth", "self-assertion", "fear", and "disgust" [Wedin, 1972]. Understanding correlations between physical antecedents and their perceptual reaction can be a difficult empirical task [Wedin, 1972]. However, strong interactions between emotional qualities and physical attributes have been studied and found most relevant within e.g. tempo, pitch, loudness, harmony, rhythm, and major-minor modes [Wedin, 1972]. Slow tempos have proven effective in eliciting dignified, calm, and serene descriptors of music. Fast tempos arouse moods more associated with excitation and happiness. High pitches lead to descriptors of a sprightly humorous character while low pitch reactions correlate more with sad, majestic, and serious quialities [Wedin, 1972].

Research within the scope of translation - or transposition - of modalities has shown that children can recognise objects visually that they have only interacted with haptically and tactile via mouthing [Gibson and Pick, 2000; Stern, 1995]. Within the perception of auditive and visual modalities shape and form are probably easier to depict through other modalities than the auditive but Stern [1995] argues that due to the multimodal perception of watching lips moving while hearing words the child's understanding is increased compared to unimodal perception of *either* speech or lips moving. On behalf of this argumentation and the dominant position of the visual modality's within perception, it is reasonable to assume that auditive differences and auditive attributes can be depicted visually. Depicting these visually will presumably occur in a more conceptual form and therefore on a higher cognitive level than immediately perceiving the auditive differences directly.

1.8 Perception of depictions, perspectives, and pictures

The reality of pictures, such as multidimensional plots, does not provide the viewer with a realistic sense of presence. Instead it puts the viewer into the scene, as Gibson [1986] describes:

> *"(...) if a picture displays the perspective of a scene it puts the viewer into the scene, but that is all. It does not enhance the reality of the scene."* [Gibson, 1986]

The viewer of the perceptual stimulus space is put into the scene when viewing the perceptual mapping, but his reality of the stimuli is not being enhanced. In other words: He does not perceive the auditive differences between stimuli although being wiser as to how they are correlated.

Language plays a huge role in the mediation of experience and memory. Gibson [1986] wrote the following criticism as to whether perspective is a language:

> *"But the essence of a picture is just that its information is not explicit. The invariants cannot be put into words or symbols. The depiction captures an*

awareness without describing it. The record has not been forced into predications and propositions. There is no way of describing the awareness of being in the environment at a certain place. Novelists attempt it, of course, but they cannot put you in the picture in anything like the painter can."

The viewer observing the stimulus space derived by multidimensional scaling has no reference of self or proprioceptive perception and the picture is therefore on a conceptual level [Gibson, 1974]. Gibson [1958] argues for the distinction between experiencing on first hand and experiencing on second hand, i.e. *becoming aware* of something and *being made aware* of something:

> *"A picture is a record of what its creator has seen or imagined, made available for others to see or imagine. Depicting should be distinguished from the decoration, ornamenting, embellishing, or beautifying of a surface considered as such. The problems of aesthetics exist in their own right (...) Let us not confuse the kind of information that has been put into words with the kind that has been simply displayed. Film is not a language with a grammar, as some film makers like to believe. A graphic depiction is not an explicit description and, similarly, a motion picture is not a verbal narration."*

However, even though Gibson [1974] differentiates between perceiving and conceiving, these classifications are not categorically different but instead of different cognitive levels [Gibson, 1974]. This corresponds to the theory of levels of cognition by Velichkovsky [1990], stating that cognitive skills developed during ontogenesis, e.g. orientation in the near environment, are on a lower cognitive level than e.g. a "*... higher-order form of spatial representation ...*" such as "*... the representation of imaginary environments as they are built, for example, according to literary descriptions.*" [Velichkovsky, 1990]

1.8.1 Childrens' depiction of perceived invariants

Several ecological psychologists within perception and cognition argue for a heterarchy of higher and lower levels of cognition [Velichkovsky, 1990; Gibson, 1974; Bærentsen, 2000; Nørager, 2009]. The visual perception psychologist, Gibson [1986], states that:

> *"The young child learning to draw ... I suggest, he depicts the invariants that he has learned to notice ... He may not yet draw in edge perspective because he has not noticed it. Hence, he may draw a table with a rectangle top and four legs at the corners because those are the invariant features of the table he has noticed. This is a better explanation than saying he draws what he knows about the table, his concept, instead of what he sees of the table, his sensation. The fatal flaw of the latter explanation is that it ought to be the other way around. The child should begin by drawing sensations and progress to drawing concepts."*

A child can be seen as a subject viewing a picture in a more intuitive manner than an adult due to the child not yet having developed its higher cognitive levels such as language, abstract representation, and conceptual thinking. When Gibson [1986] argues that the child draws the four table legs at the corners of the table, it can be reasoned that the table legs' interspatial relation is understood by the child. A reasonable assumption is thus that the child will also perceive points in a stimulus space as being further apart when the distances between these points increase. Another way of stating this could be that the child has an understanding of the spatial relations between invariants. At the stage of preschool, children begin to develop map competences and by the age of 3 years, they show understanding of abstract representations such as "X marks the sport" [Newscombe and Huttenlocher, 2000]. A few months after the age of three years, abilities of distance scaling across simple spatial representations are vivid [Newscombe and Huttenlocher, 2000]. A child may not perceive the distances in a multidimensional plot as representing differences between the stimuli but will still be able to perceive distances between points in the plot as opposed to perceiving - and *understanding* - two guitarists verbalising their degree of perceived difference between two stimuli.

1.9 Existing Reseach on perceptual mapping of guitar amplifiers

Atsushi and Martens [2001]

Atsushi and Martens [2001] argue that specialised knowledge is possesed by experienced sound production experts and musicians. These music professionals have an understanding of the perceptual results available in guitar signal processing effects. Furthermore they posses a specialised vocabulary for describing the effects but not necessarily for describing the perceptual attributes resulting from manipulating the control parameters (i.e. "knob tweaking") [Atsushi and Martens, 2001].

Atsushi and Martens [2001] conducted and analysed experiments in hope of gaining knowledge of a user-centered control structure that presents guitar effects *"in a manner accesible to a wide range of users"* by basing "*a GUI upon a mapping between perceptual dimensions of representative guitar effects and semantic differential scales used to describe these effects*" [Atsushi and Martens, 2001]. Inconcistency and potential confusion was observed in the operation of the "drive" parameter in the operation of the MIDI-controllable Boss GX-700 Guitar Effects Processor when manipulating the gain parameter in the nominal guitar effects: `Overdrive`, `Fuzz`, and `Distortion`. The inconsistency arised due to the "drive" setting only changing the loudness of the `Distortion` output while changing both loudness *and* either timbre or overdrive quality for `Fuzz` and `Overdrive`, respectively.

Analysing results of dissimilarity data and semantic differential results showed that only the scales "Wildness" and "Heaviness" gave relatively clean separation of stimuli over five levels of "drive" and the three nominal categories `Overdrive`, `Fuzz`, and `Distortion`. Atsushi and Martens [2001] interprets the following from their analyses of their results:

> "*As "drive" level is increased from 0 to an intermediate value, the perceived quality of the effect gradually becomes what can be described as "wilder". If "drive" level is further increased above this intermediate value, "Wildness" appears to*

saturate somewhat for all three nominal effect types. In contrast, "Heaviness" ratings seem to be saturated below this intermediate "drive" level, and increase to their maximum at the maximum level of the "drive" parameter. This is true for all three nominal effect types. Thefore, it is hypothesized that the full range of the drive parameter, initial increases first turn the effects on, making the output "wilder," and subsequent increases only make the output "heavier." Perhaps such a two-limb control function existed as a natural feature of the analog distortion effects emulated by modern digital multi-effects processors. Hard clipping would of course yield such behavior, since progressively boosting the input signal only begins to produce a significant change in the output as the signal peak begins to exceed the point at which the clipping occurs, the output timbre becoming "wilder". Then, as the amount of clipping grows, the increasing distortion of the input signal produces a progressively "heavier" output timbre." [Atsushi and Martens, 2001]

Martens and Marui [2002]

Martens and Marui [2002] performed a study of the timbral attributes of three characteristic amplifiers simulated by the GM-200 "Guitar Amp Modeler" [Martens and Marui, 2002]. They argue that tone coloration in no way is a unidimensional problem although easier defined for steady sounds with no spectral evolution. Furthermore they argue that tone coloration can be described by a lower dimensional structure than timbre and its definition (tone coloration) can be narrowed by excluding i.a. timbral differences as a result of the degree of inharmonic content in a sound such as *attack inharmonicity*. For the three amplifier models studied (*British Crunch*, *Twin Drive*, and *Combo 335*) Martens and Marui [2002] found that in a 3-dimensional stimulus space Dimension 3 separated the three three amplifier models in three discrete layers, while Dimension 1 and 2 were interpreted as primarily representing the tone coloration especially "thickness". Martens and Marui [2002] furthermore discussed the basis in their results for maintaining a separation between perceptual attributes derived by dissimilarity scaling and the verbal attributes related to them due to the former probably being valid across languages while the latter not being readily transferable across language and culture. However, the semantic differential analysis should not be rejected, though, due to the method aiding the interpretation of the dimensions derived by multidimensional scaling.

1.10 Discussion

Gaining knowledge of the complex perceptual attributes resulting from playing a guitar through a high-gain guitar amplifier is a difficult and cumbersome task. There is an unlimited number of variables affecting the perceived sound resulting from manipulating with guitars, amplifiers, speakers, rooms, and microphones to name a few. Even though research in the area of mapping these perceptual attributes has been conducted [Atsushi and Martens, 2001; Martens and Marui, 2002], this research has focused on describing perceptual attributes resulting from digital modelling technology. No known research has yet been conducted on perceptual attributes and differences in all tube-driven high-gain guitar amplifiers. Furthermore, descriptors of perceptual attributes found by [Martens and Marui, 2002; Atsushi and Martens, 2001] were constructed in Japanese and translated to English. The generation of program material by Atsushi and Martens [2001]; Martens and Marui [2002] were "*a combination of three nominal distortion-based effects and five "drive" levels*". Scientific research within the perceptual domain of high-gain guitar amplification is still very much virgin territory.

1.11 Problem Formulation

Due to the complex nature of auditory perception of high-gain amplification a low-dimensional representation of otherwise complex data can help prove a useful and helpful tool for users, customers, development teams, and music professionals. It could provide an increment towards a common understanding of the perceptual aspects of high-gain guitar amplifications. Representative high-gain guitar amplification settings should be the offset for generating the stimuli used as program material. Furthermore, perceptual attributes should be generated in English and not be translated from e.g. Danish or Japanese due to potential translational limitations. Acquiring dissimilarity ratings, direct ratings of perceptual attributes as well as preference measures from listeners will aid the generation of a graphical representation of the dissimilarities within the perceptual domain of the amplifiers. According to Martens and Marui [2002] "*the current assumption is that auditory events having multiple perceptual attributes have a mental structure that can be quantitatively captured in terms of a multi-dimensional perceptual space (or similarity structure) that is distinct from the words that might be used to describe the stimulus percepts occupying that space*". Such multidimensional perceptual space of a group of participants' perception of a range of high-gain guitar amplifiers will be sought generated.

Chapter 2

Methods

An experiment were designed in order to acquire dissimilarity data, attribute ratings, and preference measures of seven all tube-driven high-gain guitar amplifiers. The experiment consisted of two sessions. The independent variable of the experiment, the high-gain amplifiers, had to be manipulated to a representative setting for each amplifier hence an online survey was designed and implemented in order to have respondents answer which of the said guitar amplifiers they owned followed by a picture of the selected amplifier(s) and a user interface where the respondents could indicate their representative setting of the amplifier's control panel. A screen shot of the online survey is seen in 8.6 in Appendix. Knowledge of the representative settings for each amplifier were used to manipulate the controls for the independent variable. Before conducting listening experiments all stimuli were subjectively calibrated for loudness by 13 graduate students (see section 8.13 in the Appendix). Audiometric tests for hearing impairments were performed for all participants according to the procedure described in section 8.10 in the Appendix. None of the participants in the loudness calibration participated in the actual listening experiment.

Figure 2.1: *Photograph of the physical setup in the acoustics laboratory used for all listening tests.*

Figure 2.2: *Photograph of the setup for the listening experiment. The participant shown in the photograph participated in the loudness calibration but the setup was identical during the listening experiments.*

2.1 The Listening Experiment

2.1.1 Session 1

In session 1 the participant was first tested for hearing impairments. This was conducted in a small cabin with a double-door construction and had isololating floor, walls, and ceiling. The results were not revealed to the participant before the debriefing at the end of session 2 due to the risk of bias due to emotions and mood, which is proven to affect cognitive factors and thereby listeners rating of stimuli [Zielinski et al., 2007]. No participants were excluded from participating in the test regardless of hearing impairments. The participants were then interviewed to collect meta data such as age, instrument experience, and whether they had been exposed to loud sounds within 24 hours.

Hereafter, the participants were exposed to all stimuli in order to familiarise them with the range of stimuli, thereby indirectly encouraging them to also use the two labels at the ends of the scale. This should eliminate contraction bias [Zielinski et al., 2007]. After the familiarisation of the stimuli the participants were asked to give their affective rating of each stimuli on a scale. The participants were instructed that they should not *necessarily* assign seven stars (highest) to the most prefered amplifier.

The listeners were also asked to make personal constructs naming stimuli attributes according to the repertory grid technique (RGT). Advantages of the RGT is that it avoids asking participants directly for opposite semantic descriptors. Whether these personal constructs would opposite in meaning implicitly as a result of the RGT task can be discussed although it has been assumed by Choisel and Wickelmaier [2005]. Triads of randomised program material were presented for the participant; one triad at a time. The participant were asked to specify which amplifier differed from the other two, why it differed, and finally describe what the two remaining stimuli had in common. This procedure was repeated untill the test participants ran out of new words of perceptual attributes or began repeating themselves. As seen in the following raw data of personal constructs generated by participant 16, it is clear that he repeats himself after a while:

> *"more defined treble sound. less blanket over the speakers. less bass. it sounds like there is a blanket over the speakers. they lack definition and precence. b has more tone and less distortion. furthermore b is charicterized by having are more pronounced mid tone than the other two. more distortion than b. less midtone. more bass. a because of a is acceptable because of a clearer sound than c. c does not sound very good at all. it has much more high freq. going on. actually to a state that it is very harsh to listen to for an extended period of time. it sounds very scoped out. a lot of treble, some bass and allmost no midtone. i don't like it very much. a and b play a lot of bass and they lack definition to the tone. not enough treble and midtone i my opinion. it is the blanket/effect over again. it sounds like there is a blanket in front of the speakers and the sound does not enter the room properly. more defined almost harsh treble sound. not very pleasing to the ears. a lot of bass, and not very much besides that. a very dull and boring sound. c is slightly better*

and has a more open sound than a though. more treble, more precence, less bass and midtone less treble and precence than c. more bass... especially b. a sounds better than b. more midtone. less distortion. less bass. more tone! to much bass. not enough definition. not enough treble and midtone. more midtone, less treble, less bass. not enough midtone. too much bass or treble. it has more mid tone, less distortion and is more enjoyable to listen to. a lot of bass and/ or treble and too little midtone."

All exposure to the program material were conducted in a 60 m^2 multichannel listening room according to the ITU-R BS775-1 recommendation. The room was designed for subjective evaluation of sound. This may not provide the most ecologically valid results (as given in e.g. a typical living room or in this case - a rehearsal space) but evidence proves this scientific solution to have the advantage of higher sensitivity and accuracy [Zielinski et al., 2007]. Between the two sessions, the attributes elicited in session 1 were written on a bipolar VAS paper template

2.1.2 Session 2

In session two the same participants came back and initially were presented to the stimuli again before giving their affective ratings for the experimenter to check for response consistency. Hereafter the participants were asked to rate the dissimilarity for each pairwise combination of the range of program material presented in a randomised order to account for e.g. recency effects. The label "dissimilarity" of scale points instead of "similarity" due to the possibility of the term "similarity" affecting the results. Dissimilarity can not always be regarded as the inverse of similarity [Wickelmaier, 2003]. It was assumed that the difference between stimulus A and B is equal to the difference between stimulus B and A. This assumption is not proven but the randomisation of the stimuli pairs compensates for such effects Schiffman et al. [2007]. No repeated measurements were included as well as no same-same dissimilarity ratings were included. The dissimilarities are used for indirectly scaling the stimuli by generating a multidimensional scaling solution. For a thorough description of direct and indirect scaling in perceptual audio evaluation, please refer to section 8.3 in the Appendix. Before a debriefing the participants gave their attribute ratings for each of the attributes elicited by the RGT in session 1.

As in session 1, all exposure to the program material were conducted in a 60 m^2 multichannel listening room according to the ITU-R BS775-1 recommendation.

2.1.3 Playback of the program material

The program material was played for the participants through a pair of Genelec® 1031A studio monitors. This way all visual cues of the amplifier through which the stimuli were played were eliminated. Such visual cues are known to cause significant bias toward e.g. a significant brand [Zielinski et al., 2007]. The setting could in some way be resemble the situation in which a recording musician in a studio control room listens to the same guitar track recorded with two different microphones in order to compare the sound quality of these.

2.2 Recording of stimuli

Sound is generated by the condensation and rarefaction of molecules transmitting the original vibration of the object causing the sound. Due to reflections and refractions influencing the sound as it is moving through the surrounding medium, the sound "image" reaching the ear differs from the sound originally generated [Moore, 2003]. In order to vary only the amplifier of the different stimuli, the electrical line signal of a Gibson® Les Paul Standard was recorded through a preamp/mixer and a sound card into a computer. Hereafter, the guitar signal was played sequentially through each of seven high-gain guitar amplifiers through a 4x12" guitar cabinet and recorded with a Shure® SM57 instrument microphone. However, when scrutinising the sound originally generated by a guitar signal, played through an amplifier and a cabinet, recorded with a microphone converted into a digital signal and played back to a test participant, there is more to it than the influence of reflections and refractions mentioned above.

The recording of the high-gain amplifiers were conducted in controlled environments, i.e. a specially designed multichannel acoustics laboratory, in order to eliminate as much reverberation as possible. The heavily damped walls and ceiling in the 60 m^2 large laboratory causes a reverberation time less than 0.2 seconds. The room is well insulated from outside noise and the specially designed windows insulate towards the outside. A photograph of the laboratory is seen in Figure 2.3.

Figure 2.3: *Photograph of the specially designed acoustics laboratory.*

The signals from - and to - the multichannel acoustics laboratory could be monitered and controlled in a control room outside the laboratory. A photograph of the control room is seen in Figure 2.4. The whole recording session was facilitated by the experimenter and a professional recording engineer. A thorough interview with the recording engineer can be read in section 8.5 in the Appendix.

Figure 2.4: *Photograph of the control room where the facilitator could connect to the microphones and amplifiers in the laboratory and record the signals. The most significant hardware in the signal chain positioned in the control room is labelled.*

2.2.1 Settings of the Amplifiers

The settings of the channels on the high-gain guitar amplifiers were collected from an online survey asking guitarists questions such as which amplifier they own. The primary purpose of the web-based survey was to uncover the average high-gain settings representative of the use of each high-gain guitar amplifier. More information regarding the online survey is read in section 8.6 in the Appendix. In the survey the respondents answered how they had their settings on their own amps on a 10-point scale for each turning knob and radio buttons for on/off buttons. 192 respondents from the age 13 to 58 (average = 27, SD = 9.5) took part in the survey. The average values of the respondents' answers were used when setting the controls on the amplifiers before recording in order to reduce experimental bias and select representative recordings as recommended by Zielinski et al. [2007]. The settings were also listened to from within the recording room by the professional producer and the experimenter before accepting the specific setting for the recording. It could be argued that by accounting for amount of e.g. distortion, treble, and bass for each high-gain guitar amplifier would reduce the complexity of tone coloration and timbre and thus the stimulus space would show more subtle differences between the guitar amplifiers. It was decided, however, to compare the representative use of each high-gain guitar amplifier.

2.2.2 Materials

Hardware

- Microphone: Shure® SM57
- Sound Card: RME DIGI96 / 8 PST
- Behringer EURORACK MX802A preamp/mixer
- Fujitsu Siemens PC
- Marshall 1960A angled 300 W 4x12" cabinet with Celestion G12T 75 W units
- Monacor® SM-4 Sound Level Meter
- Voltmeter

Software

- Adobe® Audition® 3, version 3.0.1

Guitar

- Gibson® Les Paul Standard. Tuning: (lo-hi) D A D G B E. The so-called drop-D tuning is widely utilised within the musical genres utilising high-gain guitar amplification due to the ease of playing power chords combining the root and the fifth of the chord using only one finger on the fretboard.

Figure 2.5: *Photograph of the Gibson Les Paul Standard electric guitar on which the recorded guitar riff was played*

The video “VideoGibsonMeter.mov” on the Appendix DVD shows the voltage amplitude for each of the three guitars. The vide were recorded for documentation purposes as well as adjusting the output of the sound card. The objective was to make the A/D and D/A conversion in the sound card transparent, i.e. to synthesise the exact signal originating from the electric guitars when reamping the line signal.

2.2.3 The guitar line signal

In 2.6 the signal chain for recording the guitar line signal are illustrated.

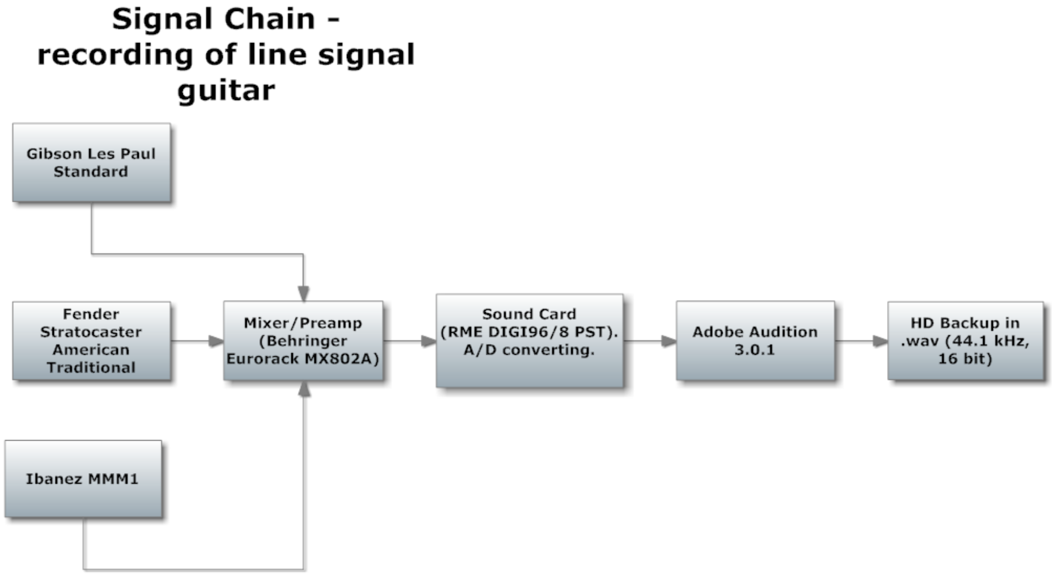

Figure 2.6: *Signal chain of the recording of the guitar line signal. Three guitars were recorded: Fender® Stratocaster American Traditional, Gibson® Les Paul Standard and Ibanez® MMM1 (Mike Mushok Signature Model)*

.

The musical figure resembling a generic rhythmic heavy-rock riff was played by an experienced guitarist on the guitar. The electrical signal was then recorded through a signal chain, which is illustrated in 2.7. When having recorded the line signal onto the PC, it showed that the Gibson Les Paul Standard was phase inverted, i.e. its peaks are downwards. This, however, does not have an impact on the further processing or future experiments.

2.2.4 Reamping the signal and recording the stimuli

The guitar line signal was played on the PC and sequentially sent to each amplifier through the line input of the amplifier. The - at the time of the recording - specific amplifier was positioned on top of the cabinet in order to simulate a real ecological environment. The output signal from the amplifier was then sent to the input of the 4x12" cabinet (8 Ω load for all amplifiers). The signal chain is illustrated in Figure 2.7. In order to acquire an approximate measure of the settings of each amplifier (e.g. tone, gain, presence) in an ecological environment (i.e. the real usage of the amps by real guitarists in real use scenarios), an online sruvey was designed and programmed in html. The survey is elaborated in Appendix 8.6. 134 respondents answered that they did not own any of the amplifiers recorded for the experiment. Only 34 respondents owned one or more of the specific amplifiers. The mean setting of each amplifier was calculated and used as an anchor point of the settings whereafter the professional recording engineer and the experimenter manipulated the settings in finer details for their joint subjective optimum sound. The exact settings for each amp is seen in Appendix 8.7.

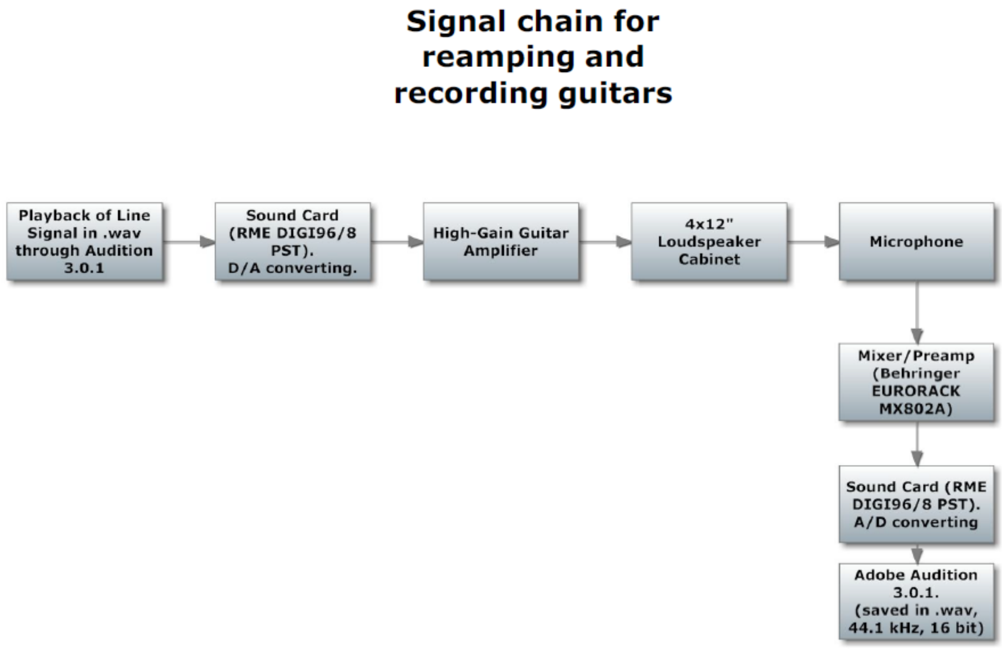

Figure 2.7: *Flowchart of the signal chain for reamping the guitar line signal and recording the sound.*

Three microphones were tried for recording the sound from the cabinet in the acoustics laboratory: Shure® SM57, Shure® SM7B, and a Sennheiser MD421. After listening to the recordings of each microphone, it was decided to use only the Shure® SM57 due to its status in the recording industry and the fact that it was the preferred microphone under these circumstances. The microphone was pointed to the centre of the loudspeaker cone of the top right loudspeaker in the cabinet. After listening to different angling positions of the microphone, the angling seen in Figure 2.8 and Figure 2.9 was chosen.

Figure 2.8: *Photograph of the approximate 40°-45° off-axis angling of the Shure® SM57 microphone. The photograph was taken from above.*

Figure 2.9: *Photograph of the approximate 40°-45° off-axis angling of the Shure® SM57 microphone. The photograph was taken from the side.*

Table 8.3 and 8.4 in Appendix illustrate the systematic variation of the amplifiers for each recorded .wav-file. Dithering was used in exporting the audio files from Adobe® Audition® 3.

2.2.5 Participants

Participants in the experimental session should be representative of people using high-gain guitar amplifiers - including but not limited to guitarists. Guitarists are the primary user of high-gain guitar amplifiers but also music professionals (e.g. studio engineers, record producers, guitar technicians, etc.) are defined to be representative of the population using high-gain guitar amplifiers. It is argued that non-guitarist musicians playing with guitarists using high-gain guitar amplifiers are also valid participants. Listening to music recordings containing high-gain guitar amplification on an almost daily basis is also argued to increase distinguishing sounds within this domain. Using engaged guitarists - or other music professionals - will reduce SQH since a low tendency to engage in cognitive thinking increases the risk of respondents using SQH. Furthermore, engaged test participants should facilitate discrimination among scale points and discimination of related scale-items [Weathers et al., 2005]. This will increase the chances of recollecting relevant information to each scale item. Expert listeners were used in the listening experiments due to novice and experter listeners' perceptual differences of music [Gromko, 1993; Novello et al., 2006].

2.3 Response scales

In order to reduce variance in response data caused by measuring inaccuracy, many respondents are asked to respond on a x-point scale such as e.g. a Likert scale representing a range of possible opinions. Research is ambiguous regarding whether the number of scale points affect scale reliability [Weathers et al., 2005]. Typically a 7- or 9-point scale is used for dissimilarity ratings [Wickelmaier, 2003]. It is desired to avoid the respondents using status quo heuristics (SQH) in their dissimilarity rating of the auditive stimuli, i.e. respondents selecting the response category that was selected in the previous response category. Weathers et al. [2005] argue that the number of scale points should not affect the use of SQH if the respondents are following a pattern shifting between SQH and making an effort giving their reponse due to them feeling a need to provide consistent answers across a group of related items. Research found that SQH usage increases with the complexity of choice environment. Weathers et al. [2005] found support of SQH being proportional to the number of scale points. Other factors, such as e.g. experiment topic, time pressure, and the tendency to engage in thinking may affect respondents' need to distinguish between response alternatives [Weathers et al., 2005]. Furthermore respondents' processing capacity should be proportional with discrimination of related scale-items and among scale points, recollecting relevant information to each scale item [Weathers et al., 2005]. Unfortunately, no clear recommendations of the number of scale points when designing a rating scale are present [Weathers et al., 2005]. Finally, processing capacity should be inversely proportional with SQH. A review of biases in modern audio quality listening tests is found in [Zielinski et al., 2007] where also biases regarding number of scale points are discussed.

Response scales used in the listening experiment were as follows. More information and illustrations are seen in section 8.14 in Appendix.

- Preference Rating: 7-point scale using stars [$\star$] and no anchor points.
- Dissimilarity Rating: 7-point unipolar scale using only end points.
- Attribute Rating: Bipolar visual analogue scale (VAS) with open ends and end points only.

A 7-point dissimilarity scale due to the wish for a rather high resolution of response rating without the risk of increasing SQH. Furthermore, the number of scale points alone only has a small impact on reliability [Weathers et al., 2005]. Zisapel et al. [n.d.] compared the VAS scale to a categorical five-grade scale and found that two scales concordant and highly correlated and concluded that the two scales were comparable in terms of capturing changes in sleep quality. It seems safe to say that there are many views on scale use, scale points, validity, continuity, and scale discrimination.

Chapter 3

Hypotheses

3.1 Exploratory Study of Dissimilarity Data

The hypotheses proposed in this chapter are generally reached through plotting metric MDS solutions for each individual participant as well as observing individual biplots of the first two principal components of attribute ratings. These plots are seen in section 8.15 in the Appendix, where the proportion of explained variance by the two principal components as well as the amount of stress are stated for the biplot and MDS solutions, respectively. The metric MDS procedure was performed due to the nonmetric MDS procedure proved itself too constrictive; eliminating both subtle - and significant - differences in the stimulus plot.

All individual stimulus plots generated from dissimilarity data were physically printed on transparent A4 sheets. This aided an exploratory study where grouping the program material was of interest. The tangible sheets allowed rotation and mirroring of plots. Examples of this study are seen in Figure 3.1 and Figure 3.2.

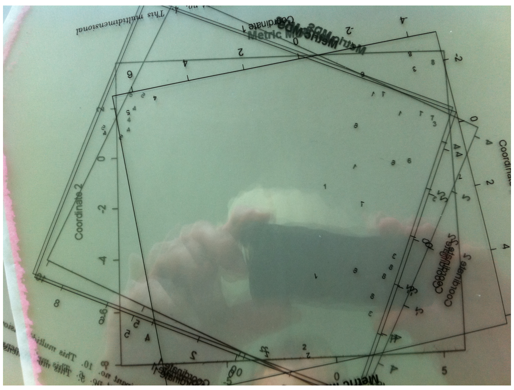

Figure 3.1: *Photo of one group of transparent A4 sheets on top of each other. Each plot is generated by individual data for the dissimilarity rating task. Each plot are rotated and in some instances mirrored.*

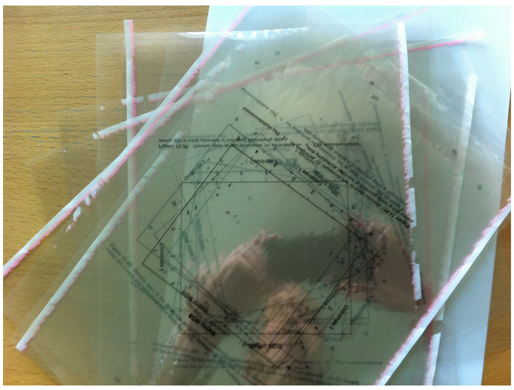

Figure 3.2: *Photo of another group of transparent A4 sheets on top of each other. Each plot is generated by individual data for the dissimilarity rating task. Each plot are rotated and in some instances mirrored.*

Performing this explorative study and by looking at individual biplots and stimulus plots, the hypotheses of clustering commenced.

***Hypothesis A*: A cluster analysis will group Amplifier 2 in a group for itself.**

***Hypothesis B*: A cluster analysis will group Amplifier 4 and 5 for themselves.**

***Hypothesis C*: A cluster analysis will group Amplifier 6 and 7 for themselves.**

***Hypothesis D*: A cluster analysis will group Amplifier 3 and 8 for themselves.**

Assuming that the participants are consistent in regards to their preference of the program material from session 1 to session 2 induced Hypothesis e).

***Hypothesis E*: The pairwise preference ratings for each amplifier does not differ significantly from the first preference rating to the second preference rating.**

No hypotheses were stated in regards to attributes.

Chapter 4

Results

4.1 Participants

16 male participants took part in the experiment over two sessions. The participants were from 21 to 58 years old (Mean = 33, Standard Deviation = 8.9). All participants were amateur musicians - mostly guitarists - having played their instrument between 8 and 44 years (Mean = 18, Standard Deviation = 9). Ten participants answered "yes" in at least one of the sessions to whether they had been exposed to loud sound pressure or loud noise within the last 24 hours.

4.2 Assumptions of the Dissimilarity Rating data

The dissimilarity data for each participant is measured on at least ordinal level but is assumed to be measured at a ratio scale. The respondents were able to rate from "exact same" to "as different as possible" on a 7-point scale. The assumption of ratio measurement seems valid due to the participant being able to respond in which degree each stimuli pair were dissimilar. However, a nonmetric MDS solution seemed to be the only correct procedure for perfoming a MDS solution to the averaged dissimilarity data due to the dissimilarity scale not being continuous. A nonmetric MDS solution also resulted in a lower stress measure (5.84×10^{-7}) than the metric MDS solution (0.023), which is less than a threshold once set to be 0.15 for an acceptably precise MDS solution [Borg and Groenen, 2010]. This threshold is set for ordinal MDS only [Borg and Groenen, 2010]. More information on stress measures are seen in 8.2.2 in the Appendix. The individual metric MDS plots in the Appendix (section 8.15) are metric MDS plots due an attempt of applying the nonmetric solutions to the individual data eliminated both subtle and major differences. All MDS solutions were performed using the SMACOF package [de Leeuw and Mair, 2009] for the statistical computing software, R.

4.3 Common Stimulus Plot

Averaging the dissimilarity matrices for all participants generates a stimulus plot showing the general perception of the program material. The averaged dissimilarity ratings are seen in Table 4.1. Plotting an MDS solution of these dissimilarities yields the plot in Figure 4.1. However, the plot does not represent deviations from the mean points of the stimuli. In order to present the deviations, Table 4.2 shows a matrix where each entry point shows the standard deviation from the mean value of this specific entry. A further risk is that an average of dissimilarity plots will eliminate - or cancel out - obvious groupings and patterns. Comparing the common stimulus plot with the biplot of the two principal components of all direct ratings of all attributes show a high similarity between the two plots and the groupings of the points. The biplot of the two principal components for all 16 participants' attribute ratings (Figure 4.3) also indicate the same groupings as the common stimulus plot. According to Ramsay [1980] joint estimation of the position of the points in a MDS-derived stimulus space will in general produce less biased parameter estimates with a smaller sampling variance. This, due to the joint analysis using all available information in the estimation of stimuli-similarity. Direct ratings such as attribute ratings for each stimulus contain some information regarding coordinates that dissimilarity ratings do not [Ramsay, 1980]. A joint estimation has not been conducted but due to the stimuli points' positions in both biplot common stimulus plot indicating same groupings of the amplifiers. Based on this knowledge the stimulus plot derived by the averaged dissimilarity ratings (Figure 4.1) is accepted as a valid representation of the participants' perception of these amplifiers' dissimilarities.

Table 4.1: *Averaged Dissimilarity Matrix for the 16 participants.*

	Amp 1	Amp 2	Amp3	Amp4	Amp5	Amp6	Amp7	Amp8
Amp 1	0.0000	5.1250	3.3125	5.5625	5.6250	3.6875	3.6875	4.5000
Amp 2	5.1250	0.0000	5.0625	5.9375	6.0625	5.0000	5.0625	5.6875
Amp 3	3.3125	5.0625	0.0000	5.6875	6.1875	2.4375	3.3125	1.5625
Amp 4	5.5625	5.9375	5.6875	0.0000	2.0000	5.1250	5.2500	5.8125
Amp 5	5.6250	6.0625	6.1875	2.0000	0.0000	5.3750	5.9375	5.9375
Amp 6	3.6875	5.0000	2.4375	5.1250	5.3750	0.0000	2.7500	3.1875
Amp 7	3.6875	5.0625	3.3125	5.2500	5.9375	2.7500	0.0000	3.0000
Amp 8	4.5000	5.6875	1.5625	5.8125	5.9375	3.1875	3.0000	0.0000

Table 4.2: *Standard Deviation Matrix of the Averaged Dissimilarity Matrix. All values are rounded to three decimal places.*

	Amp 1	Amp 2	Amp3	Amp4	Amp5	Amp6	Amp7	Amp8
Amp 1	0.000	1.218	1.102	1.059	1.409	1.402	1.402	0.866
Amp 2	1.218	0.000	1.298	1.088	0.899	1.173	1.345	1.102
Amp 3	1.102	1.298	0.000	1.102	0.726	0.933	1.102	0.704
Amp 4	1.059	1.088	1.102	0.000	0.612	1.363	1.250	1.130
Amp 5	1.409	0.900	0.726	0.612	0.000	1.111	0.555	1.088
Amp 6	1.402	1.173	0.933	1.364	1.111	0.000	0.968	1.014
Amp 7	1.402	1.345	1.102	1.250	0.555	0.968	0.000	1.225
Amp 8	0.866	1.102	0.704	1.130	1.088	1.014	1.225	0.000

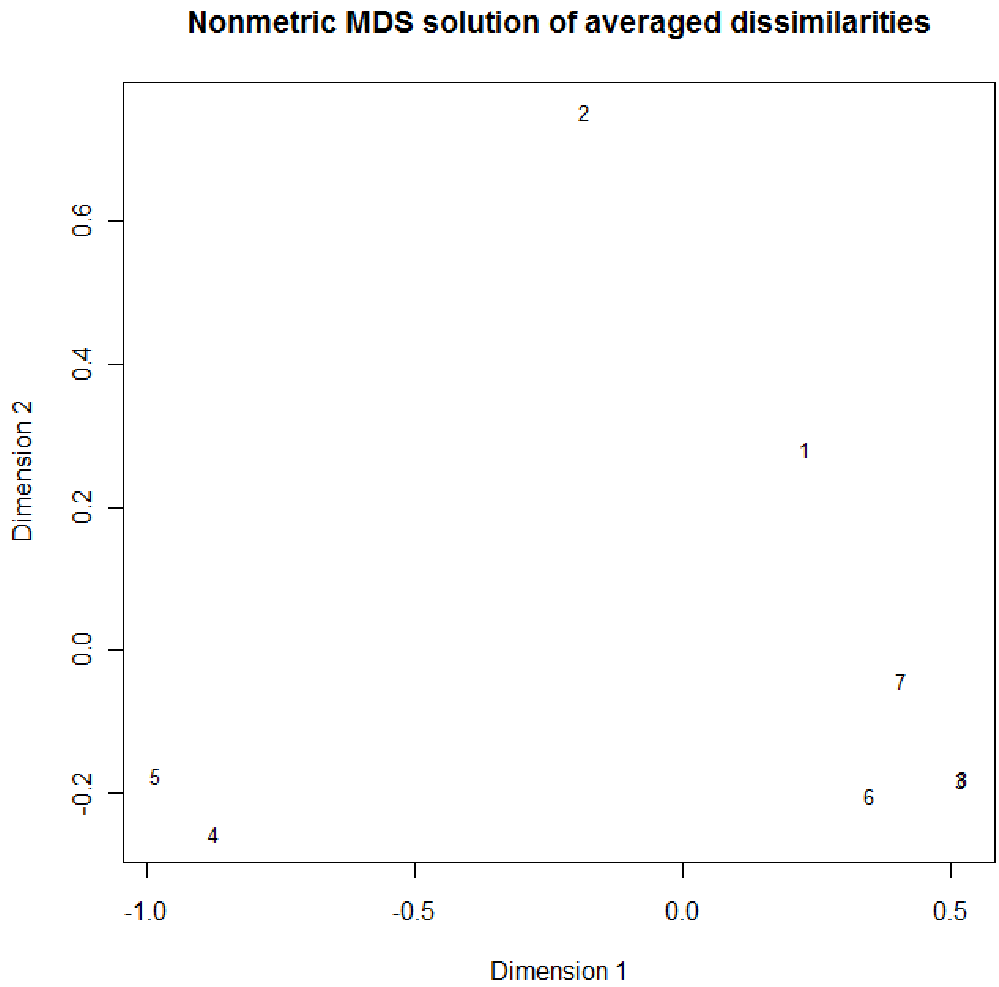

Figure 4.1: *Nonmetric MDS solution based on the 16 averaged dissimilarity matrices. All data has been divided by seven in order to have dissimilarity values ranging from 0 to 1. The amplifiers are - from 1 through 8 - as follows: (1) Mesa Boogie Dual Rectifier, (2) Engl e625 Fireball, (3) Engl e635 Fireball, (4) Marshall Ma50H, (5) Marshall TSL60, (6) Peavey 6505+, (7) Randall Kirk Hammet RM100H, and (8) Engl e635 w/ midboost on. The coordinates of Amplifier 3 and 8 are similar and the two amplifiers are therefore plotted in the same position. None of the individual biplots or MDS plots positions point 3 and 8 as close to each other as this common stimulus plot. The positions may be an expression of the error-like measure when averaging dissimilarity ratings. The two points do represent the same amplifier but auditive difference as a result of the Engl e635 Fireball midboost should be perceivable.*

The plot in Figure 4.1 show major difference in the perceived sound of the amplifiers. Amplifier 4 and 5 seems to form a group, which could be explained by both of the amplifiers being manufactured by Marshall. As mentioned in the caption of Figure 4.1 amplifiers 3 and 8 are positioned in the exact same point. The two points represent the Engl e635 with - and without - the midboost parameter set to 'on' and thus are likely to be positioned close to each other. However, none of the individual biplots or individual MDS plots for the 16 participants (seen in section 8.15 in the Appendix) positions these amplifiers as close as the common stimulus plot. The two amplifiers being positioned in the same point may be an expression of an error-like measure when averaging the dissimilarity ratings. A group containing amplifier 6 and 7 also seems to be present as well as amplifier 2 seeming very different from the other amplifiers.

Figure 4.2 illustrate the transformation of the dissimilarities into configuration distances represented in the common stimulus plot. The diagram resembles a monotonically increasing line, i.e. the perfectly fitting solution [Wickelmaier, 2003], and thus provide a measure of the 2-dimensional MDS solution plot as being an acceptable fit of the averaged dissimilarity ratings.

4.3.1 Individual Differences

A three-way MDS procedure could be performed to study the individual differences of the 16 participants. However, data for such analyses typically require individual ratings of the same attributes for all participants. Scrutinising the individual MDS plots and biplots in section 8.15 in the Appendix gives an idea of the individual differences in the data. The groupings of the amplifiers described in the previous section are also observed in most of the individual plots.

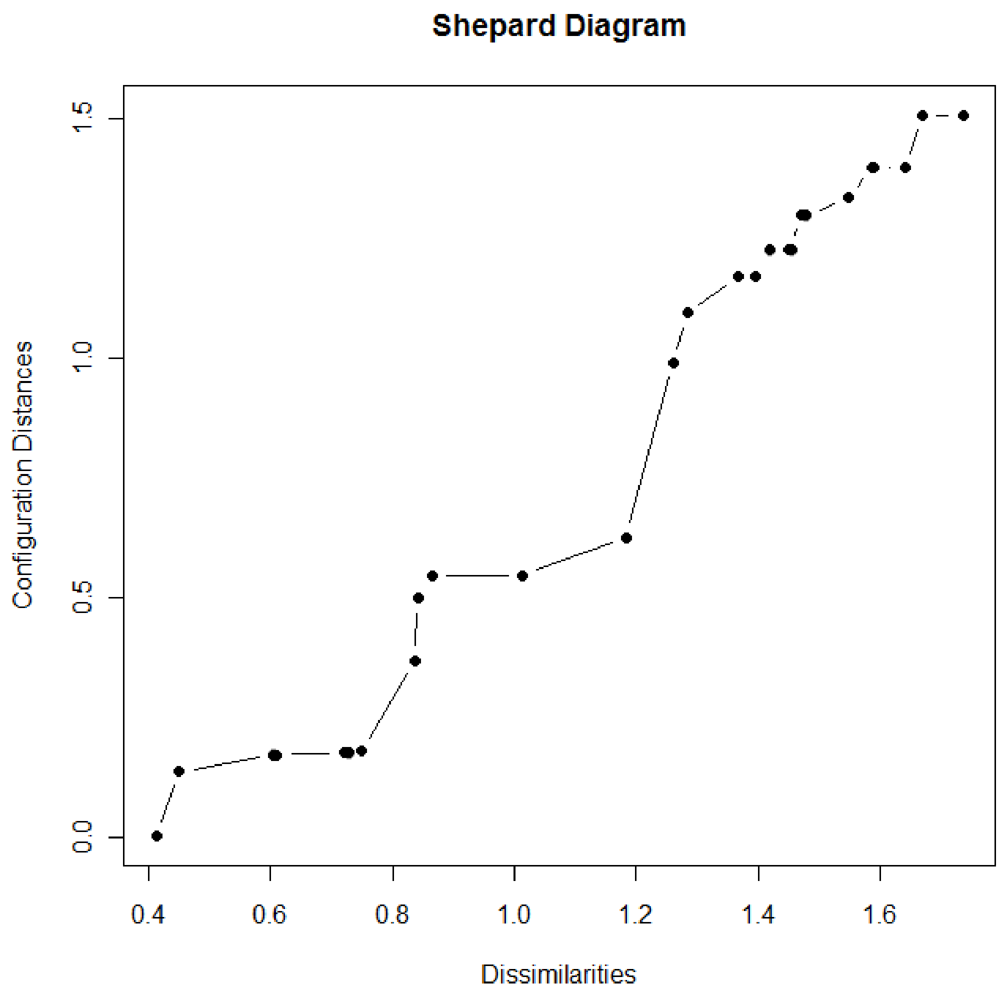

Figure 4.2: *Shapard Diagram of the relationship between dissimilarities and configuration distances.*

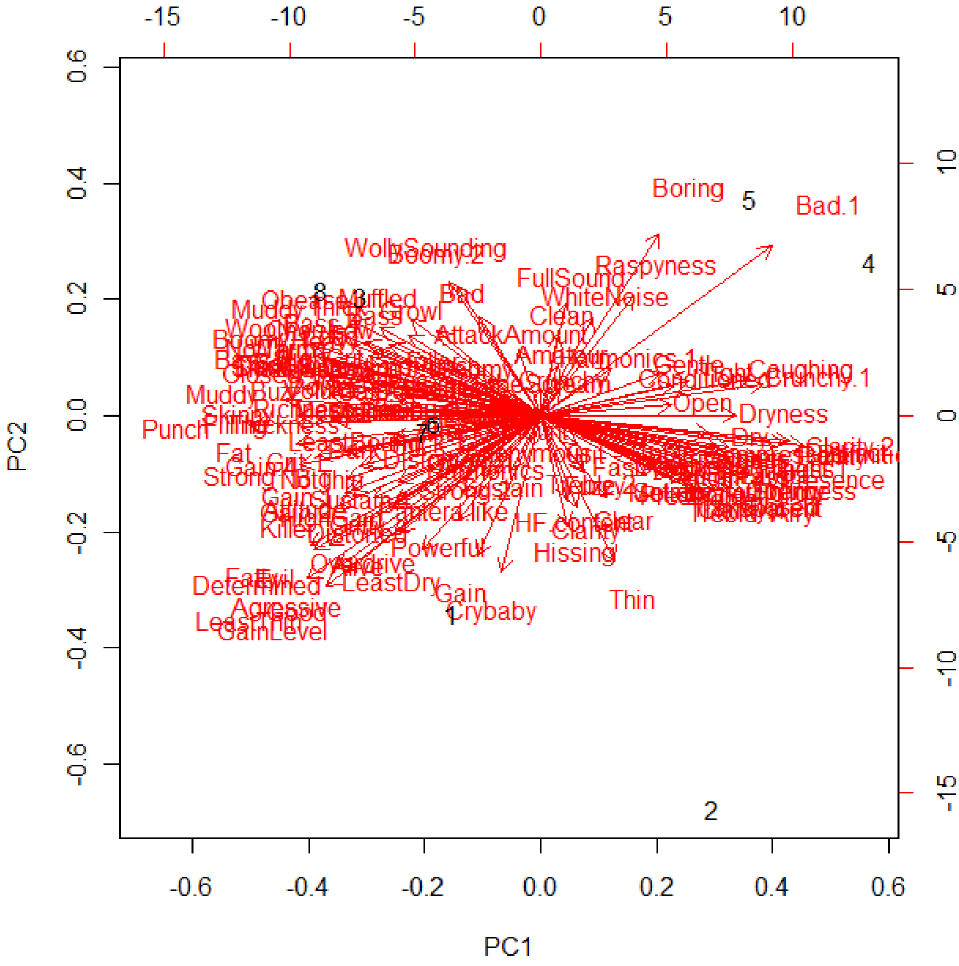

Figure 4.3: *Biplot of the two principal components of all participants' attribute ratings. The plot seems to cluster the amplifiers in the same patterns as seen in the common stimulus plot, derived by dissimilarity data, i.e. amplifier 2 being very different from the other amplifiers, amplifier 3 and 8 being very similar, amplifier 6 and 7 being very similar, and amplifier 4 and 5 being very similar.*

4.4 Interpreting the common stimulus plot

Dimension 1 in Figure 4.1 seems to differentiate the amplifiers with amplifier 4 and 5 in one end and amplifier 6, 7, 3, and 8 in the other end. An equivalent dimension, related to bass in one end and treble/brightness in the other, is indicated through the biplots of participant 1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 13, 14, and 16 (found in Appendix). Dimension 2 seems to differentiate the amplifiers with amplifier 4 and 5 in one end and amplifier 2 in the other end. Observing the biplots for indications of what differentiates amplifier 4 and 5 from amplifier 2 in no way implies a clear answer but rather a quite ambiguous one suggesting amplifier 2 to be more metallic (participant 4), aggressive (participant 10), hissing and harsh (participant 8 and 16, respectively), and have more gain (participant 1, 3 and 10). Due to the amplifiers in the positive end of Dimension 1 seeming to possses more 'thickness', 'bass', and being more ‘boomy’, Dimension 1 may be related to amount of low-frequency content. Dimension 2 may be related to amount of high-frequency content due to amplifier 2 being described as hissing, harsh, metallic, and having more gain.

4.4.1 Hypothesis A, B, C, D

As previously described the amplifiers are believed to be clustered into groups. In order to measure the inter-cluster dissimilarity, the commonly used procedure *average linkage clustering* [Everitt and Hothorn, 2010] was used. A so-called *dendrogram*, i.e. a two-dimensional diagram illustrating the fusions made at each stage of the cluster analysis, of the average linkage clustering is seen in Figure 4.4.

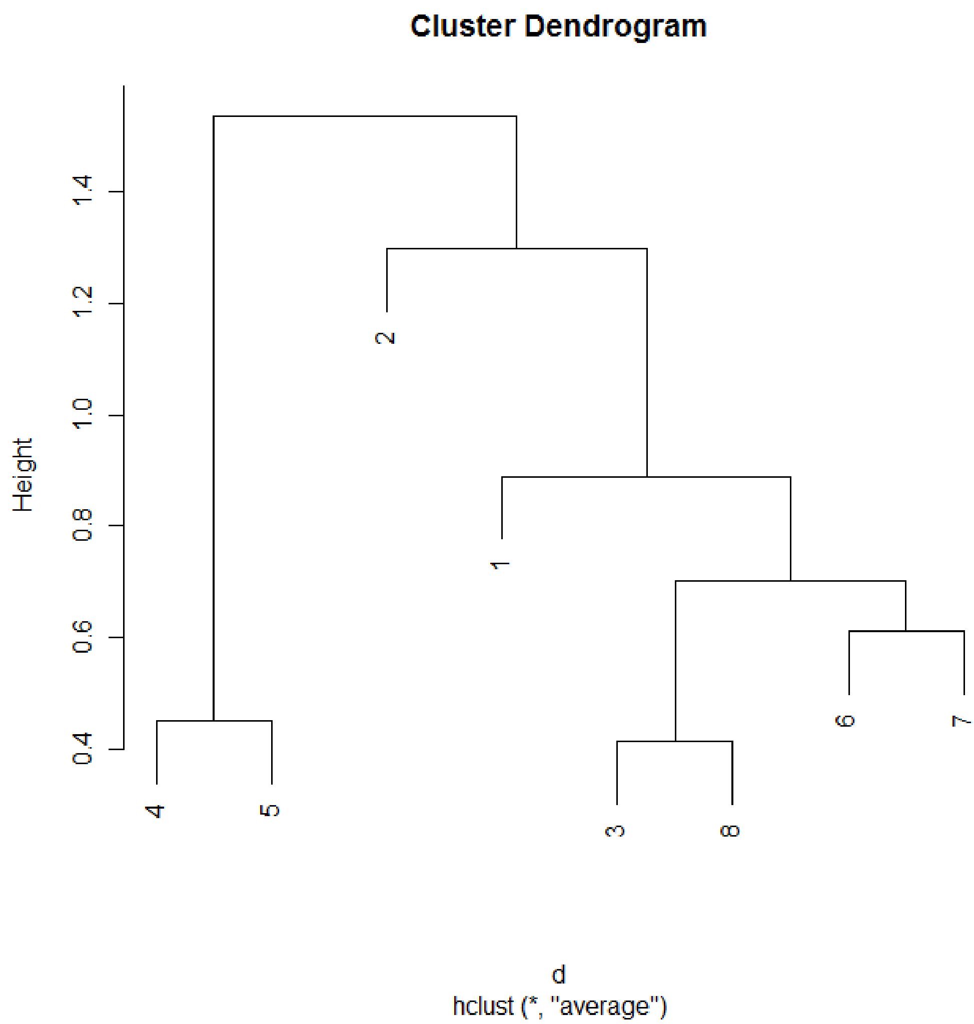

Figure 4.4: *Dendrogram of the average linkage clustering.*

The dendrogram in Figure 4.4 confirms the clustering of the stimuli seen in the common stimulus plot (Figure 4.1), i.e. amplifier 4 and 5 being similar, amplifier 3 and 8 being similar, and amplifier 2 being very dissimilar from the other amplifiers. Furthermore, amplifier 6 and 7 seem to form a group. Hypotheses A stating that “***a cluster analysis will group***

Amplifier 2 in a group for itself" is confirmed by the dendrogram in Figure 4.4, which also confirms Hypothesis B ("***a cluster analysis will group Amplifier 4 and 5 for themselves***"), Hypothesis C ("***a cluster analysis will group Amplifier 6 and 7 for themselves***"), and Hypothesis D ("***a cluster analysis will group Amplifier 3 and 8 for themselves***").

4.5 Preference Ratings

Figure 4.5 shows boxplots of the participants' preference for each of the eight amplifiers. Each amplifier is listed twice due to the amplifier preference being rated in both session 1 and session 2.

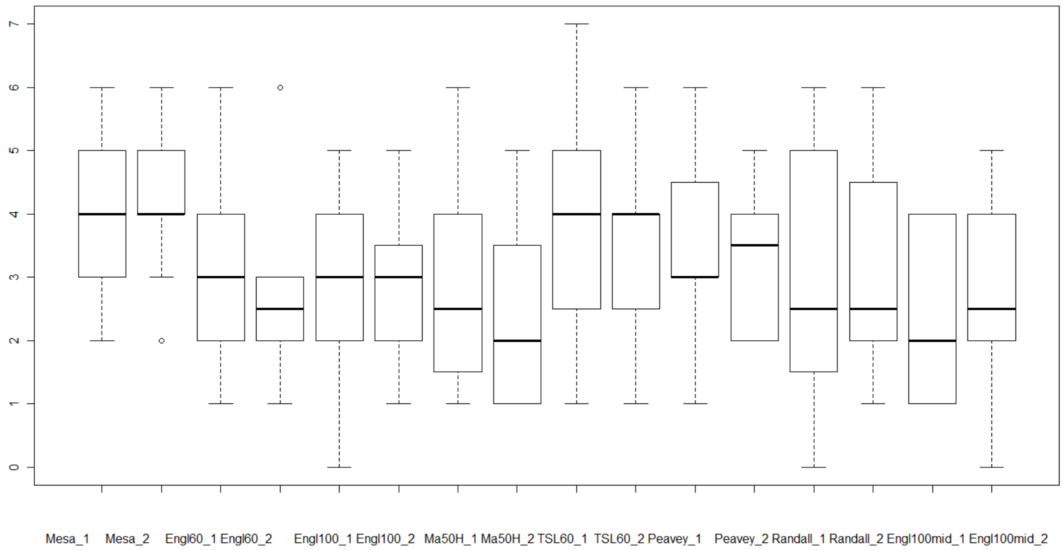

Figure 4.5: *Boxplot of the participants' preference of each amp for session 1 and session 2. Small deviations are noticable and indicate a small inconsistency in the participants' preference ratings.*

Bonferroni adjusted paired comparisons by t tests revealed no significant differences in preference between the amplifiers in session 1. In session 2 the same statistical procedure revealeded the Mesa Boogie Dual Rectifier ($M = 4.25, SD \approx 1.18$) being prefered significantly over the following amplifiers: Engl e635 Fireball (midboost off) ($M = 2.6875, SD \approx 1.20, p = 0.048$), Engl e625 Fireball ($M = 2.5, SD \approx 1.21, p = 0.015$), and Marshall Ma50H ($M = 2.375, SD \approx 1.45, p = 0.037$).

4.5.1 Hypothesis E

A Wilcoxon rank sum test with continuity correction revealed no significant difference in preference between session 1 and session 2, (V = 1261, p = 0.1981, p > 0.05) - hence Hypothesis E ("***The pairwise preference ratings for each amplifier does not differ significantly from the first preference rating to the second preference rating***") cannot be rejected.

Chapter 5

Conclusion

5.1 Multidimensional Scaling

Under the given circumstances significant dissimilarities were observed between the amplifiers recorded for the listening experiment. The Engl e625 Fireball 60 W amplifier seemed to be the most different from the other amplifiers. Not much difference was perceived between the Engl e635 Fireball 100 W amplifier when the midboost parameter was set of 'on' and 'off' compared to the differences observed between other amplifiers. The Marshall MA50H amplifier was perceived quite similar to the Marshall TSL60 amplifier. The Peavey 6505+ and the Randall Kirk Hammet RM100H amplifier was also perceived to sound quite similar whereas the Mesa Boogie Dual Rectifier also seemed to belong to this group; although peripherally. The hypotheses A, B, C, and D were confirmed concluding that *"a cluster analysis will group Amplifier 2 in a group for itself"*, *"a cluster analysis will group Amplifier 4 and 5 for themselves"*, *"a cluster analysis will group Amplifier 6 and 7 for themselves"*, and finally that *"a cluster analysis will group Amplifier 3 and 8 for themselves"*.

5.2 Amplifier Preference

No significant differences in amplifier preference was observed in the initial session of the listening experiment (session 1) although the Mesa Boogie Dual Rectifier ($M = 4.25, SD \approx 1.18$) was prefered significantly over the Engl e635 Fireball (midboost off) ($M = 2.6875, SD \approx 1.20, p = 0.048$), the Engl e625 Fireball ($M = 2.5, SD \approx 1.21, p = 0.015$), and the Marshall Ma50H ($M = 2.375, SD \approx 1.45, p = 0.037$) in session 2. The participants seemed fairly consistent in their preference ratings for the specific amplifiers in both sessions of the experiment (V = 1261, p = 0.1981, p > 0.05) which led to not being able to reject Hypothesis E ("*The pairwise preference ratings for each amplifier does not differ significantly from the first preference rating to the second preference rating*").

5.3 Perceptual Attributes

In regards to attribute relating to Dimension 1 and Dimension 2 in the common stimulus plot generated by averaging dissimilarity ratings nothing is concluded but indications of the dimensions being related to timbre are suggested, namely Dimension 1 being related to amount of low-frequency content and Dimension 2 being related to amount of high-frequency content.

Chapter 6

Discussion

In the problem formulation (section 1.11) it is stated that the mental structure of multiple perceptual attributes can be quantitatively captured in a multidimensional space. This has been proven in this report; not only by generating a perceptual representation of the stimuli by using dissimilarity ratings but also that a perceptual representation, very much alike the one derived by dissimilarity ratings, can be derived by generating a biplot of only two principal components of the attribute rating. These two methods aid each other in interpreting each other. The use of perceptual mapping thus seems to be a helpful tool for users, customers, audio professionels, and development teams in visualising the complex perceptual structures of these products' attributes.

6.1 Attributes, RGT, and semantic differential

Some participants reported the task of eliciting attributes as "healthy" for them due to them not being used to describing perceived auditory attributes this thoroughly. In general, the participants reported selecting an attribute that differed from the other two in the RGT task as being fairly easy while the task of selecting an attribute making the other two stimuli alike was more difficult. One participant reported it helpful to having completed the RGT task in session 1 before the rating of dissimilarities in session 2 due to him relating the differences more to the elicited attributes from session 1.

Many attributes elicited by the RGT were either cognitive or affective. An interesting trend in these attributes were the correlation between sensation and descriptors related to visuo-haptic features, e.g. "wooly", "thick", "hollow", "heavy", "fuzzy", "gritty", "metallic", "crisp", "closed", "buzz", "dark", "mushy"

Some attributes also tended to related to specific bands such as the attribute "Pantera-like" elicited by participant no. 6.

Initially, attributes such as "Darkness" were put on semantic differential VAS scales by the facilitator with the end-points "darkest" and "brightest". However, this seemed to bias the ratings of the attribute towards the end-point semantically more similar to word elicited in the RGT task. This was solved by being fairly consistent in defining end-points as either *more* or *less* of the specific attribute in order to increase the resolution of the VAS scale.

Another argument for presenting scale end-points followed by "least" and "most" is the risk of interpreting a participant's meaning of the word erroniously. Participant 4 constructed the attribute "High-frequency content" which - on a semantic differential scale - could move from "low-frequency content" to "high-frequency content" or from "least amount of high-frequency content" to "most amount of high-frequency content", i.e. the position of frequency emphasis contra a measure of the magnitude of high-frequency content.

6.2 Intra-subject related constructs

In order to find constructs related from participant to participant Berg and Rumsey [1999] analysed each subject's grid by cluster analysis. Berg and Rumsey [1999] implied that similar constructs "*are linked together at their level of match, thus forming a 'new' construct.*". They used the number of unrelated constructs as an indication of the approximate number of latent variables.

Due to the 16 participants having rated their own personally constructed attributes, the biplots generated by individual attribute rating also represent relations between stimuli in 16 *different* geometrical spaces, making quantitative comparisons using only attribute ratings impossible. This impossibility in comparison using only attribute ratings could be avoided in the experimental design by having participants rate the same attributes. These attributes could be elicited by e.g. cluster analyses [Berg and Rumsey, 1999], previous research [Gromko, 1993; Tessarolo, 1981], or through stages of Quantitative Descriptive Analysis [Bech and Zacharov, 2006]. Making a group of participants have a common understanding of a set of attributes can require quite some ressources [Bech and Zacharov, 2006] but can be seen as an investment. The necessity of ensuring a common set of attributes is up for discussion. In some cases it may be assumed that the participants rate attributes in the same perceptual domain.

6.3 Program Material

The stimulus plot is to a high degree dependent on the settings of the high-gain guitar amplifiers. Data on these settings was acquired through an online survey allowing the respondent to assign a value for each control parameter between 1 and 10. This resolution may not have been sufficient as large variations in sound output are produced varying a control parameter less than a tenth of its control range. Furthermore, not as much data on each amplifier was collected as wanted. The representative aspect of these settings is therefore of questionable character.

6.3.1 Loudness Differences

Several participants pointed out that the levels of volume did not seem to be equal for the eight stimuli. Especially amplifier 2 was perceptually less loud and amplifier 1 seemed to be loudest. It is uncertain in which degree this influenced the data.

6.4 Separating timbre and gain

In the attempt to interpret Dimension 1 and Dimension in the common stimulus plot, it was proposed that the dimensions are dependent on the amount of low and high frequencies, respectively, suggesting timbre to be a salient dimension. In their study, Martens and Marui [2002] argue for the attempt to separate the distortion-related timbral quality from the variation in tone coloration associated with the complex perceptual differences between modeled guitar amplifier outputs. Martens and Marui [2002] stated that "*the three modeled guitar amplifiers had characteristic timbre associated with the nonlinear distortion introduced by each, varying in a manner not unlike the differences in interharmonic distortion characteristic of onset transients of brass horns (...)*" It is therefore plausible that the variations in timbre in the common stimulus space are also a products of varying amount of gain between the amplifiers.

6.5 Using boostrapping to generate pseudo-confidence regions in MDS solutions

In order to predict where a specific amplifier may be positioned in the stimulus plot applying bootstrapping can generate pseudo-confidence regions [Potts, 1999]. Applying MDS to each participant's data individually is of limited interest when wanting to conclude something general.

Assuming that each participant's dissimilarity matrices come from a population of dissimilarity matrices, the averaged dissimilarity matrix represents a population mean. In order to visualise a standard-error like statistic of this mean graphically, the empirical method of boostrapping is utilised. Estimating sampling properties of some statistic can be performed by repeatedly drawing a sample from the population distribution and calculating the median value. Doing this a few thousand times will show sampling properties of the median allowing for giving a confidence interval [Potts, 1999].

When not knowing the true population distribution, a simulation based on this distribution is not possible. Instead the population distribution can be estimated by the maximum likelihood estimate, which is known as the empirical distribution function [Potts, 1999]. Simulating from this empirical distribution functions is known as bootstrapping. The same observation may be present more than once in a bootstrap sample where other observations will not be represented.

An empirical multivariate distribution function can be formed by the sample containing the individual participants' dissimilarity matrices. Bootstrapping from this distribution function and re-calculating the stimulus plot for each bootstrapped averaged dissimilarity matrix produces an MDS output for each bootstrap. Limitations are, however, that this is only sufficient for a scalar statistic represented by a single number.

Due to the invariance of stimulus plots under rotation, reflection, and translation, each bootstrapped stimulus plot has an arbitrary rotation, reflection, and translation, which limits the combination of outputs [Potts, 1999]

Holding all items constant except one enables comparison between resulting plots. Not boostrapping complete MDS solutions but keeping most of the solution constant and only re-fitting few individual items for each boostrap, conditioning of the rest of the map can remain constant [Potts, 1999].

Repeating this will eventually generate a point cloud, which represents intersubjective variability for each item. Due to the stimulus plot being represented in two dimensions, a confidence region - not interval - must be calculated and represented. Instead of constructing a circle with center on the original estimate Potts [1999] argues for estimating the multivariate distribution of all possible boostraps. Hereby the distribution may be estimated using two-dimensional kernel density estimation. "*The process of kernel density estimation is like plotting the data in a histogram and then smoothing the histogram boxes a continuous distribution*" [Potts, 1999]. Having estimated such as distribution, isocontour lines can be drawn joining points of equal probability. Potts [1999] suggests selecting an isocontour containing alpha of the volume of probability under the distribution giving an aplpha confidence region.

The calculated confidence regions will aid interpreting what is a result of chance and what is genuine effects of the stimuli. Potts [1999] stresses the care needed in interpreting these results as (a) the output should strictly be considered pseudo-confidence regions and (b) the position of all points in the stimulus plot are uncertain and moving one point will affect all the others. Another disussion is the number of test participants where Potts [1999] states that 43 participants are no large population and subtle (but real) differences may go unnoticed.

6.6 Joint Estimation Analysis

The relationship between direct ratings and the configuration underlying the dissimilarities has been expressed in a variety of ways. According to Ramsay [1980] the simplest of these is the scalar product model:

$$u^*_{igr} = \sum_{m}^{k} a_{grm} x_{im} + c_{gr},$$

where a_{grm} defines the ideal direction of attribute g and subject r. "*According to this model, if a_{grm} is positive, the more of aspect* m *underlying the dissimilarities the stimulus has, the more highly it will be rated with respect to property* g *by subject* r" [Ramsay, 1980]. x_{im}

denotes the coordinates, c_{gr} is a coefficient, and a u^*_{igr} is the direct rating of stimulus i for attribute g by participant r. However, a disadvantage of the model is that it does not take saturation of an attribute into account and would e.g. not be able to predict the shift from "Wildness" to "Heaviness" for the `drive` parameter in the study by [Atsushi and Martens, 2001] mentioned in Section 1.9. Due to this limitation of the model, the ideal point model has emerged. In the ideal point model it is postulated that the rating for a specific stimulus is a function of the distance between the ideal point and the coordinates of the rated stimulus [Ramsay, 1980].

Chapter 7

Perspectives

7.1 Perceptual mapping as a tool in Audio Development

A graphical representation of the perceptual differences of existing products within the category of interest can help the development team to a common knowledge. Combining this representation with scaled dimensions of attributes will also increase understanding in communication within the team. Furthermore, when knowing the point in the stimulus space to be occupied by the future product, data on preference, attributes, and physical parameters will help the engineers reach this specific perceptual goal. For instance, if a preference for “thick” high-gain distortion sounds is observed and a physical parameter is highly correlated with “thickness”, manipulating this physical parameter will most likely be prefered by the customer. Here it is important to know whether an ideal point of this parameter exists. Preference is likely to decrease again if overemphasising a parameter as Ramsay [1980] points out using the example of an optimum amount of sugar in food.

7.2 Perceptual research as a foundation of user-centered user interfaces

Further thorough and systematic research regarding perception of high-gain guitar amplifier could also as proposed by Atsushi and Martens [2001] act as the foundation of creating more user-centered user interfaces of amplifiers. With the rise of digital modulation of high-gain guitar amplifiers, some of the control parameters related directly to the electrical circuit within the amplifier can be seperated more into perceptually unidimensional control parameters, i.e. one control parameter changing only one perceptual characteristic. Clearly uncorrelated perceptual dimensions could controlled by each their knob on the user interface. Typically control of gain, tone, and volume of high-gain amplifiers has been designed in a 1:1 mapping of the amplifier’s electrical circuit. MDS and perceptual mapping can therefore be utilised as a mean of adapting the controls to the guitarists rather than professionals within the field of electrical circuit engineering.

7.3 Perceptual mapping as a tool for the customer

Perceptual mapping of the products of interest can potentially increase customer understanding of the perceptual differences and the relations between amplifiers within the same category. A similar marketing tool has been introduced by Boss, illustrated in Figure 7.1.

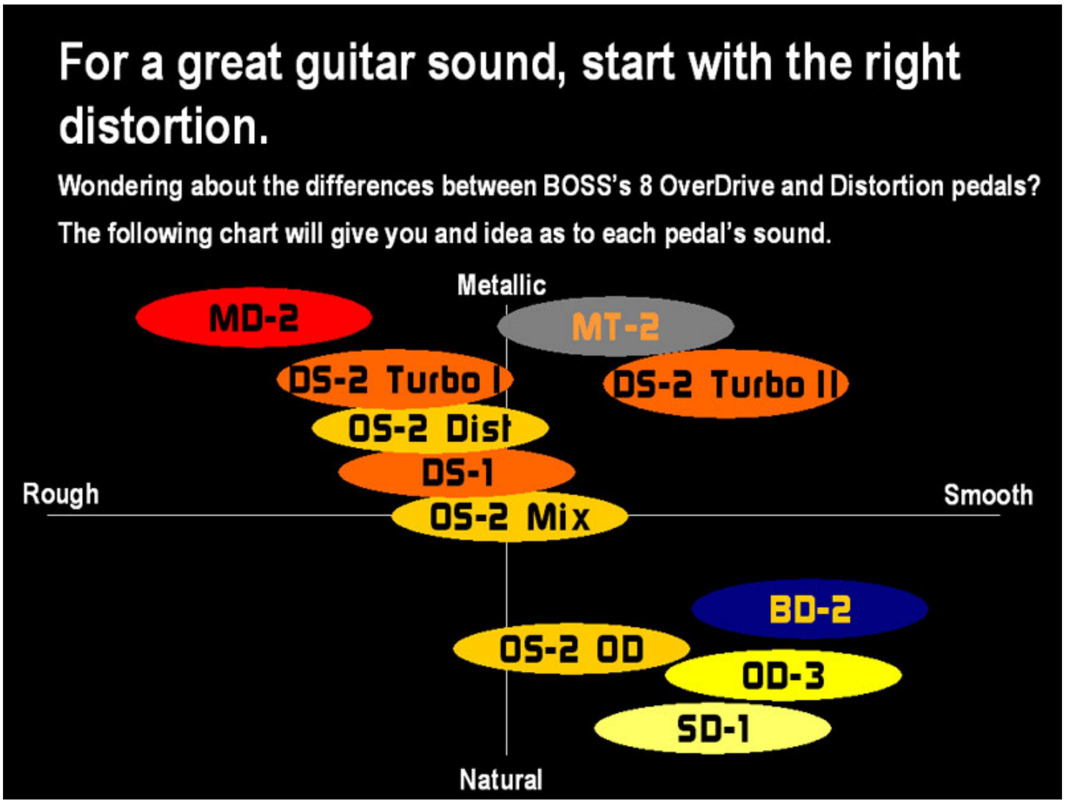

Figure 7.1: *This distortion chart from Boss depicts the diffences between Boss products within the category of overdrive and distortion. It is assumed that the two bipolar semantic differential dimensions "Metallic-Natural" and "Rough-Smooth" help the consumer to differentiate between the vast amount of products within the distortion category. Screen dump of illustration by Boss® [Boss Coorporation, 2007].*

Bibliography

Atsushi, M. and Martens, W. L. [2001], Perceptual and Semantic Scaling for User-Centered Control Over Distortion-Based Guitar Effects, *in* 'Audio Engineering Society Convention Paper 5387', Audio Engineering Society.

Bærentsen, K. [2000], 'Special issue on information technology in human activity - "Intuitive User Interfaces"', *Scandinavian Journal of Information Systems* **12**(1-2), 29–60.

Bech, S. and Zacharov, N. [2006], *Perceptual Audio Evaluation: Theory, Method and Application*, John Wiley & Sons Inc.

Bell, R. C. [2003], *International Handbook of Personal Construct Psychology, Chapter 9: The Repertory Grid Technique*, John Wiley & Sons, Ltd. chapter 9: The Repertory Grid Technique.

Berg, J. and Rumsey, F. [1999], 'Spatial attribute identification and scaling by repertory grid technique and other methods', *Proceedings of the 16th AES International Conference on Spatial Sound Reproduction* pp. 51–66.

Black Sabbath [1971], *Sweet Leaf.* from the album 'Master of Reality' released by Vertigo (UK) and Warner Bros. Records (US).

Borg, I. and Groenen, P. J. F. [2010], *Modern Multidimensional Scaling - Theory and Applications*, 2nd edn, Springer.

Boss Coorporation [2007], *BOSS GUITAR EFFECTS GUIDE BOOK Vol. 20*, `http://www.bossarea.com/pocket/2007/Vol20/4247_GuitarEffects_Cata_72.pdf`. Downloaded: Februart 14th, 2011.

Bradshaw, J. M., Ford, K. M., Adams-Webber, J. R. and Boose, J. H. [1993], 'Beyond the Repertory Grid: New Approaches to Constructivist Knowledge Acquisition Tool Development', *International Journal of Intelligent Systems* **January and February**.

Bruner, J. S., Goodnow, J. J. and Austin, G. A. [1956], *A Study of Thinking*, Transaction Publishers.

Choisel, S. and Wickelmaier, F. [2005], 'Extraction of auditory features and elicitation of attributes for the assessment of multichannel reproduced sound', *Convention of the Audio Engineering Society, Barcelona, Spain, 2005 May 28-31* . article available in Sylvain Choisel's ph.D. thesis referred to elsewhere in this bibliography.

Cooper, L. G. [1983], 'A review of multidimensional scaling in marketing research', *Applied Psychological Measurement* **78**(4), 427–450.
URL: *http://apm.sagepub.com/content/7/4/427.full.pdf+html*

de Leeuw, J. and Mair, P. [2009], 'Multidimensional scaling using majorization: SMACOF in R', *Journal of Statistical Software* **31**(3), 1–30.
URL: *http://www.jstatsoft.org/v31/i03/*

Deutsch, D. [1972], 'Mapping of Interactions in the Pitch Memory Store', *Science* **175**, 1020–1022.

Deutsch, D. [1975], 'Disinhibition in pitch memory', *Perception & Psychophysics* **17**(3), 320–324.

Deutsch, D. [n.d.], 'Tones and Numbers: Specificity of Interference'.

Ekman, G. [1954], 'Dimensions of color vision', *The Journal of Psychology* **38**, 467–474.

Everitt, B. and Hothorn, T. [2010], *A Handbook of Statistical Analyses using R*, Chapman & Hall/CRC.

Eysenck, M. W. and Keane, M. T. [2005], *COGNITIVE PSYCHOLOGY - A Student's Handbooks*, Psychology Press.

Fallman, D. and Waterworth, J. [2005], 'Dealing with User Experience and Affective Evaluation in HCI Design: A Repertory Grid Approach', *CHI 2005* pp. 1–5.

Field, A. [2009], *Discovering Statistics Using SPSS*, 3nd edn, SAGE Publications.

Gibson, E. J. and Pick, A. D. [2000], *An Ecological Approach to Perceptual Learning and Development*, Oxford University Press.

Gibson, J. J. [1958], 'A theory of pictorial perception', *HFORL Memorandum Report* (85), 3–23.
URL: *http://www.springerlink.com/content/k82801l0r85u1416/fulltext.pdf*

Gibson, J. J. [1974], *A Note on the Relation Between Perceptual and Conceptual Knowledge*. The World Wide Web distribution of James Gibson's "Purple Perils" is for scholarly use with the understanding that Gibson did not intend them for publication. References to these essays must cite them explicitly as unplublished manuscripts. Copies may be circulated if this statement is included on each copy.".
URL: *http://www.trincoll.edu/depts/ecopsyc/perils/folder6/perceptc*

Gibson, J. J. [1986], *The Ecological Approach to Visual Perception*, Psychology Press.

Goldstein, E. B. [2001, 2005], *Blackwell Handbook of Sensation and Perception*, Blackwell Publishing. Handbooks of Experimental Psychology.

Gromko, J. E. [1993], 'Perceptual differences between expert and novice music listeners: A multidimensional scaling analysis', *Psychology of Music* **21**(1), 34–47.

Hornsby, A. S. [2000], *OXFORD Advanced Learner's DICTIONARY of Current English*, 6th edn, OXFORD UNIVERSITY PRESS. (hardback).

Karapanos, E. and Martens, J.-B. [2008], 'The quantitative side of the Repertory Grid Technique: some concerns', *CHI 2008* pp. 1–5.

Karapanos, E., Martens, J.-B. and Hassenzahl, M. [2009], 'Accounting for Diversity in Subjective Judgments', *CHI 2009 ∼ Understanding UI 1* pp. 639–648.

Kortum, P. [2008], *HCI Beyond the GUI, Designed for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces*, Elsevier. A Morgan Kaufman Series edited by Philip Kortum.

Martens, W. L. and Marui, A. [2002], Multidimensional Perceptual Scaling of Tone Color Variation in Three Modeled Guitar Amplifiers, *in* 'Audio Engineering Society Convention Paper 5552', Audio Engineering Society.

Martin, G. and Bech, S. [2005], 'Attribute Identification and Quantification in Automotive Audio - Part 1: Introduction to the Descriptive Analysis Technique', *Audio Engineering Society Convention Paper 6360* pp. 1–14.

Meulman, J., Heiser, W. J. and Carrol, J. D. [1986], *PREFMAP-3 User's Guide.* Tech. Rep., Bell Telephone Laborattories, Murray Hill, NJ, USA.
URL: *http://www.worldfoodscience.org/cms/?pid=1005419*

Mewaldt, S. P. and Hinrichs, J. V. [1973], 'The Ranschburg effect: Stimulus variables and scoring criterion', *Memory & Cognition* **1**(2), 177–182.

Moeslund, T. B. [2009], *Image and Video Processing*, 2nd edn, Computer Vison and Media Technology, Aalborg University.

Moore, B. C. J. [2003], *An Introduction to the Psychology of Hearing*, 5th edn, Academic Press, An imprint of Elsevier Science.

Newbold, P., Carlson, W. L. and Thorne, B. [2006], *Statistics for Business and Economics*, 6th edn, Pearson Prentice Hall.

Newscombe, N. S. and Huttenlocher, J. [2000], *Making Space - The Development of Spatial Representation and Reasoning*, The MIT Press. A Bradford Book'.

Nørager, R. [2009], *Low level cognition in user interfaces, Ph.D. dissertation*, Department of Psychology, University of Aarhus, Aarhus.

Novello, A., McKinney, M. F. and Kohlrausch, A. [2006], 'Perceptual evaluation of music similarity', pp. 1–4.

Pantera [1992], *Walk.* from the album 'Vulgar Display of Power' released by Eastwest Records.

Potts, H. W. W. [1999], *Human Food Rejections (ph.D. dissertation)*, University of London, Department of Psychology, Institue of Psychiatry.

Ramsay, J. O. [1980], 'The Joint Analysis of direct ratings, pairwise preferences, and dissimilarities', *Psychometrika* **45**, 149–165.

Roads, C., Strawn, J., Abbott, C., Gordon, J. and Greenspun, P. [1996], *the computer music tutorial*, The MIT Press, Cambridge, Massachussets; London, England.

Schiffman, S. S., Reynolds, M. L. and Young, F. W. [2007], *Introduction to Multidimensional Scaling THEORY METHODS, AND APPLICATIONS*, 1st edn, JAI Press of Emerald Group Publishing Limited.

Silverchair [1997], *Freak.* from the album Freak Show released by Murmur.

Stern, D. [1995], *Barnets interpersonelle univers*, 2nd edn, Hans Reitzels Forlag A/S. Barnets interpersonelle univers er oversat efter "The Interpersonal World of the Infant".

Tessarolo, M. [1981], 'The musical experience. the semantic differential as a research instrument', *International Review of the Aesthetics and Sociology of Music* **12**(2), 153–164.

Trettvik, J. [2001], 'En økologisk tilgang til perception og aktivitet', *Psyke & Logos* **22**(2), 485–513.

Turner, M. L., Johnson, S. K. and McNamara, D. S. [2001], 'Effects of same-modality interference on immediate serial recall of auditory and visual information', *The Journal of General Psychology* **119**(3), 247–263.

Ueda, K. [2004], 'Short-term auditory memory interference: The Deutsch demonstration revisited', *Acoust. Sci. & Tech.* **25**(6), 457–467.

van de Kerkhof, M. [n.d.], 'Repertory Grid Technique (RGT)', *Unpublished* pp. 1–5.

Velichkovsky, B. M. [1990], 'The vertical dimension of mental functioning', *Applied Psychological Measurement* **52**, 282–289.
URL: *http://www.springerlink.com/content/m07833110273l204/fulltext.pdf*

Weathers, D., Sharma, S. and Niedrich, R. W. [2005], 'The impact of the number of scale points, dispositional factors, and the status quo decision heuristic on scale reliability and response accuracy', *Journal of Busness Research* **58**(2), 1516–1524.

Wedin, L. [1972], 'A Multidimensional Study of Perceptual-Emotional Qualities in Music', *The Scandinavian Journal of Psychology* **13**(4), 241–257.

Wickelmaier, F. [2003], *An Introduction to MDS.*
URL: *http://homepages.uni-tuebingen.de/florian.wickelmaier/pubs/Wickelmaier2003SQRU.pdf*

Wickelmaier, F. [2005], *Indirect scaling methods applied to the identification and quantification of auditory attributes*, Sound Quality Research Unit (SQRU), Department of Acoustics, Aalborg University, Denmark.

Zielinski, S., Rumsey, F. and Bech, S. [2007], 'On some biases encountered in modern audio quality listening tests - a review', *Journal of the Audio Engineering Society* **56**(6), 427–451.

Zimmer, K. and Ellermeier, W. [2003], 'Deriving ratio-scale measures of sound quality from preference judgments', *Noise Control Engineering* **51**(4), 210–215.

Zisapel, N., Tarrasch, R. and Laudon, M. [n.d.], *A Comparison of Visual Analog Scale and Categorical Ratings in Assessing the Patient's Estimate of Sleep Quality*, Landes Bioscience/Eurekah.com and Springer Science+Business Media.

Zölzer, U., Amatriain, X., Arfib, D., Bonada, J., De Poli, G., Dutilleux, P., Evangelista, G., Keiler, F., Loscos, A., Rocchesso, D., Sandler, M., Serra, X. and Todoroff, T. [2002], *DAFX Digital Audio Effects*, John Wiley & Sons, ltd. Edited by Udo Zölzer.

Chapter 8

Appendix

8.1 Terminology

Affective rating *is also known as "hedonic rating"* [Martin and Bech, 2005].

Attribute: *"An attribute, in brief, is any discriminable feature of an event that is susceptible of some discriminable variation from event to event."* [Bruner et al., 1956]

Attack inharmonicity *is "the physical correlate that can aid the human listener in differentiating between tones played by musical instruments belonging to different families (e.g., strings, woodwinds, and brass horns)"* [Martens and Marui, 2002].

Bel *One Bel corresponds to the ratio between two intensities as 10:1. Often decibel [dB] is used, i.e. 1 Bel = 10 dB.* [Moore, 2003].

Bias *is a term used to describe the systematic errors affecting the results of a listening test* [Zielinski et al., 2007].

Centering Bias *denote the bias in which the listeners tend to centre their ratings along the mid value of the scale. The bias does not typically affect the relative distances between judgments but the way in which judgments are projected onto the grading scale.* [Zielinski et al., 2007].

Clipping *is defined as severe distortion of the signal because the amplitude is larger than the processing system can handle* [Zölzer et al., 2002].

Configuration *"is a particular organization of a set of points, that, a map"* [Schiffman et al., 2007].

Contraction Bias *"can be regarded as a conservative tendency ... (where) listeners normally avoid the extremes of the scale." The bias is mostly observed in listening tests where the respondent is not familiar with the stimuli* [Zielinski et al., 2007].

dB *(decibel) is calculated as: number of decibels* $= 10log_{10}(I_1/I_0)$ *where* I_1 *and* I_0 *are different sound intensities measured in SPL, e.g. 60 dB SPL means that* I_1 *is 60 dB* *higher in level than the 0 dB reference level (I_0). The 60 dB SPL has an intensity of* $10^{-6}\ W/m^2$.

Dimension *is "a characteristic that serves to define a point in a space; an axis through the space"* [Schiffman et al., 2007].

Direction *is "a vector through a space that relates to an attribute. A vector is a quantity which possesses both magnitude and direction"* [Schiffman et al., 2007].

Disparities *are "monotonic transformations of the data which are as much like the distances (usually in a least-square sense) as possible)"* [Schiffman et al., 2007].

Euclidean Distance *is "the distance that corresponds to everyday experience. The distance between two stimuli can be calculated from their coordinates according to the Pythagorean formula. In a three-dimensional map, for example, the distance between stimulus A (coordinates X_A, Y_A, Z_A and stimulus B (coordinates X_B, Y_B, Z_B) is*

$$[(X_A - X_B)^2 + (Y_A - Y_B)^2 + (Z_A - Z_B)^2]^{1/2}$$

" [Schiffman et al., 2007].

Factor Analysis

> "*(A factor analysis is a) multivariate technique for identifying whether the correlations between a set of observed variables stem from their relationship to one or more latent variables in the data, each of which takes the form of a linear model.*"

> ''*Are these different variables driven by the same underlying variable? (...) Factor analysis explain the maximum amount of common variance in a correlation matrix using the smallest number of explanatory constructs.*" [Field, 2009].

Hedonic rating is "*a hedonic measurement of a stimulus provides a rating of how much the subject likes the stimulus. Consequently, this is also known as a measurement of the "degree of liking"*" [Martin and Bech, 2005].

Interface Bias *denote bias as a result of the design of the user interfaces used in listening tests. "Listeners seem to use the points of the scale that are associated with labels, numbers, or ticks more frequently than the remaining part of the scale." This may result in quantization effects where distinct peaks in the response histogram are observed exactly where labels, number, or ticks are positioned on the scale* [Zielinski et al., 2007].

Internal judgments *is used by i.a. Zielinski et al. [2007] to denote the listeners' purely psychological implicit judgment made in their minds and not affected by their motor control. However, these two distinctions of judgment may overlap* [Zielinski et al., 2007].

Loudness *corresponds to the subjective impression of the magnitude of a sound; "that attribute of auditory sensation in terms of which sounds can be ordered on a scale* *extending from quiet to loud"* [Goldstein, 2001, 2005].

MDS *is Multidimensional scaling* [Schiffman et al., 2007].

Metric *is "the type of measuring system. The word is used very widely in different contexts which can be confusing. It is common in MDS to refer to metric and nonmetric solutions for the stimulus space. The distances in metric solutions preserve (as far as possible) the original similarity data in a linear fashion. The distances in* nonmetric *solutions preserve only the rank order of the original similarity data. The actual computation of the coordinates of the stimulus space is, of course, a metric (numerical) operation. Monotone transformations of the original similarity data provide the bridge between rank order and distances in the stimulus space."* [Schiffman et al., 2007].

MIDI *is an abbreviation of "Musical Instrument Digital Interface" and was designed for real-time control of music devices* [Roads et al., 1996].

Modality *should not be confused with modality in music! Modality is typically denoted the "sensation" (e.g. hearing, feeling, seeing, tasting, smelling)* [Martin and Bech, 2005].

Objective measurements - *in terms of an audio reproduction system - include e.g. "its physical descriptions as well as evaluations of perceived qualities such as loudness, brightness, harshness, or spaciousness"* [Martin and Bech, 2005].

Orthogonal *means "perpendicular to. Most MDS spaces are developed with orthogonal axes"* [Schiffman et al., 2007].

Perceptual attribute *is also known as "sensory attribute"* and is *"an objectively measurable description of a perceived quality of a stimulus without qualification of the subject's opinion of the description"* [Martin and Bech, 2005]

Perceptually Nonlinear Scale Bias *denote the bias in which "the scales used in listening tests are not perceptually linear". For example the respondent may not perceive the distance between labels of equal distances as being equal (e.g. there may be perceptually longer between "Fair" and "Poor" than between "Poor" and "Bad". This bias is very different internationally and may vary significantly between e.g. Italians on one side and Swedish and American English-speaking people on the other side* [Zielinski et al., 2007].

Physical description *can e.g. be SPL (sound pressure level)* [Martin and Bech, 2005].

Pitch *is defined as "that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale" and is related to the physical repetition rate of the waveform of a sound* [Goldstein, 2001, 2005]. "Variations in pitch create a sense of melody" [Moore, 2003].

Point *is "a position in a space that is an abstract representation of a stimulus"* [Schiffman et al., 2007].

Preference *can mean either:*

– *which of a number of stimuli is liked (preferred).*

– *selecting one stimuli over another.* [Martin and Bech, 2005; Zimmer and Ellermeier, 2003]

Program material - *in this report - is a term for the audio signal, which can be either pre recorded or synthesised* [Martin and Bech, 2005].

Random errors *are commonly observed in data from e.g. listening tests and are easily averaged out by calculating the mean value. These errors can be due to e.g. listeners' individual differences* [Zielinski et al., 2007].

Range Equalizing Bias *denote a bias in which the responded tend to stretch or compress the grading scale. "As a result the scores span the whole scale, regardless of the range of the stimuli"* [Zielinski et al., 2007].

Repertory Grid Technique (RGT) *is a technique designed to elicitate so-called personal constructs. The participant is exposed to randomised triads of the stimuli and asked to point out which stimuli stands out, why, and what makes the two other stimuli similar. The technique is continued for as long as the participant can come up with new constructs and does not repeat himself. Typically the RGT is followed up by a rating of the individual attributes elicited for each participant. More can be read by Fallman and Waterworth [2005], van de Kerkhof [n.d.], Karapanos and Martens [2008], Karapanos et al. [2009], Bradshaw et al. [1993], and Bell [2003].*

Sensation Level *is the intensity level with a reference to the participant's or subject's absolute threshold for the specific stimulus.*

Sensory attribute (see: Perceptual attribute).

Shepard Diagram. *"Also scattergram or scatter diagram. A plot comparing the distances derived by MDS and the transformed data (disparities) with the original data values or proximities."* [Schiffman et al., 2007]

Space *is a set of all potential point defined by a set of dimensions* [Schiffman et al., 2007].

Stimulus *equals that which is perceived or sensed using one or more modalities. The stimulus in perceptual audio evaluation is the sound presented for the participant* [Martin and Bech, 2005]. *It should be noted that a stimulus is not always perceived due to e.g. the sinusoid (stimulus) being below or above the subject's hearing threshold level.*

Stimulus Spacing Bias *is a bias in which the scores obtained in e.g. a listening test "will be spaced at more or less equal intervals, regardless of the distribution of the stimuli in the perceptual domain (i.e. internal judgements).* [Zielinski et al., 2007].

Stress *is "a particular measure that shows how far the data depart from the model. There are several stress formulas available in the various algorithms."* [Schiffman et al., 2007]

Subjective measurements is *"based on or influenced by personal feelings, tastes, or* *opinions"* [Martin and Bech, 2005]. A subjective measurement could be whether the subject prefers a stimulus with a higher loudness measure (objective measure) over a stimulus with a lower loudness measure.

Systematic Errors *are difficult to identify and are likely to cause a repeatable and consistent shift in the data. The errors may thus go unnoticed by researchers and are also difficult to eliminate using statistical postprocessing of data* [Zielinski et al., 2007].

Timbre *has been defined as "that attribute of auditory sensation in terms of which a listener can judge two sounds similarly presented and having the same loudness and pitch are dissimilar (...) Unlike pitch and loudness, which may be considered unidimensional, timbre is multidimensional; there is no single scale along which the timbres of different sounds can be compared and ordered* [Goldstein, 2001, 2005].

- time-variant timbral components:
- more global timbral components: Attributes of a whole sound event, rather than its components that are discriminable over time, e.g. tone coloration, "which may be more narrowly defined than is the term timbre" [Martens and Marui, 2002].

Tone Coloration: *"Though tone coloration is certainly not a unidimensional perceptual attribute, it can be described by a lower dimensional structure than timbre can be. (...Of orchestral instruments...) This perceptual dimension was associated with the verbal attribute "brightness", which was shown to depend more upon temporally-integrated spectral energy distribution than it does upon spectral evolution (...) Furthermore, the definition of tone coloration can be narrowed even more by excluding timbral diffeences due to the attack inharmonicity identified (...) as the physical correlate that can aid the human listener in differentiating between tones played by musical instruments belonging to different families (e.g. strings, woodwinds, and brass horns)."* [Martens and Marui, 2002].

Transformations. *"Nonmetric MDS programs apply* monotone *transformations to the original data to allow performance of arithmetic operations on the rank orders of proximities. A monotone transformation need only maintain the rank orders of proximities. The logarithm function is an example of a monotone transformation."* [Schiffman et al., 2007]

8.2 Multidimensional Scaling (MDS)

Multidimensional scaling is a mathematical procedure used for representing percetions. The output of a multidimensional scaling can be plotted in a stimulus space. A stimulus space containing the perceptual dimensions can *also* be created through factor analysis, discriminant analysis, and correspondence analysis, cluster analysis, multiple discriminant analysis, conjoint measurement [Cooper, 1983].

Indirect elicitation methods have been chosen to identify the degree of dissimilarity between perceptual attributes in high-gain guitar amplifiers since this method does not allow for a *facilitator bias*. A multidimensional scaling technique performed on dissimilarity neither requires the test participants to evaluate complex stimuli on - a priory knowledge-based - semantic differential scales. The only question that the participant in a MDS-based experiment is given, is to which degree (s)he perceives the stimuli pairs to be different. Multidimension scaling techniques of dissimilarity data are thus low in experimenter contamination [Schiffman et al., 2007].

Multidimensional scaling (MDS) is not a statistical method but rather a powerful mathematical procedure used in a large variety of application areas to systematise data. The MDS process represents similarities spatially as in a map [Schiffman et al., 2007]. In order to visually illustrate color perception, Ekman [1954] conducted a session, in which similarity measures among pairs of 14 colors varying in hue[1] were obtained [Ekman, 1954; Schiffman et al., 2007]. Ekman [1954] analysed his data using factor analysis [Schiffman et al., 2007]. Eight years later a MDS procedure was applied to Ekman's analyses [Schiffman et al., 2007]. This resulted in a geometrical representation in two dimensions of the 14 colors, which, if drawing the contour of the representations with a marker, approximates a circle (Figure 8.1) [Schiffman et al., 2007].

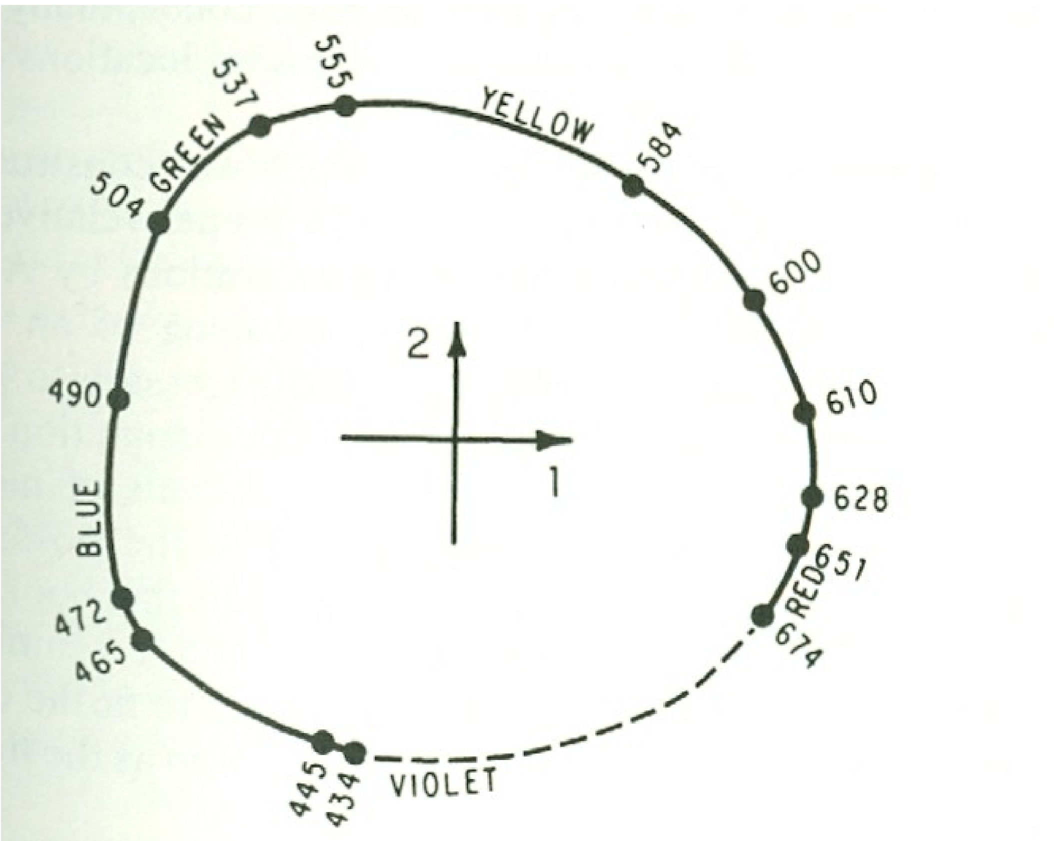

Figure 8.1: *Spatial representation of similarities in color perception derived by applying MDS to Ekman [1954]'s similarity judgments. Illustration by [Schiffman et al., 2007].*

[1]"Hue is the dominant wavelength in the perceived light and represents the pure color, i.e. the color located on the edges of the chromaticity plane." [Moeslund, 2009]

8.2.1 How a stimulus space is created

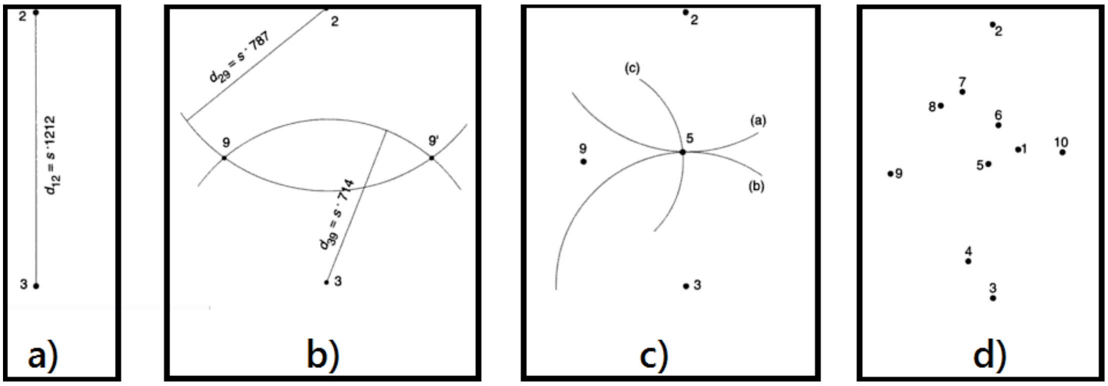

Figure 8.2: *Example of the construction steps in creating the stimulus space. a) The distance between point 2 and 3 are drawn. b) From each of point 2 and point 3, their respective distances to point 9 are drawn. c) Another point (point 5) is added. d) The final stimulus plot containing all points. Illustration modified from Borg and Groenen [2010].*

In Figure 8.2 an example of how a stimulus space can be created simply by a pencil, a ruler and a compass. This is in theory what a MDS method does iteratively, which is elaborated in the following example of classical MDS. It is not hard to measure the distances between several larger cities in North America when given a map. The distances between city A (x_{11}, x_{12}) and city B (x_{21}, x_{22}) can be measured euclidian by a ruler or the following equation:

$$d_{AB} = \sqrt{(x_{11}) - (x_{21})^2 + (x_{12}) - (x_{22})^2}$$

If one was to create a map of the cities in North America, knowing only the distances, a start would be to line all cities ordinally on a line (unidimensionally, as illustrated in Figure 8.3). Due to a large measure of error, iterations are hereafter performed as seen in Figure 8.4 and Figure 8.5 untill the representation begins to resemble a map of the cities in North America.

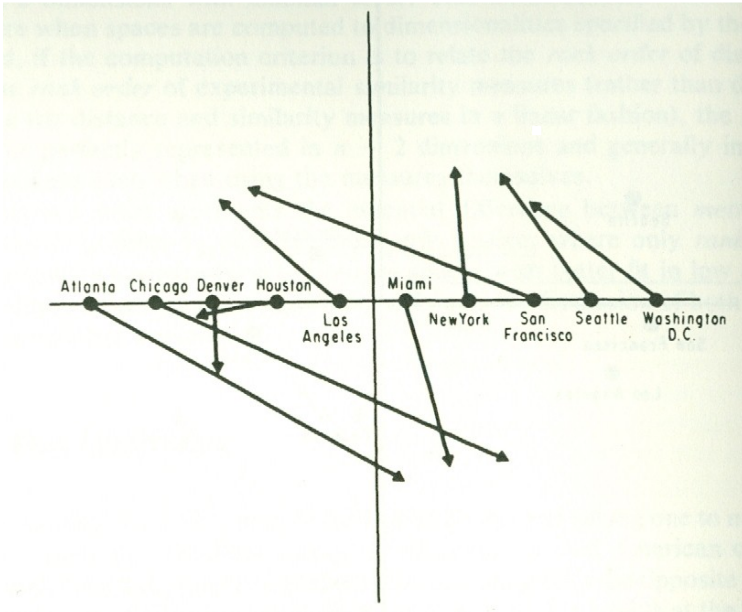

Figure 8.3: *First iteration of creating a stimulus space for the distances between cities in North America. The stimuli are placed on a line in only one dimension. The arrows points to where the stimulus points are moved by the specific MDS procedure in order to reduce the error. Illustration by [Schiffman et al., 2007].*

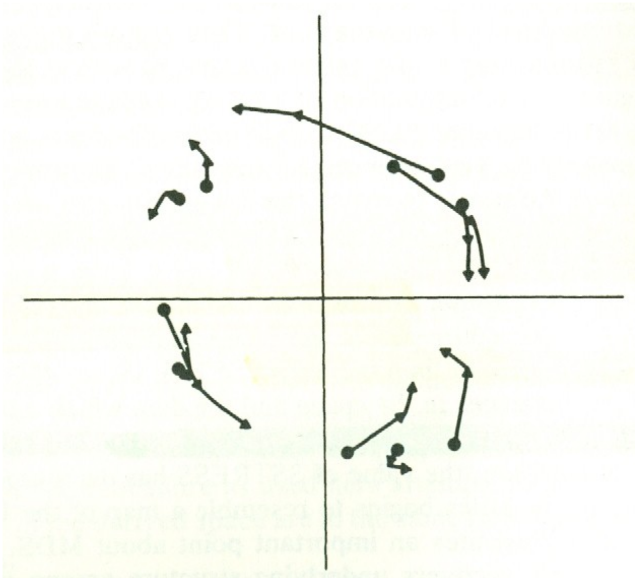

Figure 8.4: *Next iteration of creating the stimulus space for the distances between cities in North America. Illustration by [Schiffman et al., 2007].*

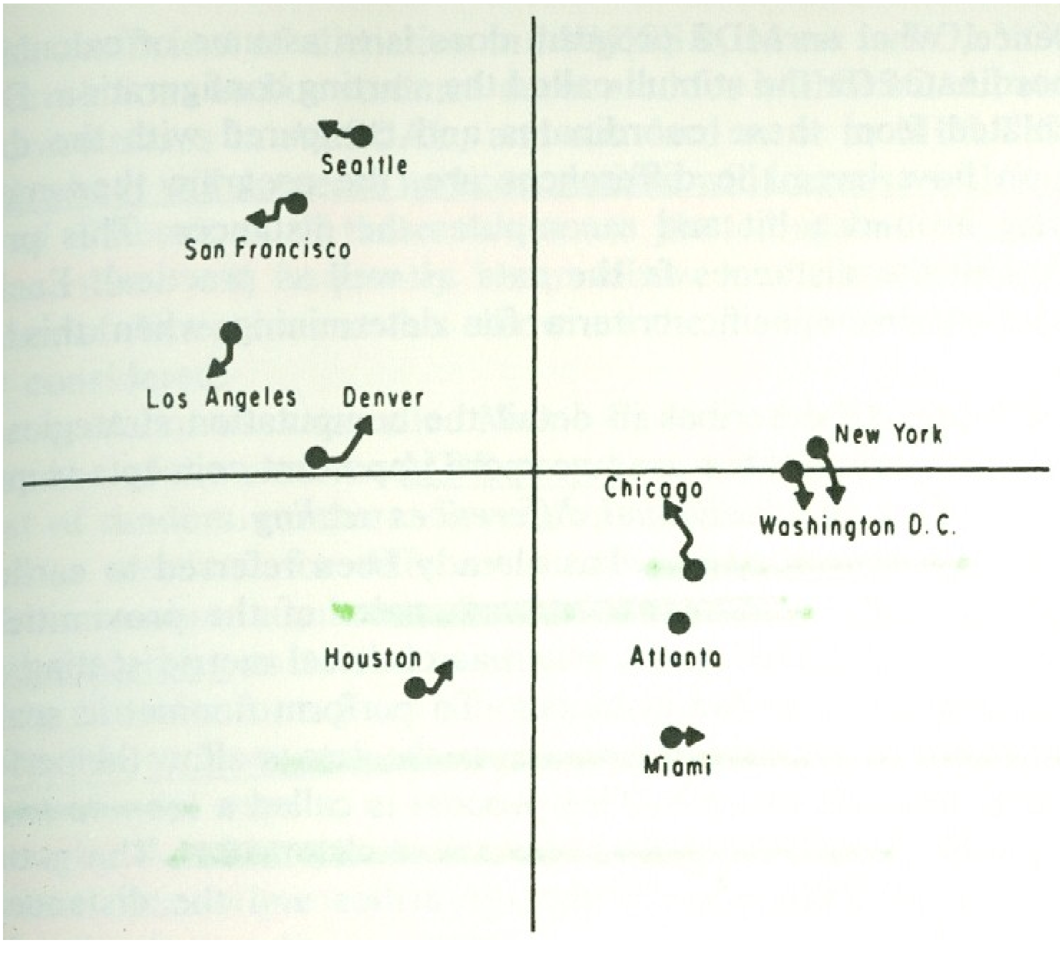

Figure 8.5: *Last iteration before ending up with the final stimulus space (in this case a map of North American cities). Illustration by [Schiffman et al., 2007].*

8.2.2 Evaluating Stress

Stress is a measure similar to 'Goodness of fit' and the lower the stress - the better the model fits the data. Chosing to represent the stimulus plot in too many dimensions may blur the structure due to overfitting noise components [Borg and Groenen, 2010]. On the other hand scaling with too few dimensions may overcompress the structure this distort the true MDS structure. In 1964, Kruskal proposed the following benchmarks for ordinal MDS based on his "experience with experimental and syntehtic data":

.20 = poor

.10 = fair

.05 = good

.025 = excellent

.00 = perfect

A more thorough critique of these benchmark and the use of them is written in [Borg and Groenen, 2010].

8.2.3 Combining MDS with other methods

Meulman et al. [1986] and Ramsay [1980] have provided methods which - when combined - largely circumvent the limitations of MDS and semantic differential scaling Martens and Marui [2002]. These methods include joint analyses and external unfolding. Martens and Marui [2002] argue as follows:

> "*MDS analyses of dissimilarity ratings are often included in investigations of complex perceptual phenomena in order to indicate the involvement of stimulus parameters for which direct ratings might not be collected. Conversely, the dissimilarity-based perceptual space can reveal which stimulus parameters do not enter into the listener's global evaluative reactions. Also, wide variation in ratings on a particular adjective scale might no correspond to large perceptual differences. Therefore, while adjective ratings aid in the interpretation of the MDS-derived perceptual dimensions, the dissimilarity ratings aid in identifying which of the adjective ratings scales correspond to salient perceptual attributes.*"

Furthermore, Martens and Marui [2002] describe that by analysing the MDS output there is a basis for determining how perceptually distinguishable the stimuli are from each other with respect to the adjective scale response values.

8.3 Perceptual Audio Evaluation

As technologies within audio and music develop, it is increasingly essential to evaluate these technologies both technically and perceptually [Bech and Zacharov, 2006] as human listeners are the ultimate judge of sound quality [Zimmer and Ellermeier, 2003]. [Martin and Bech, 2005] argues that traditional measures of audio systems in general provide an incomplete representation of a system's overall perceived "sound". Any given perceived "sound" is often said to be of "subjective" quality. However, the term "subjective testing" appears to induce confusion, which has led to clarification of commonly used definitions relating to subjective and perceptual testing of i.a. automotive audio systems [Martin and Bech, 2005].Figure 8.6 illustrates the so-called filter model. Among methods for identification and quantification of perceptual attributes in automotive audio systems is the Descriptive Analysis technique [Martin and Bech, 2005].

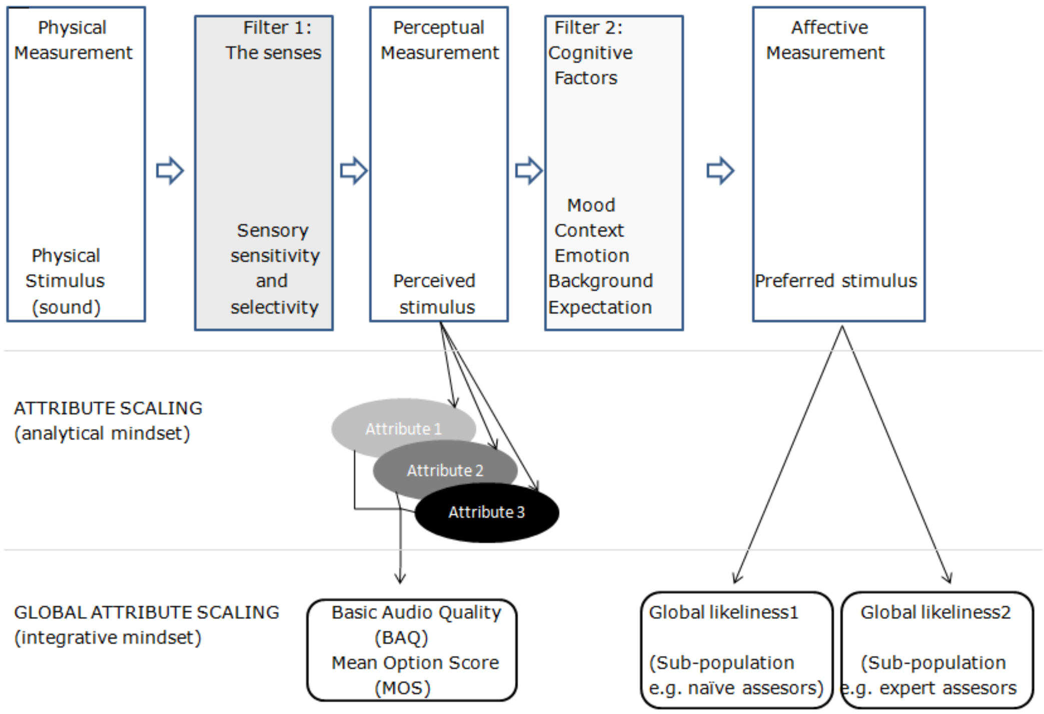

Figure 8.6: *Illustration of the so-called filter model that show the division between analytical and integrative mindset. Re-illustrated with reference to the filter model in [Bech and Zacharov, 2006].*

8.3.1 Direct Scaling

Level of measurement

The measurement of the response variable can be of the following four levels:

- **Nominal scale:** Variables can be either *equal* or *inequal*, which means that any variable can be compared one-to-one. "Yes" and "no" responses are examples of nominal data.

- **Ordinal scale:** Variables can be either *greater than* or *less than* each other and are monotonically increasing (e.g. Product A preference $<$ Product B preference $<$ Product C preference) but there is *no measurable meaning* to the "difference" between responses.
- **Interval scale:** Variables can be responses on an *ordered scale* where there is *meaning* to the difference between measurements. Unlike ratio scale data, a person being [18-20] years old is not necessarily half as old as a person being [36-40] years old.
- **Ratio scale:** Variables are rank ordered and indicate both rank and distance from a natural zero. A person can e.g. weigh 80 kg which is twice the weight of a person weighing 40 kg.

[Bech and Zacharov, 2006; Newbold et al., 2006]

According to Bech and Zacharov [2006], debate exists whether the type and degree of statistical analysis is linked to the level of measurement [Bech and Zacharov, 2006]. However, data based on scales *assumed* to be interval scales is analysed by quantitative statistics unless severe violations of the statistical assumptions or scale properties are present. If certain statistical assumptions are not validated, data can be analysed using methods based on categorial analyses, e.g. contingency tables. Valid conclusions can be drawn by comparing the results of quantitative statistics with the results of the categorical based methods [Bech and Zacharov, 2006].

Partition scaling

- The **equisection method** presents the listener to a group of stimuli and asks him (or her) to select a limited number of stimuli produing equidistant[2] sensations on the attribute of interest.
- The **category scaling method** asks the listener to assign a category to each stimuli in the presented stimuli set. Categories (or labels) are determined by the experimenter. According to Bech and Zacharov [2006] this method is often used in evaluations of audio.

Ratio scaling

- **Ratio production:** The listener's task is to adjust the magnitude of a manipulatable stimulus to be equal to a prescribed ratio of a reference stimulus [Bech and Zacharov, 2006].
- **Ratio estimation:** The listener's task is to describe the ratio between two stimuli for the attribute of interest [Bech and Zacharov, 2006].
- **Magnitude estimation:** The listener is asked to assign a number representing the sensory magnitude of the stimulus. When the listener perceives the initial stimulus he

[2] *"equally far from two or more places"* [Hornsby, 2000]

is told the magnitude of this. After perceiving a new stimulus he is asked to report the magnitude of this relative to the magnitude of the initial stimulus. A variation of this method is where the listener is first asked to assign a magnitude value to the initial stimulus before reporting the magnitude of the new stimulus relative to the initial stimulus [Bech and Zacharov, 2006].

- **Magnitude production:** The listener's task is to adjust the auditory sensation of a manipulatable stimulus to a magnitude specified on forehand [Bech and Zacharov, 2006].

Bech and Zacharov [2006] recommends a standardised scale for allowing the listener to report his - or her - perceived degree of the response variable.

Direct elicitation methods assume a close relationship existing between a given sensation and the verbal descriptors used by the test participant to describe the sensation [Bech and Zacharov, 2006].

> *Consensus vocabulary techniques* rely on a commong terminology developed and agreed upon by a group of highly trained subjects [Bech and Zacharov, 2006].
>
> *Individual Vocabulary techniques* use the vocabularies developed by the individual subject and a set of principal components representing the common attributes is then identified using statistical procedures [Bech and Zacharov, 2006].

Indirect elicitation methods focus on seperating the stimulus and its verbal discriptor and count methods as MDS, drawing, and perceptual structure analysis (PSA) [Bech and Zacharov, 2006]. As verbalisation *can* depend significantly on the 'size' and availability of suitable terms in the subject's lexicon, the indirect elicitation methods try to separate sensation and verbalisation [Bech and Zacharov, 2006].

Wickelmaier [2005] argues for three disadvantages of direct scaling methods:

1. They demand expertise and training.
2. They involve the risk of biasing the judgments by a priori presenting verbal categories.
3. They prevent the detection of latent and unlabeled auditory attributes.

8.3.2 Indirect scaling

Assumption

The amount of times a stimulus is judged different from another stimulus is directly related to the degree by which the two sensations are different [Bech and Zacharov, 2006].

Measure

The basic measure is the probability of the two stimuli being considered different. Averaging many presentations provides the estimated probability. Basic methods within indirect scaling are difference threshold (e.g. JND) and paired comparisons [Bech and Zacharov, 2006].

Advantage of indirect scaling

Asking subjects to assign a number to an impression can be a complicated problem but is avoided if employing methods of indirect scaling [Bech and Zacharov, 2006]. Furthermore, indirect scaling methods - such as e.g. multidimensional scaling - is not biased by the experimental design, facilitator, or test conductor [Schiffman et al., 2007]. However, the is less efficient compared to direct scaling of attributes.

[Bech and Zacharov, 2006] argues that the general acceptance of complex stimuli such as audio being a multidimensional problem. This multidimensional problem includes a number of individual auditory attributes. These auditory attributes can be identified and elicited through i.a. interviews, experience, experiments, multidimensional scaling, and multivariate techniques[3] [Bech and Zacharov, 2006]. In order to elicit perceptual attributes - and quantify the users' impressions - both *direct* and *indirect* elicitation methods can be utilised.

Wickelmaier [2005] argues that indirect scaling methods require only simple qualitative judgments which is easier for the test participant to answer. The task to be performed by the test participant could be to judge which of two stimuli is greater than the other with respect to a certain attribute. The numerical representation depends on certain structural conditions, e.g. transitivity. Testing for transitivity can conclude whether judgements are even ordinal, e.g. if C is larger than B and B is larger than A, C must be larger than A. According to Zimmer and Ellermeier [2003] the validity of e.g. ratio scaling and magnitude estimation is unkown whereas indirect scaling methods generally are seen as being more valid.

Chi^2 tests for the possibility of the result being by chance, i.e. has the respondent chosen the loudest stimuli of A and B by chance without being able to tell a difference in loudness. Three degrees of stochastic transitivity have been defined:

> Weak Stochastic Transitivity (WST) evaluates the consistency of the cumulative matrix, i.e. pooled data. If consistent, the stimuli can be ordered ordinal on a unidimensional scale. $a > b > c \Rightarrow a > c$
>
> Moderate Stochastic Transitivity (MST) holds if $p_{ac} \geq min(p_{ab}; p_{bc})$
>
> Strong Stochastic Transitivity (SST) holds if $p_{ac} \geq max(p_{ab}; p_{bc})$

Zimmer and Ellermeier [2003] suggest the use of the following unidimensional models under the described conditions:

> If SST holds, the Bradley, Terry, and Luce (BTL) model is an accurate model:

$$p_{(ab)} = \frac{v(a)}{v(a) + v(b)}$$

[3] Multivariate means 'many variables'. An example of a multivariate technique is MANOVA (multivariate analysis of variance), which can be considered an ANOVA for experimental designs having several dependent variables [Field, 2009].

where v(a) denotes the scale value for a.

If SST does not hold, but MST is ok, then a preference tree is still an accurate model.

Zimmer and Ellermeier [2003] conclude that contrary to direct scaling methods choice models are falsible and do not generate an outcome by default. Although harder to interpret, the validity is higher in indirect scaling procedures. Zimmer and Ellermeier [2003] also argue that the choice models can be a helpful complementary methods to the technique: *Multidimensional Scaling* (MDS) where the latter focuses not only on fitting the best prediction, the former is more concerned about the cognitive processes involved with differentiating attributes' characteristics.

8.4 Table of Sound level, Intensity, and Pressure

Table 8.1: *The relationships between Sound level, Intensity, Pressure and their exampled description (re-illustrated from [Moore, 2003]).*

Sound level (dB SPL)	Intensity ratio (I/I_0)	Pressure ratio (P/P_0)	Typical Example
140	10^{14}	10^7	Gunshot at close range
120	10^{12}	10^6	Loud rock group
100	10^{10}	10^5	Shouting at close range
80	10^8	10^4	Busy street
70	10^7	3160	Normal conversation
50	10^5	316	Quiet conversation
30	10^3	31.6	Soft whisper
20	10^2	10	Country area at night
6.5	4.5	2.1	Mean absolute threshold at 1 kHz
3	2	1.4	
0	1	1	Reference level
-10	0.1		

8.5 Interview with a professional music producer

An interview was conducted with the professional recording engineer who helped facilitate the recording of the stimuli described in Section 2.2. The recording engineer is seen monitoring the recording process in Figure 8.7.

Figure 8.7: *Photograph of the professional recording engineer who helped facilitate the recording of the stimuli.*

Would you please state your responsibilities at the recording studio in which you work?

> "I'm in charge of the final product and the proces leading up to it. The point is to make the record sound technically and emotionally suiting for the material the band brings to the studio. Thus, I do everything from vocal-coaching to microphone-placement. The most significant part, however, is probably the mixing-proces in which the most hearable creative choices are made."

What got you started in sound engineering, production, and mixing?

> "A growing interest in music in general combined with a visit to a recording studio some ten years ago. It got to a point where I couldn't listen to music, without imagining what must have been going on in the studio at the time of recording."

Years of experience?

> "I've been recording music for 10 years. Professionally for about 6."

How often do you record high-gain guitar amplifiers in the studio?

> "A couple of times a month."

Could you please describe the process, the mics, and the signal chain that you use for recording high-gain guitar amplifier? (and why)

“Usually, I begin by talking to the client about what it is they want, and why. If, for some reason, I disagree with this, I try to explain why I think this or that would be better - or more fitting for the record. Once we agree to this in theory, we begin listening to setups. Usually, we start off with the clients usual (or live) setup, and try to figure out if that it suiting for the current record. From this point, there’s a lot of tweaking going on until we decide it’s time to hear it on the record. From what I’ve learned during the conversation and tweaking, I choose a signal path. Usually, the hints that really get me going are very subliminal; a facial expression, increasement in will to play or a musical reference made in another context. Most ’common’ (untrained in studio engineering, that is) have a hard time describing what they want, and what they like. So I’ve somewhat developed a talent for figuring this part out when they don’t expect it.

Thus, I use many different microphones, preamps, compressors and so forth. However, personally, I like to keep it down to one microphone (to reduce phase-issues), dynamic (to reduce extensive high-freq response), or even ribbon (to actively reduce hi-freq response). Not rarely is this a shure SM57, a sennheiser 606/609, a Sennheiser MD 421 or a T-Bone RB 500 ribbon-mic. Among other mics I occasionally use on guitars are Electro Voice PL-20 (aka RE 20), Shure SM7B, Beyer Dynamics M55, Sennheiser MD 441, AKG "The Tube" (Large diaphragm condenser) Superlux FS6 and Sontronics Helios (Large diaphragm condenser).

Preamps vary alot, but among others I use; (Listed roughly after frequency of use) Midas Pro 40 T channel strip (Preamp+eq / old console), Gyraf (gyratech) G9 (Made by BD labs), JoeMeek ThreeQ Channelstrip, GA Pre-73, Studio Projects VBT-1, Calrec UA-8000 channel strip (Preamp+eq / old console).”

What are the advantages and disadvantages of recording the way you have just described?

“I figure thats stated in the above. Otherwise, I’ll gladly expand. ”

Are there processes in the mixing of the recorded high-gain guitar that influence the original sound of the recorded guitar significantly please elaborate.

“Some productions include digital preamps (or amp-simulators). Personally, I use this extremely rarely on hi-gain guitars, but in the industry a such, it’s common practice. This process greatly alters the sound - even to a point where you wouldn’t be able to recognize the original recording. This could even make a clean blues-like recording sound like a heavily gained guitar.

A normal mix-chain for me would be:

Gate/expander - simply a tool for cutting out the noise when the guitar is not playing. Expansion can also be used for a better signal-to-noise-ratio. High-gain guitars are very noisy, especially when they are not being played, so in passages where there are a lot of breaks or rapid dynamic shifts going on, expanding the source can really help clean up the signal. Meanwhile, when the guitar is actually playing, it should remain unaffected by this process.

EQ - cutting below 70-80 Hz in a steep slope to reduce rumble and get rid of subharmonic noise. Cutting above 5k-10k Hz in a softer slope to reduce extensive overtones caused by the amount of gain in a hi-gain guitar recording. Besides this, there might be some smaller adjustments (boosting and cutting) to make the guitar 'sit' right in the mix. I.e. giving presence, reducing rumble, cutting standing waves and so on.

The EQ process can definitely be heard, but the source should remain easily recognizable during this.

Compression - mostly bringing the guitar 'out of the speakers', giving more presence, thickness and definition to the signal. Unless extensive, this process shouldn't be very obvious, but helps the guitar fit into the mix and the music.

FX - I'll call the last category fx, and I'm hereby including all processes that actually adds something to the signal that wasn't there in the first place. (EQ and compression only alters the signal). Most common are so called 'wet' effects such as reverb, delay, flanger, chorus, wah, flutter, wow and leslie. However these are all very obvious and are usually used occasionally, so they don't add to the overall sound as much as to specific passages. However, personally, I like to use tape-saturation (either real or simulated), to generate overtones and help blend the guitar with the other instruments. Which leads us to:

Full mix. The final guitar sound of a professional recording is inseparable from the rest of the instruments. Therefore, we have to consider the guitars relation to the bass, the drums the vocals and what else might be in the mix. In terms of bass, it's the bottom of the guitars that are commonplace for the two. The bass and the guitars need to be adjusted in relation to one another in terms of equalization and placement. Thus, the guitars usually ends up being panned hard L/R, while the bass stays centered. To expand this separation, a stereo expander might be used. Also, it's common to have a compressor 'ducking' the guitars (and the bass) in relation to the kick, so that this stands out more. This is also known as compression side-chaining. Sometimes this type of compression even goes on on the master channel."

8.6 The online survey regarding guitar amplifier settings

The online survey designed and programmed to investigate how real owners of the recorded amplifers use the settings. The online survey is available at http://www.mortenpurup.dk/ampsurvey/. A representative screendump of the survey is seen in Figure 8.8.

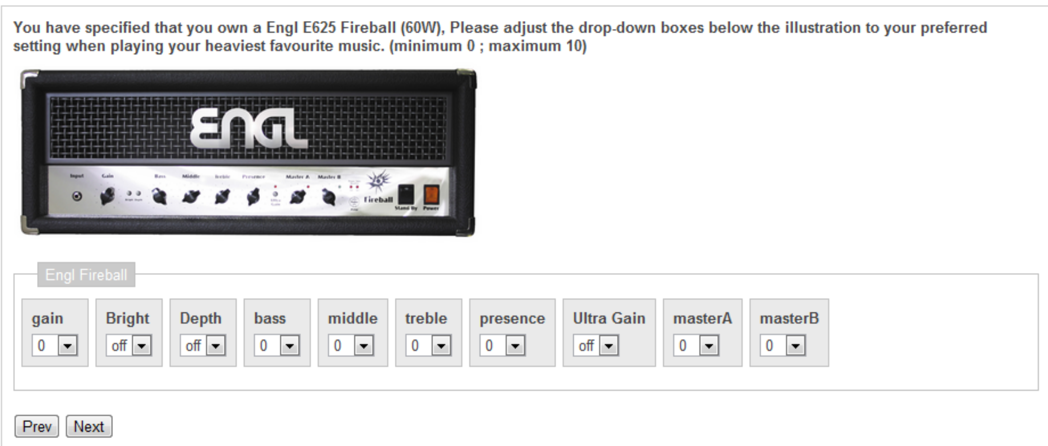

Figure 8.8: *Screendump example of the online survey.*

8.7 Recorded settings for each amplifier

A complete list of the differences in recorded settings for each guitar is listed in Table 8.3 and 8.4. The photographs of the main settings recorded for each amplifier are listed on the following pages:

Figure 8.9: *The settings for the Mesa Boogie Dual Recifier. The only channel photographed was the recorded channel with the highest amount of nonlinear distortion.*

Figure 8.10: *Photograph of the back of the Mesa Boogie Dual Rectifier where the settings "Vacuum Tubes / Silicon Diodes" and "Bold / Spongy" could be manipulated.*

Figure 8.11: *Photograph of the main settings for the ENGL e625 Fireball guitar amplifier.*

Figure 8.12: *Photograph of the main settings for the ENGL e635 Fireball guitar amplifier.*

Figure 8.13: *Photograph of the main settings recorded for the Peavey 6505+ guitar amplifier.*

Figure 8.14: *Photograph of the main settings recorded for the Randall Kirk Hammet RM100H guitar amplifier. Only the channel with the highest degree of nonlinear distortion was recorded and photographed.*

Figure 8.15: *Photograph of the main settings recorded for the Marshall TSL60 guitar amplifier. Only the channel with the highest degree of nonlinear distortion was recorded and photographed.*

Figure 8.16: *Photograph of the recorded channel on the Marshall MA50H guitar amplifier.*

8.8 Program Material

Table 8.2: *Table showing the coherent recording name for each stimulus of the program material. The amplifier settings for the recordings are seen in Table 8.3 and Table 8.4 in Appendix (Chapter 8).*

Program Material	Recording Name
Amplifier 1	mesa2Gibson.wav
Amplifier 2	engl602Gibson.wav
Amplifier 3	engl1002Gibson.wav
Amplifier 4	marshallMA50H1Gibson.wav
Amplifier 5	marshallTSL602Gibson.wav
Amplifier 6	peavey65051Gibson.wav
Amplifier 7	randallKH2Gibson.wav
Amplifier 8	engl1005Gibson.wav

Table 8.3: *Table showing the differences between all recordings of each guitar amplifier.*

Recording Name:	Remark 1:	Remark 2:	Remark 3:
Mesa1Gibson.wav	Bold	Vacuum tubes	
Mesa2Gibson.wav	Bold	Silicon tubes	
Mesa3Gibson.wav	Spongy	Vacuum tubes	Gave lower output than the two above, which were quite similar in amplitude.
Mesa4Gibson	Spongy	Silicon tubes	Gave lower output similar to Mesa3Gibson.wav
Engl601Gibson.wav	Bright: On	Depth: On	Ultra Gain: On
Engl602Gibson.wav	Bright: On	Depth: Off	Ultra Gain: On
Engl603Gibson.wav	Bright: Off	Depth: On	Ultra Gain: On
Engl602Gibson.wav	Bright: Off	Depth: Off	Ultra Gain: On
Engl1001Gibson.wav	Bright: On	Buttom: On	Mid Boost: Off
	Utra Gain: On		Volume adjusted to SPL dB *and* adjusted at the meter (-6 dB FS)
Engl1002Gibson.wav	Bright: On	Buttom: Off	Mid Boost: Off
	Utra Gain: On		
Engl1003Gibson.wav	Bright: Off	Buttom: Off	Mid Boost: Off
	Utra Gain: On		
Engl1004Gibson.wav	Bright: On	Buttom: On	Mid Boost: On
	Utra Gain: On		Ibanez MMM1 caused clipping*
Engl1005Gibson.wav	Bright: On	Buttom: Off	Mid Boost: On
	Utra Gain: On		
Engl1006Gibson.wav	Bright: Off	Buttom: Off	Mid Boost: On
	Utra Gain: On		

**The clipping was overcome by switching input channel and decreasing gain for the preamp in order to match the input level for the other settings and make the change in input gain transparent. Midboost increased the peak amplitude by approximately 6 dB measured the input level of the DAW (digital audio workstation).*

Table 8.4: *Table showing the differences between all recordings of each guitar amplifier (continued from last page).*

Recording Name:	Remark 1:	Remark 2:	Remark 3:
Peavey65051Gibson.wav			was only recorded with one setting
randallKH1Gibson.wav	Bright: Off		
randallKH2Gibson.wav	Bright: On		
marshallTSL601Gibson.wav	Deep: On OD	Lead Shift: Off OD2	
marshallTSL602Gibson.wav	Deep: Off OD	Lead Shift: Off OD2	(decreases the input level measured at level meter)
marshallTSL603Gibson.wav	Deep: Off OD	Lead Shift: On OD2	(decreases the input level even further
marshallTSL604Gibson.wav	Deep: On OD	Lead Shift: On OD2	
marshallMA50HGibson.wav			was only recorded with one setting

8.9 Nonlinear Distortion

Moore [2003] has a great argumentation on what is known within the psychology of hearing as well as describing and debating the certainty of these theories. One cannot adress linear distortion except - when talking about a change in the frequncy spectre. The output of a auditive system must be discussed as a result of adding input and something extra[4]:

$$y(n) = A_0 + A_1 sin(2\pi f_1 Tn) + A_2 sin(2 * 2\pi f_1 Tn) + ... + A_N sin(N * 2\pi f_1 Tn),$$

if the input signal is a sinusoid of known amplitude and frequency according to $x(n) = Asin(2\pi f_1 * Tn)$.

[4]By something extra, something not present in the input is thought of.

8.10 Audiometry Procedure used for Participants

All participants in the loudness equalisation test as well as the listening experiment were initially tested for hearing impairments. The audiometry tests were conducted in a small listening cabin designed with subjective audio evaluations in mind. The cabin is a double-door constructions and has sound isolating walls, floor and ceiling. The room consists of a steel box resting on springs for anti-vibrating purposes.

The participant was positioned in a chair pointing away from the facilitator in order to eliminate any visual bias or other visual cues. The participant was asked to press a button for each perceived tone. The button press lit up a diode on the control panel monitored by the facilitator. The following procedure was followed for each ear of participant, starting with the right ear:

1. First each participant were exposed to a 1 kHz tone of -10 dB.
2. The amplitude of the tone was then adjusted by the facilitator until a button press was observed.
3. After assuring that the participant could not perceive this tone played with an amplitude of 5 dB lower than the amplitude observed, the amplitude value was written down for the specific frequency.
4. Next, step 1 through 3 was performed for a 2 kHz tone.
5. Next, step 1 through 3 was performed for a 4 kHz tone.
6. Next, step 1 through 3 was performed for a 8 kHz tone.
7. Next, step 1 through 3 was performed for a 500 Hz tone.
8. Finally, step 1 through 3 was performed for a 1 kHz tone again to check for consistency.

8.11 Principal Component Analysis

> "*The basic aim of principal component analysis (PCA) is to describe variation in a set of correlated variables, x_1, x_2,...,x_q, in terms of a new set of uncorrelated variables, y_1, y_2,...,y_q, each of which is a linear combination of the correlated variables. The new variables are derived in decreasing order of 'importance' in the sense that y_1* accounts for as much of the variation in the original data amongst all linear combinations of x_1, x_2,...,x_q. Then y_2 is chosen to account for as much as possible of the remaining variation, subject to being uncorrelated with y_1 - and so on, i.e., forming an orthogonal coordination system. The new variables defines by this process, y_1, y_2,...,y_q, are the principal components. The general hope of principal component ananlysis is that the first few components will account for a substantial proportion of the variation in the original variables, x_1, x_2,...,x_q, and can, consequently ne used to provide

a convenient lower-dimensional summary of these variables that might prove useful for a variety of reasons." [Everitt and Hothorn, 2010]

8.12 Theory of Ekman [1954]'s method

"The theory of the method is based upon the reasonable assumption that the degree of perceived similarity is a function of the degree of overlap between those primary experiences (sensations, emotions) which are evoked by the stimuli. Under certain condictions the methods of factor analysis may be directly applied to the similarity matrix. The entries of the resulting factor matrix indicate the relative contribution of any primary experience to the complex experience under consideration." [Ekman, 1954]

8.13 Loudness Calibration

Due to risk of biasing the respondents' preference for the program material a loudness calibration test was conducted. 13 participants - all except one were graduate students from Aalborg University - took part in the test. None of the participants took part in the actual listening test. Initially an audiometric test was conducted for each subject before moving on to the loudness calibration. In one speaker, the participant could hear the reference stimuli. A pilot test conducted before the loudness calibration test revealed that a stimuli was preferred as a reference over pink noise. In the other speaker the participant was asked to adjust the level of the stimuli to an equal loudness of the reference stimuli. When the participant had adjusted and approved the loudness of the first stimulus, the track playing the stimulus was muted and a second stimulus was adjusted and approved. This procedure was conducted for all stimuli using the first stimuli (sound 1) as a reference. The coherent amplifier for sound 1 through 8 is seen in Table 8.2 in Chapter 2.

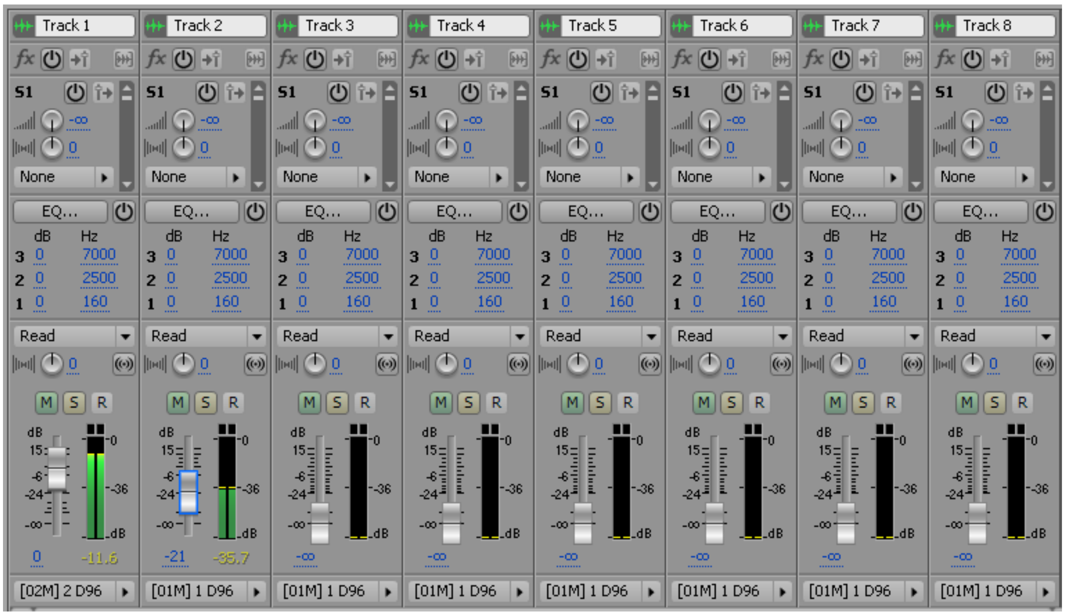

Figure 8.17: *Screendump of the user interface in Adobe® Audition® 3 where the participants moved the slider to adjust Track 2 (sound 2) to be equally loud as Track 1 (reference: sound 1). When the loudness of each track was adjusted and approved by the participant, the track was muted (clicking M) and the loudness of the next track was adjusted.*

Table 8.5: *Table showing means and standard deviations of loudness differences between the reference (sound 1) and sound 1 through 8.*

dB deviation	Mean:	SD:
Amplifier 1	0 dB	0 dB
Amplifier 2	-0.225 dB	2.482 dB
Amplifier 3	3.375 dB	1.708 dB
Amplifier 4	-0.403 dB	2.035 dB
Amplifier 5	3.425 dB	2.481 dB
Amplifier 6	-0.883 dB	1.985 dB
Amplifier 7	1.358 dB	2.259 dB
Amplifier 8	-0.033 dB	2.984 dB

Final sound pressure levels of the program material measured from the listening position are seen in Table 8.6. The levels are measure with a Monacor® SM-4 Sound Level Meter.

Table 8.6: *Table showing final sound pressure levels for the program material measured at the speaker and in the listening position*

Program Material	$SPL_{speaker}$	$SPL_{listener}$
Amplifier 1	83.9 dB	70.2 dB
Amplifier 2	76.2 dB	68.2 dB
Amplifier 3	82.0 dB	68.3 dB
Amplifier 4	71.1 dB	64.4 dB
Amplifier 5	80.7 dB	68.9 dB
Amplifier 6	79.9 dB	67.0 dB
Amplifier 7	80.9 dB	67.1 dB
Amplifier 8	81.8 dB	69.1 dB

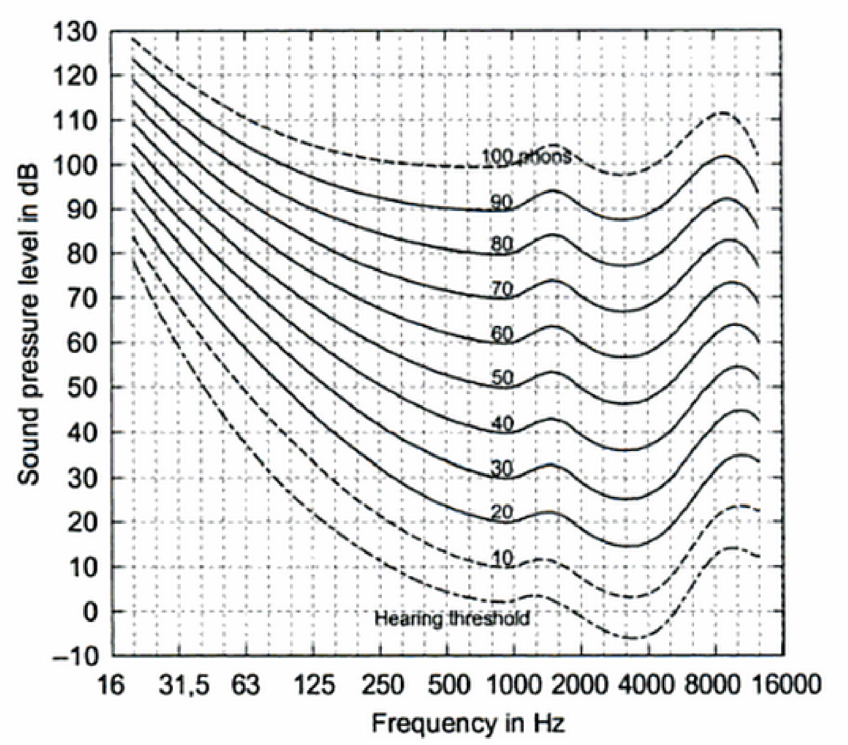

Figure 8.18: *Equal loudness curves showing and illustrating the relationship between dB SPL and frequency in Hertz for perceived equal loudness [Moore, 2003].*

8.14 Making participants understand the test

As an introduction to the participants the following text and illustrations were presented before initialising the given experiment.

8.14.1 Presenting the sounds / Preference (1st session)

You will hear all the sounds in the first step of the experiment. After you have heard them all you will be asked to rate them according to your degree of liking. You will be giving an amount of stars to each sound. You may give the same amount of stars to more than one sound.

Figure 8.19: *Screendump of the application for rating preference.*

I would like you to remember that all people judge sound differently. There are no right or wrong answers. All results are important as I am interested in knowing exactlyhow *you* hear these sounds.

8.14.2 Repertory Grid Interview

Dear participant,

In this experiment you have to name exactly what characterises the sound of different high-gain guitar amplifiers you are presented with. There is no right or wrong answers. I am interested in exactly how *you* perceive the characteristics of the sounds. You will be presented with three sounds at a time (A, B, and C) and asked the following three questions:

- Which sound is different from the other two?
- In one word, what makes this sound stand out?
- In one word, what makes the two other sounds similar?

Note that you might find there to be several characteristics making the sounds sound different or similar. I would like you to choose the characteristic which you find strongest. Please write in English. If you can not come up with the name in English but you know the exact word in Danish, please write the Danish word. You are allowed to ask questions during the test if you get stuck.

To illustrate how you can answer these questions, I present the following example:

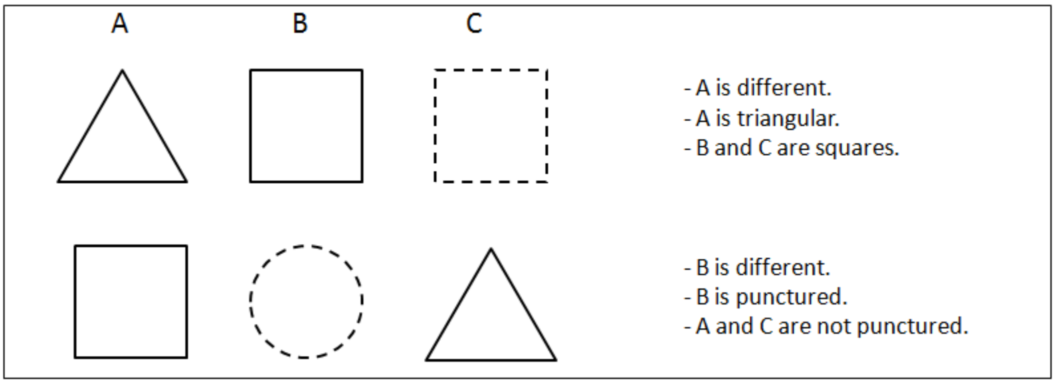

Figure 8.20: *Example of how to answer the questions in the Repertory Grid Interview. The stimuli are illustrated in simple geometrical shapes to enhance the understanding of the task.*

Any questions?

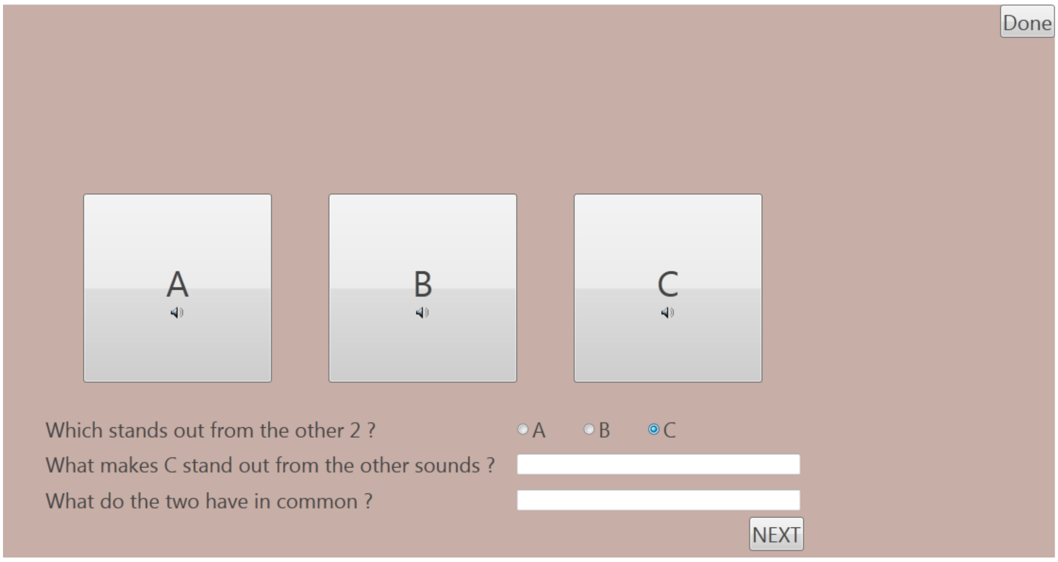

Figure 8.21: *Screendump of the GUI developed for the RGT task in session 1.*

8.14.3 Presenting the sounds / Preference (2nd session)

You will hear all the sounds in the first step of the experiment. After you have heard them all you will be asked to rate them according to your degree of liking. You will be giving an amount of stars to each sound. You may give the same amount of stars to more than one sound.

Figure 8.22: *Screendump of the application for rating preference.*

I would like you to remember that all people judge sound differently. There are no right or wrong answers. All results are important as I am interested in knowing exactly how *you* hear these sounds.

8.14.4 Dissimilarity Rating

In this experiment you have to judge the degree of similarity or dissimilarity between sounds of high-gain guitar amplifiers. You will give your judgement on the computer in front of you.

On the screen you see a scale with seven points. You will be presented with two sounds at a time and asked to judge the dissimilarity between these two sounds on the scale. In this figure, you see the scale and the buttons for playing the two sounds:

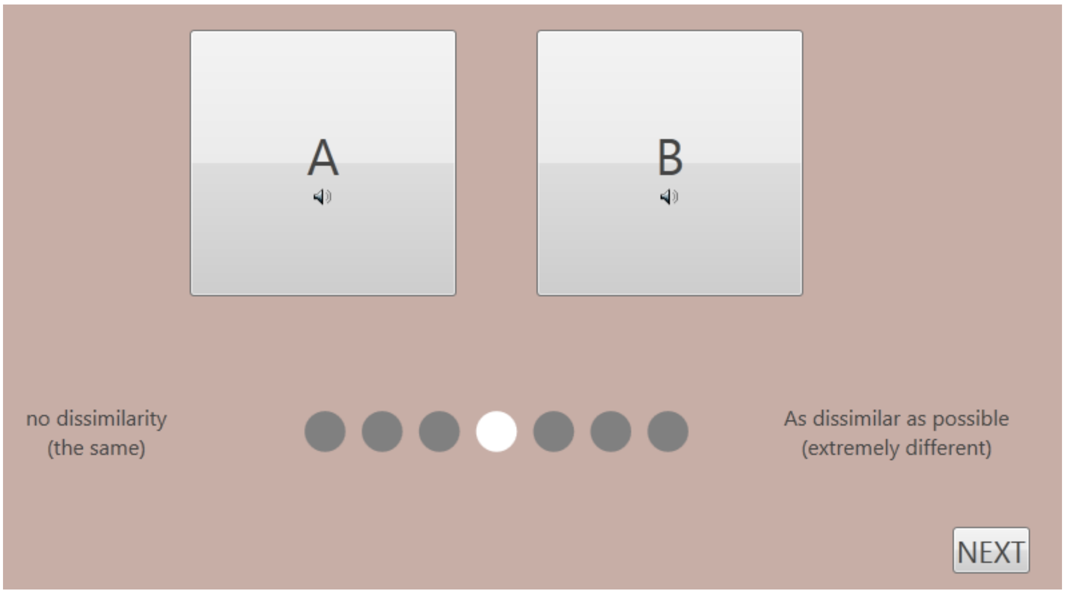

Figure 8.23: *Screendump of the application for judging dissimilarity.*

If you think there is no difference between between the two sounds you select the scale point to the left saying '*no dissimilarity (the same)*'. If you think there is a difference, you select a scale point on the scale indicating the degree of dissimilarity.

I would like you to remember that all people judge sound differently. There are no right or wrong answers. Two sounds may sound very much alike to one person and another person will hear them as being different. Both results are important as I am interested in knowing exactly how *you* hear these sounds.

8.14.5 Attribue Rating

After having rated the dissimilarities between all pairwise combinations of the program material, the participant were handed sheets on which his personal constructs were written (see Figure 8.24). The participant was instructed to rate each attribute for each amplifier according to his understanding of the word when generating it during the RGT task. The participant was informed of the purpose of the open ends, i.e. if having given ratings in the extremes of the scale, he was still able to rate a stimuli more extreme than his previous ratings.

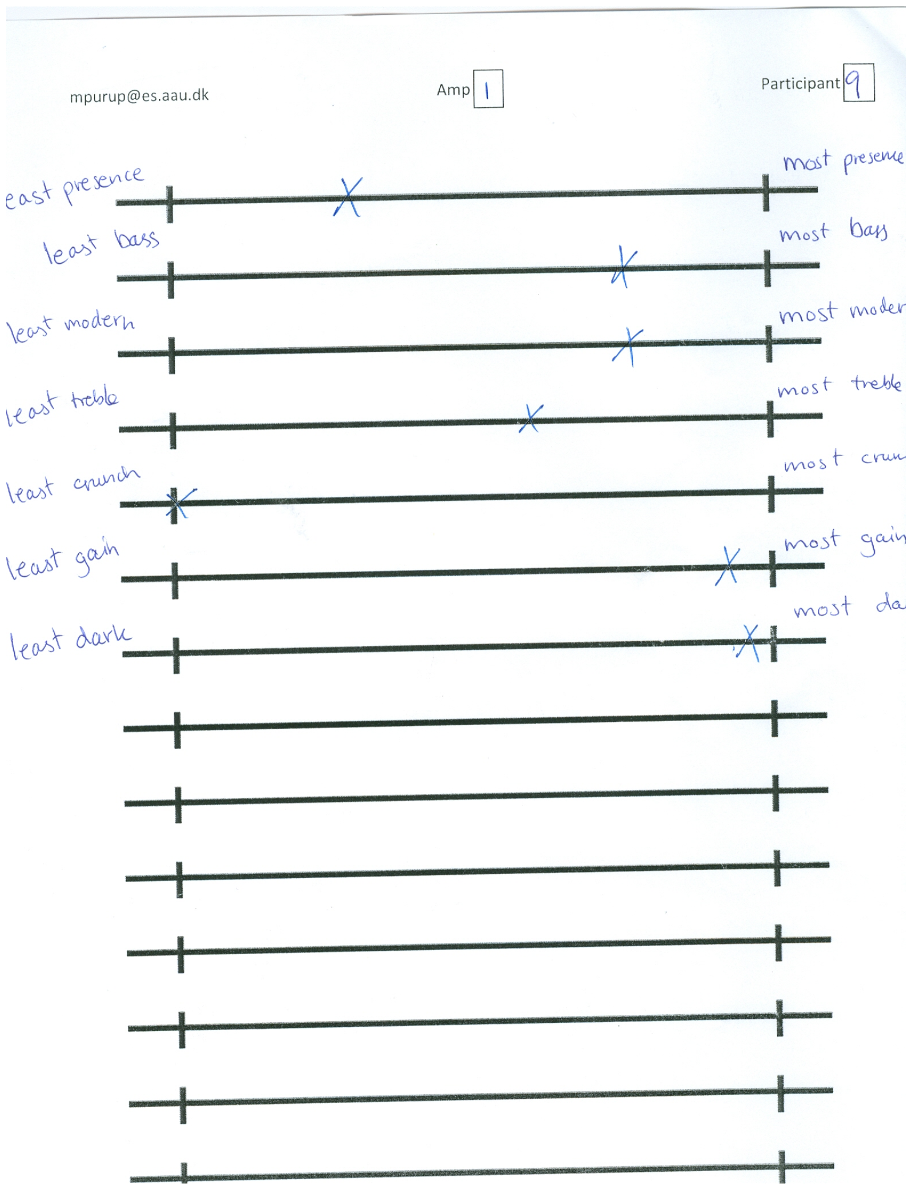

Figure 8.24: *Example of the VAS on which participant 9 rated his personal bipolar constructs.*

The value of the ratings were measured in centimeters rounded to one decimals place from the left anchor point. This method of measuring was chosen due to none of the participant going beyond the left anchor point and only one participant went beyond the right anchor point of the VAS.

8.15 Individual perceptions of high-gain guitar amplifiers

A biplot is a graphical representation of the information in an n x p data matrix [Everitt and Hothorn, 2010]. This biplot aims to represent both the observations and variables of the output of a principal component analysis. The biplot is solely generated of individual attribute scaling data. The length of the vector in the biplot determines the amount of variance explained by the vector. Metric MDS solutions were conducted for the individual dissimilarity data due to a nonmetric MDS solution eliminating too many subtle nuances and significant differences in the data. For each plot the amount of variance and measure of metric stress are given in the caption, respectively.

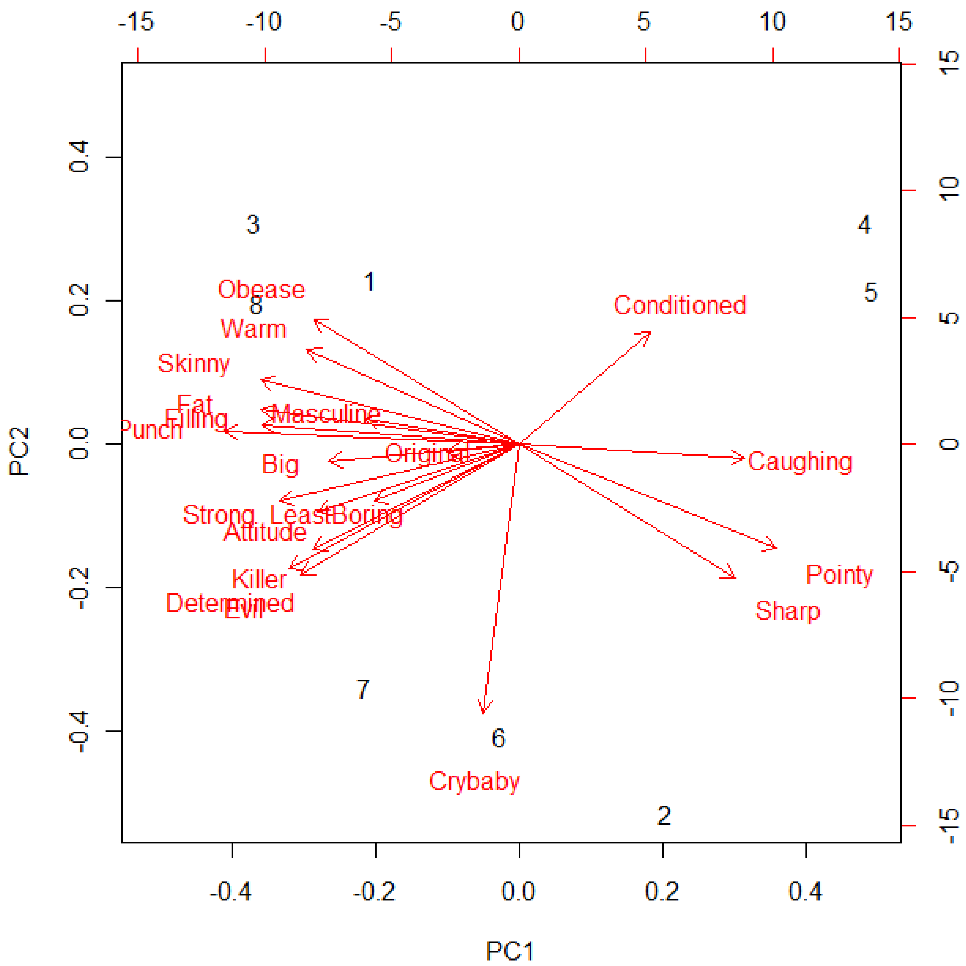

Figure 8.25: *Biplot of the first two principal components for participant no. 1. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 90.42 % of the variance observed in the data.*

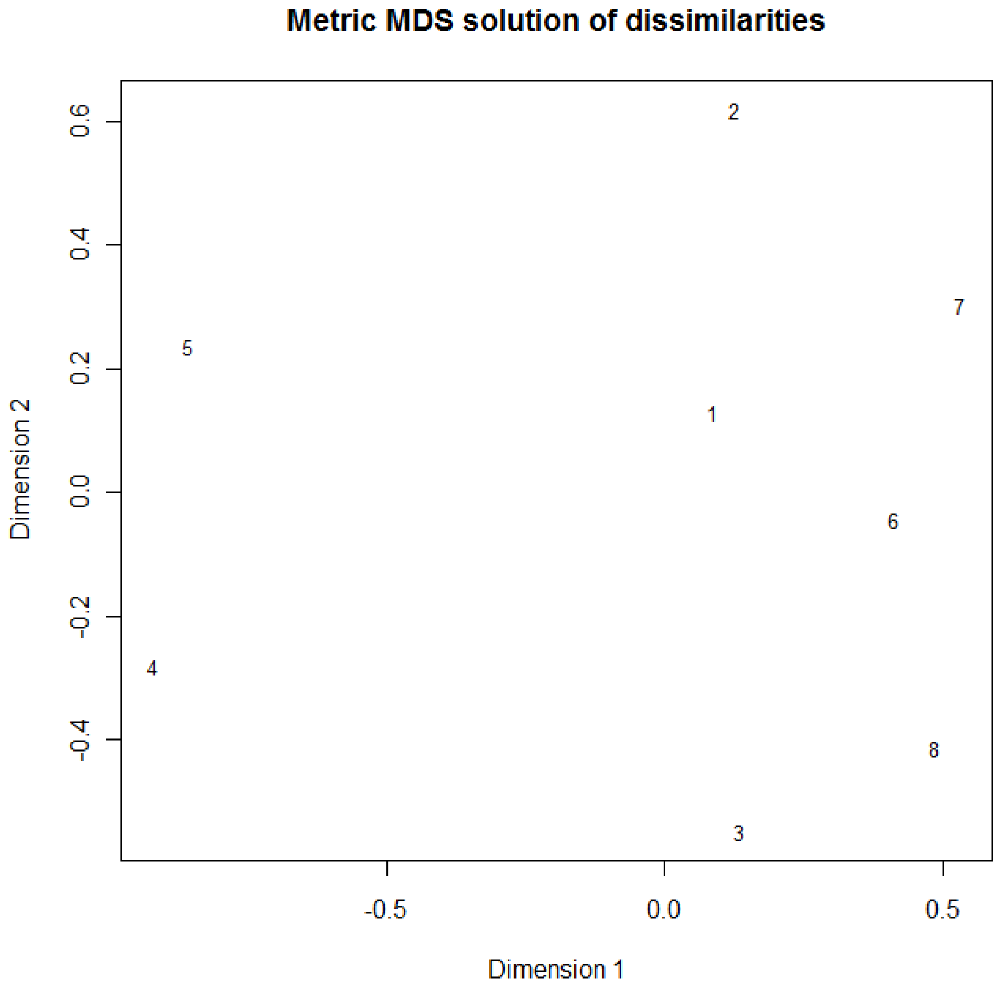

Figure 8.26: *Stimulus space in two dimensions for participant no. 1. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.028.*

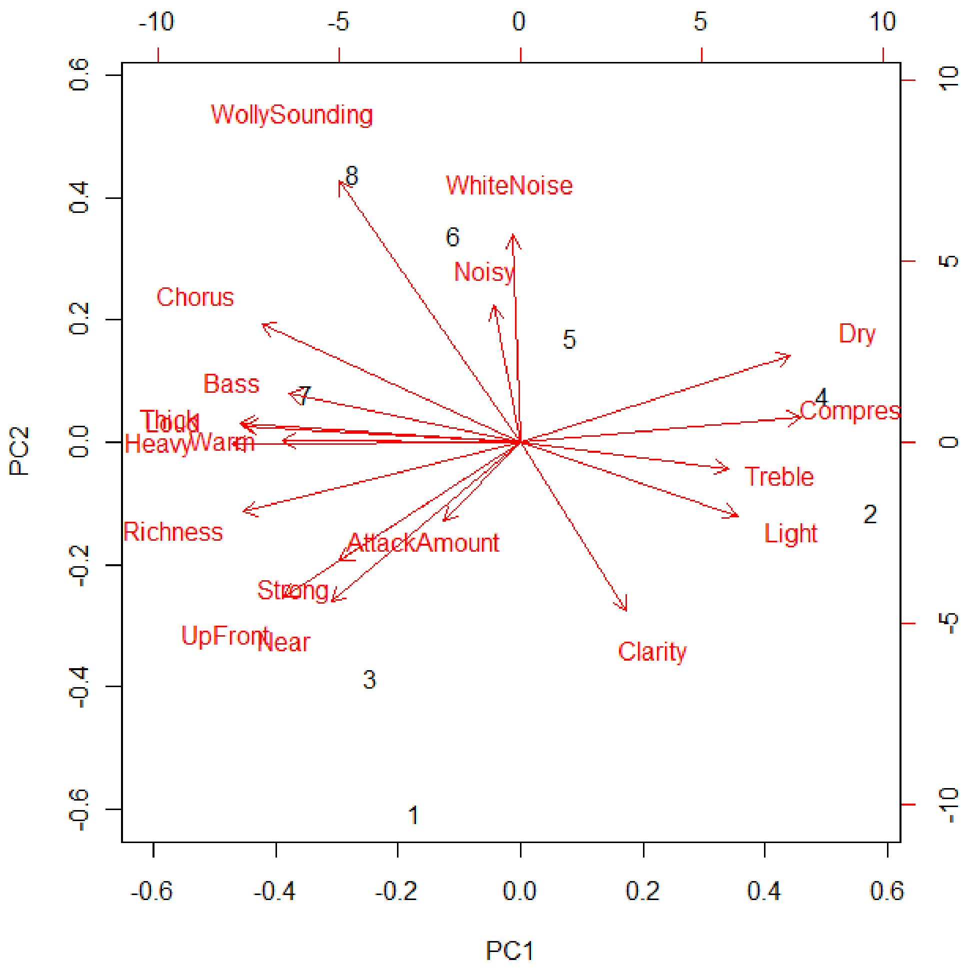

Figure 8.27: *Biplot of the first two principal components for participant no. 2. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 83.27 % of the variance observed in the data.*

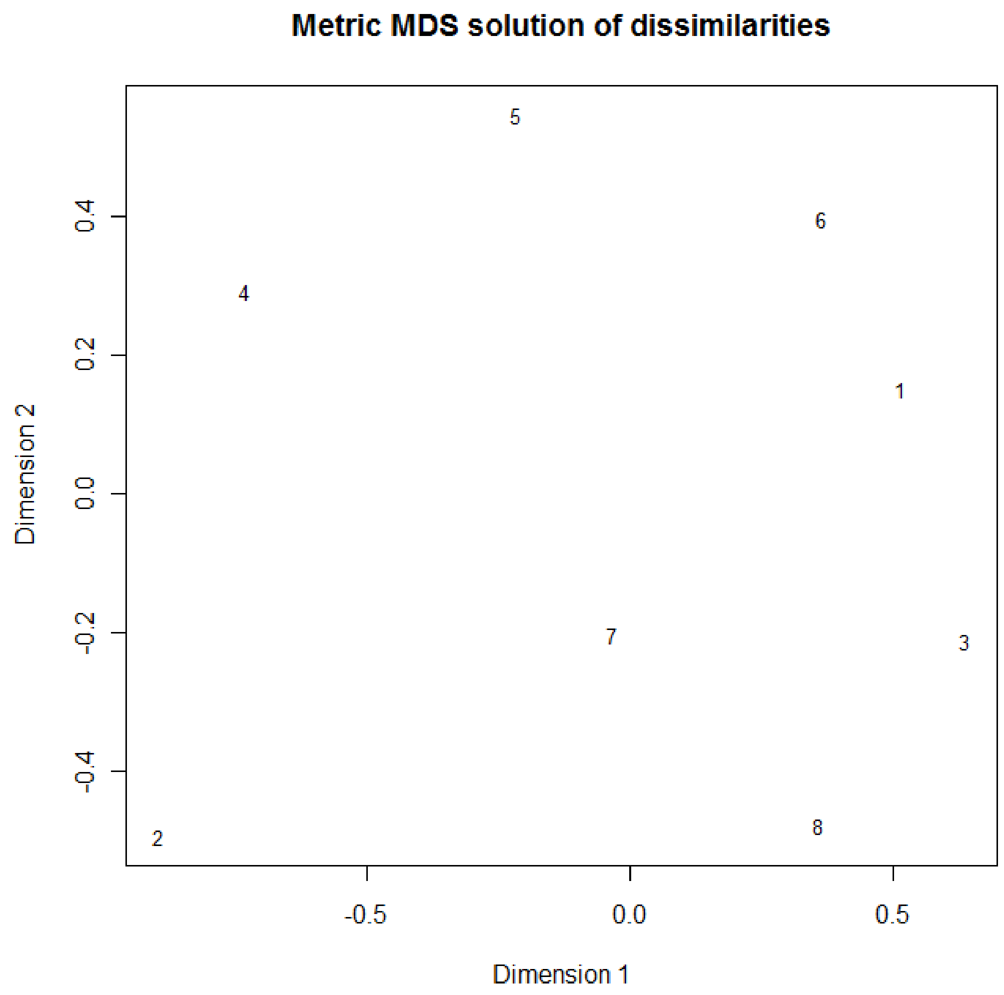

Figure 8.28: *Stimulus space in two dimensions for participant no. 2. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.017.*

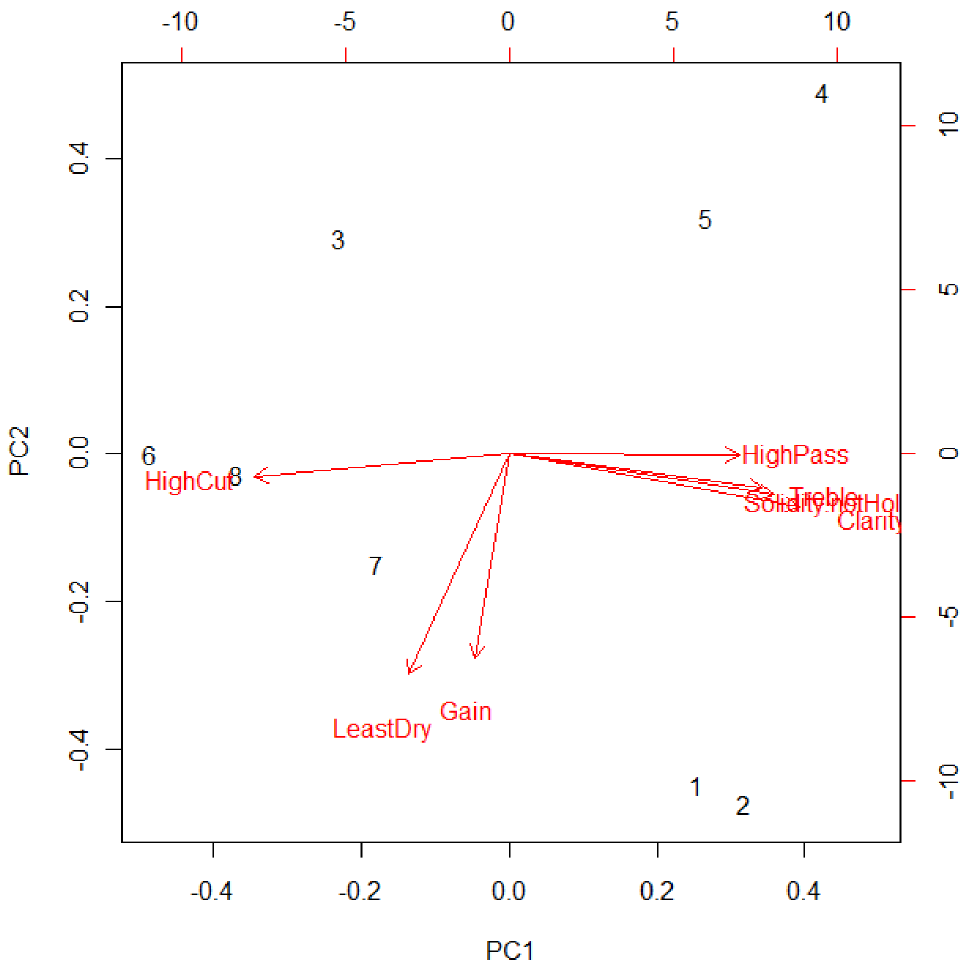

Figure 8.29: *Biplot of the first two principal components for participant no. 3. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 93.57 % of the variance observed in the data.*

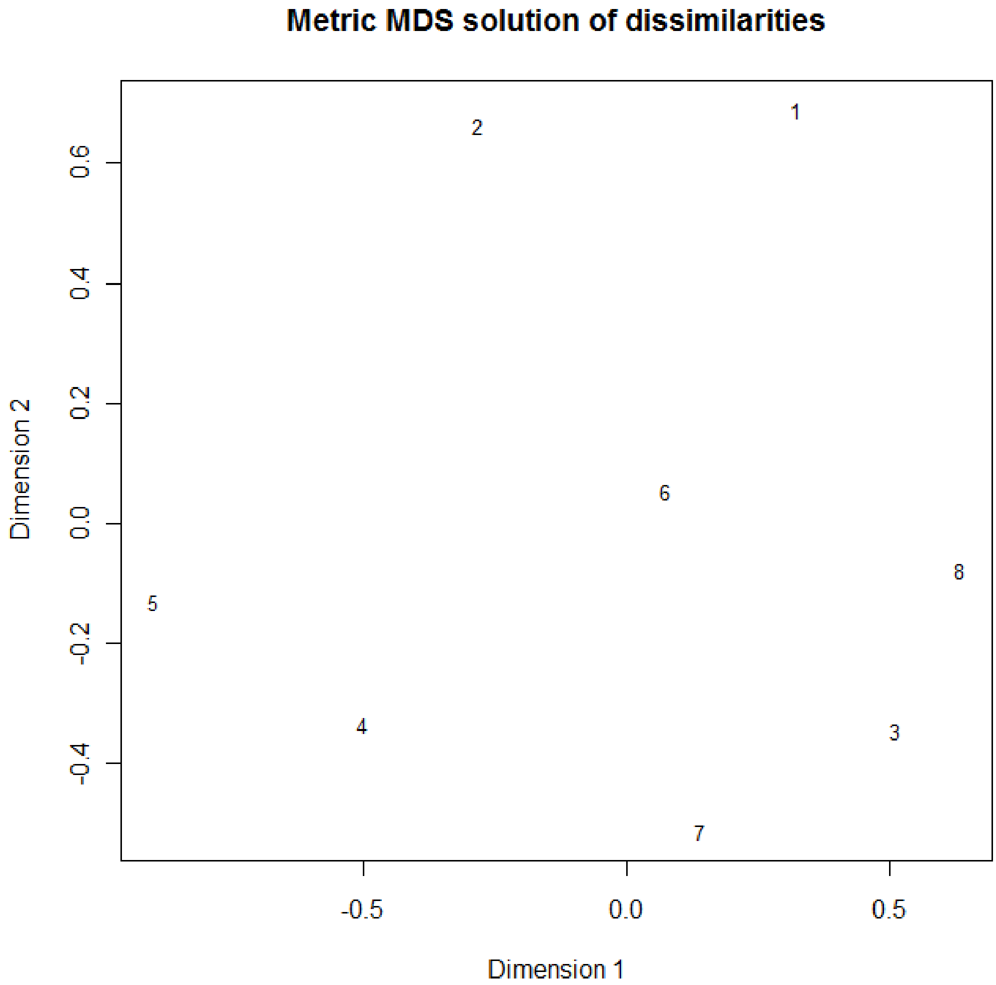

Figure 8.30: *Stimulus space in two dimensions for participant no. 3. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.040.*

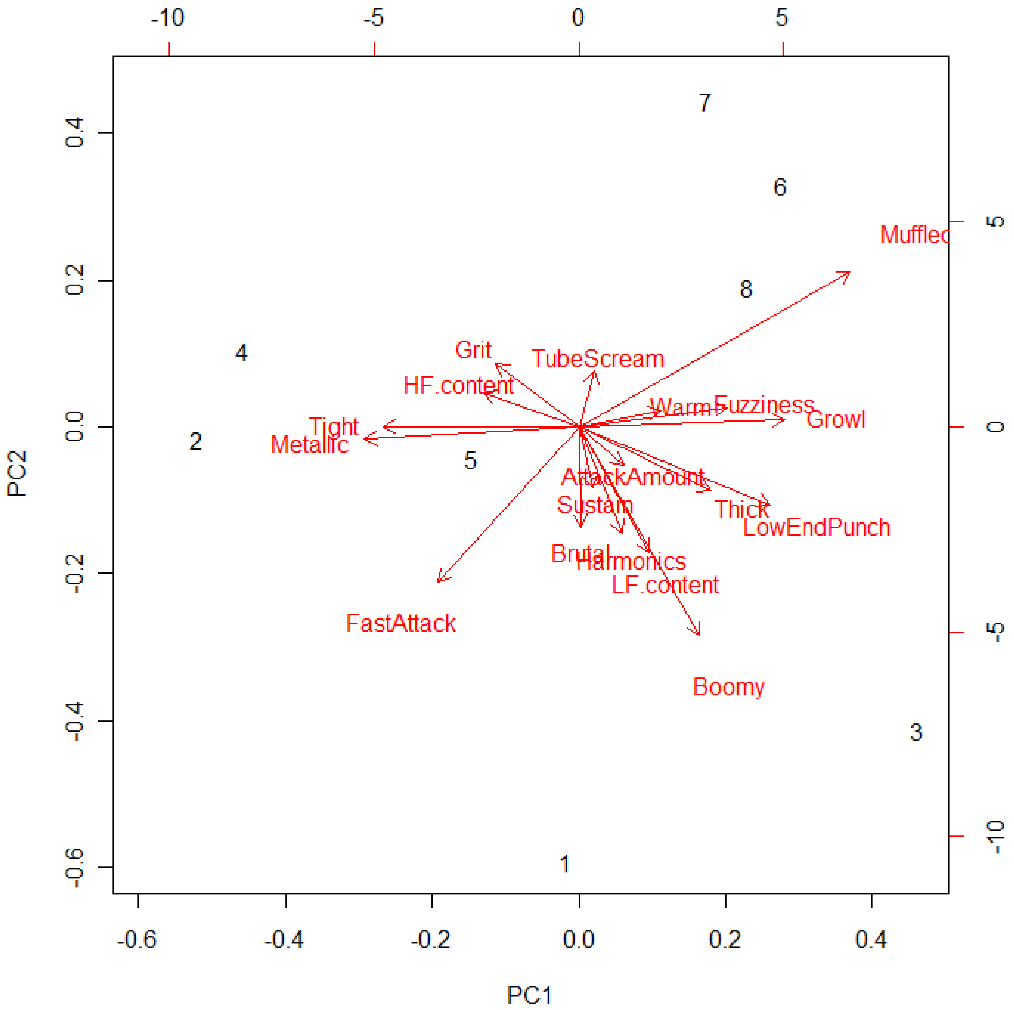

Figure 8.31: *Biplot of the first two principal components for participant no. 4. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 68.49 % of the variance observed in the data.*

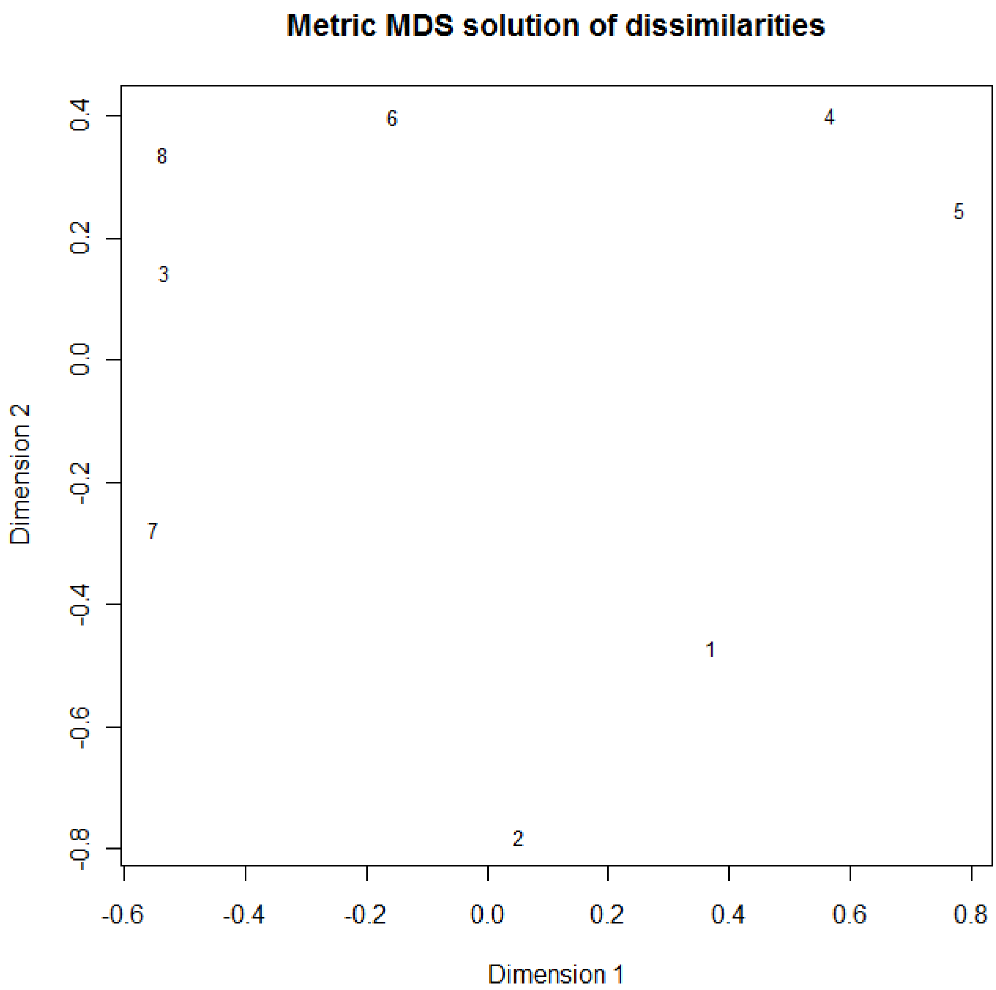

Figure 8.32: *Stimulus space in two dimensions for participant no. 4. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.030.*

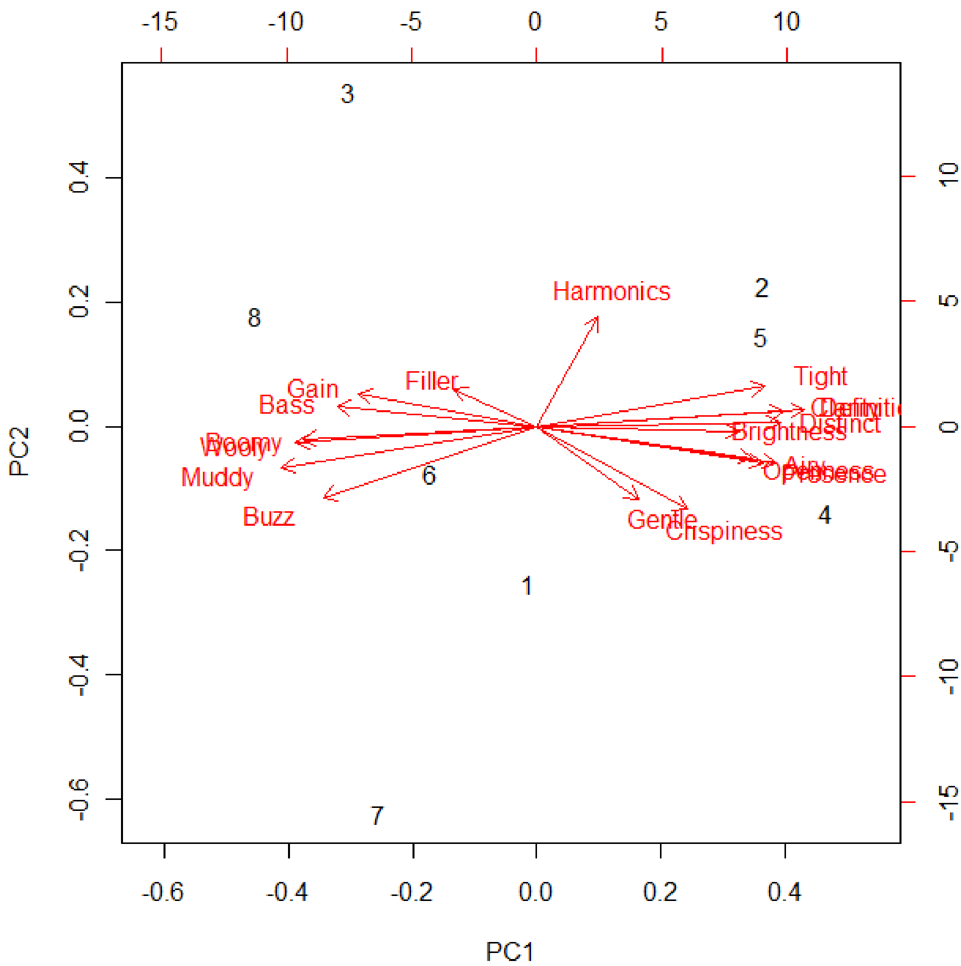

Figure 8.33: *Biplot of the first two principal components for participant no. 5. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 92.15 % of the variance observed in the data.*

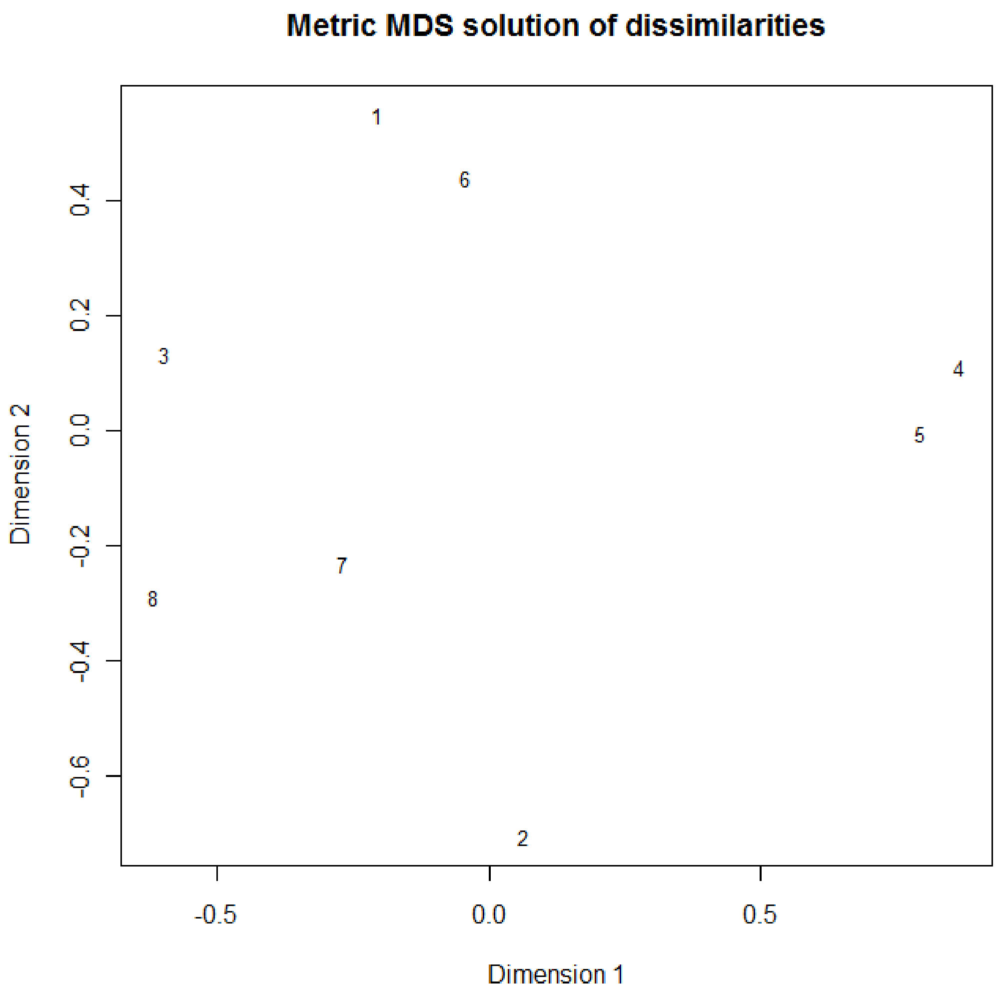

Figure 8.34: *Stimulus space in two dimensions for participant no. 5. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.029.*

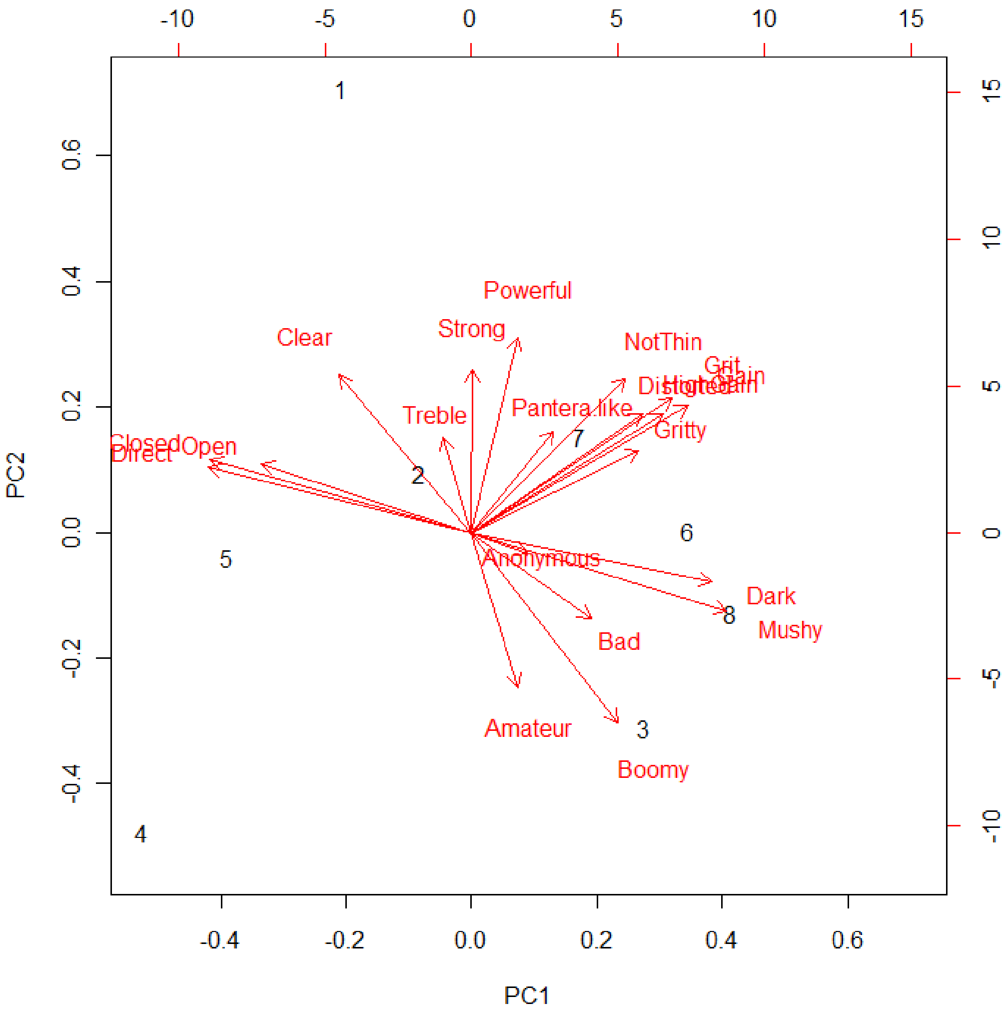

Figure 8.35: *Biplot of the first two principal components for participant no. 6. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 80.32 % of the variance observed in the data.*

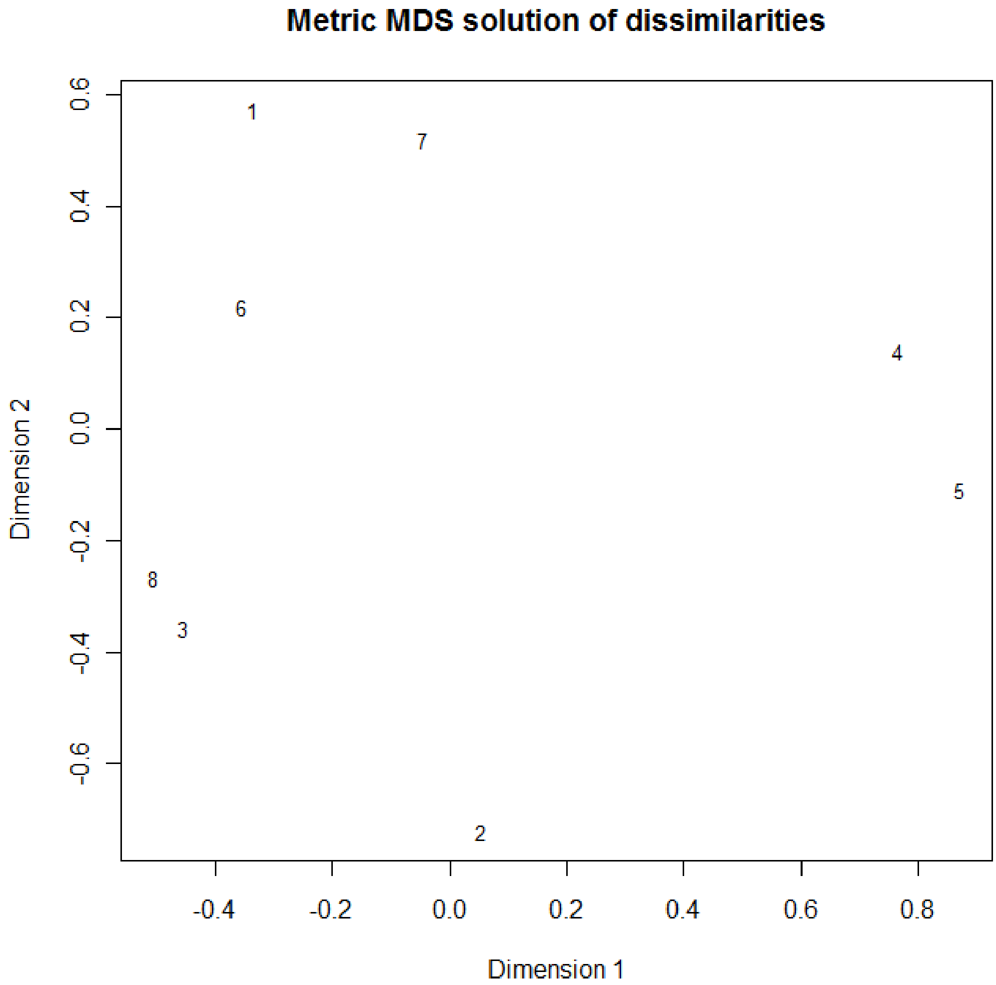

Figure 8.36: *Stimulus space in two dimensions for participant no. 6. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.016.*

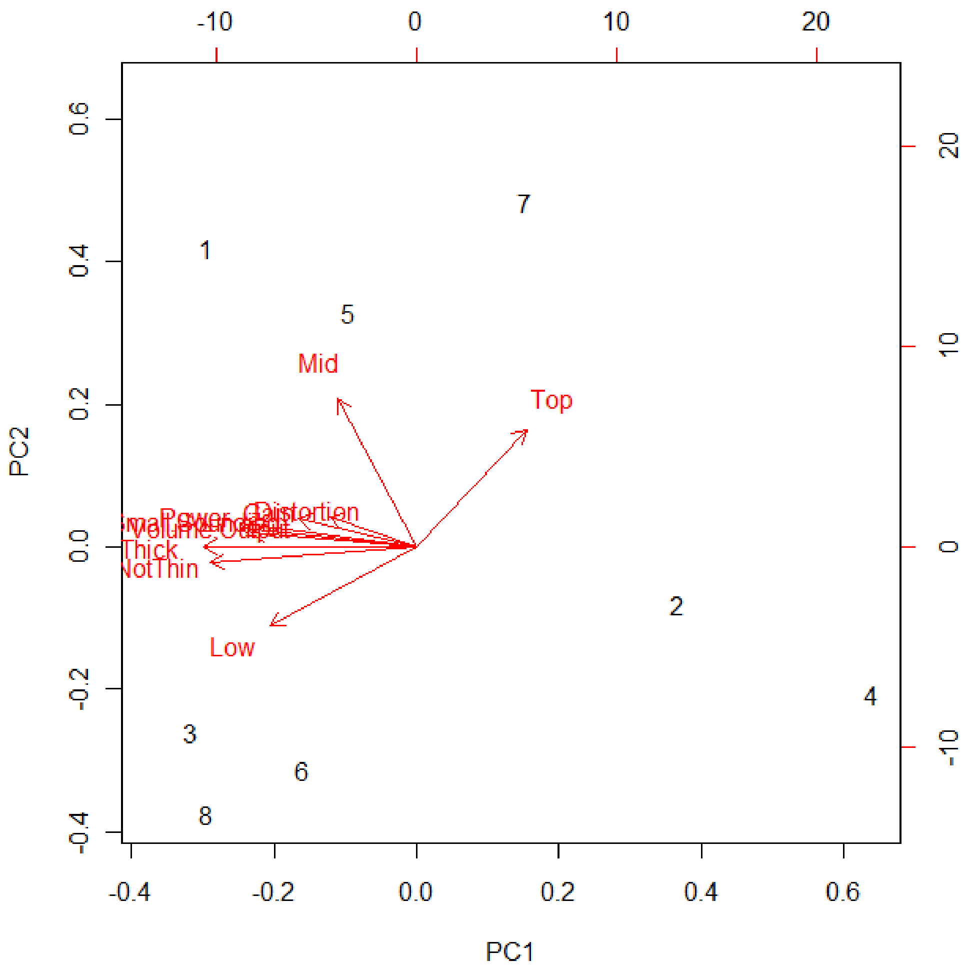

Figure 8.37: *Biplot of the first two principal components for participant no. 7. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 95.27 % of the variance observed in the data.*

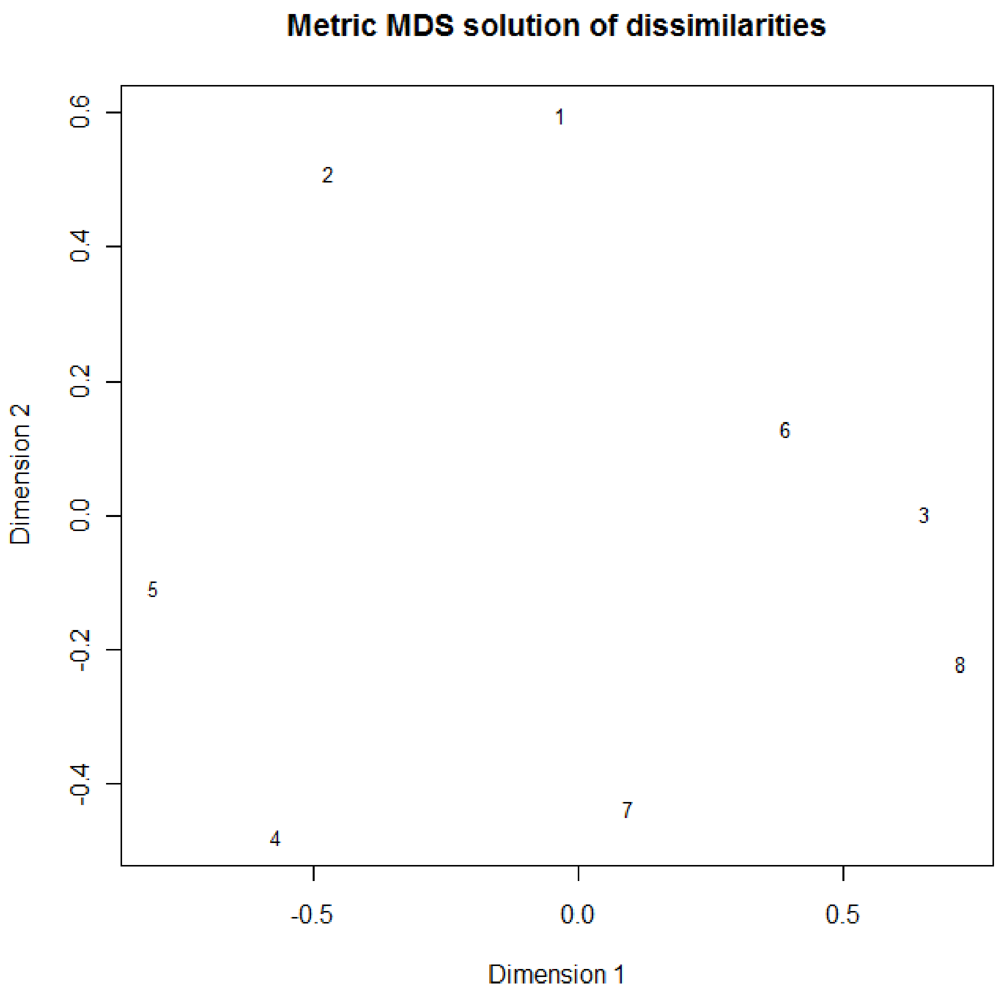

Figure 8.38: *Stimulus space in two dimensions for participant no. 7. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.025.*

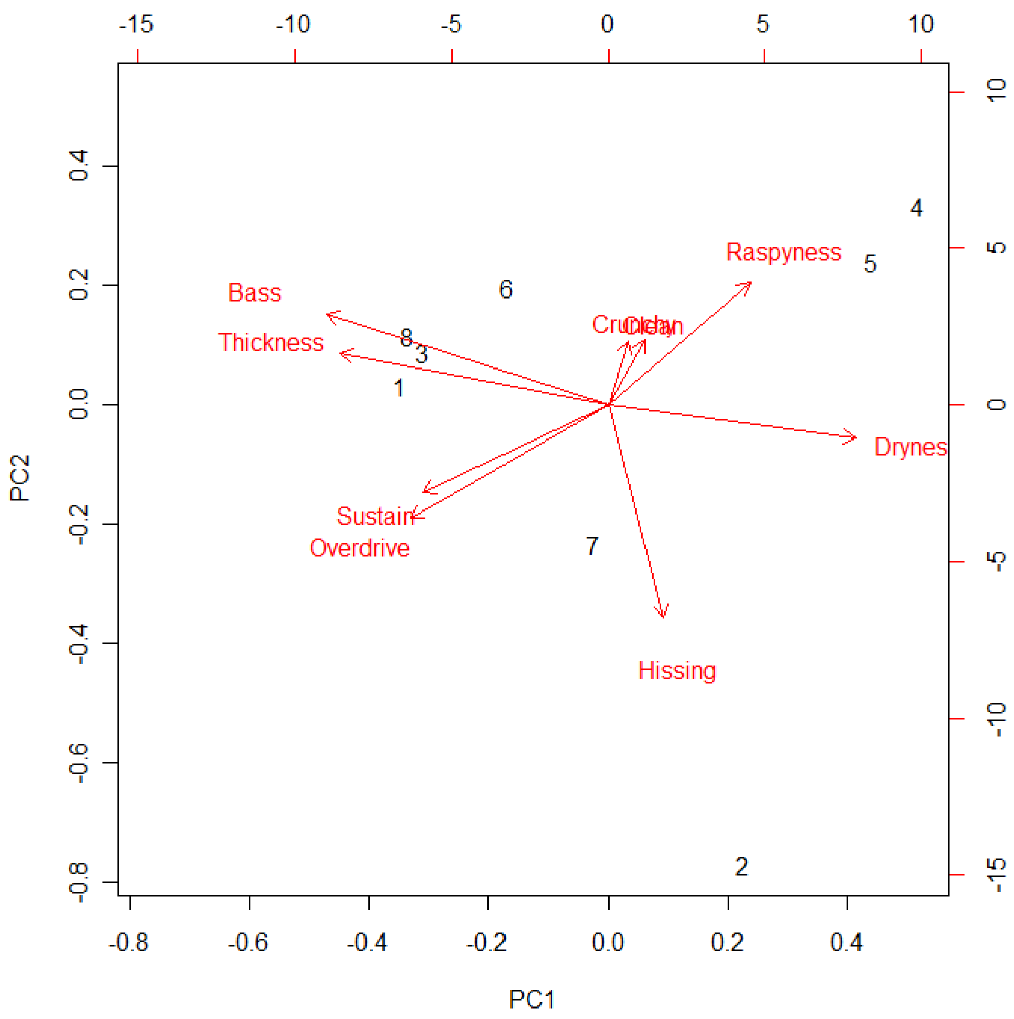

Figure 8.39: *Biplot of the first two principal components for participant no. 8. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 82.22 % of the variance observed in the data.*

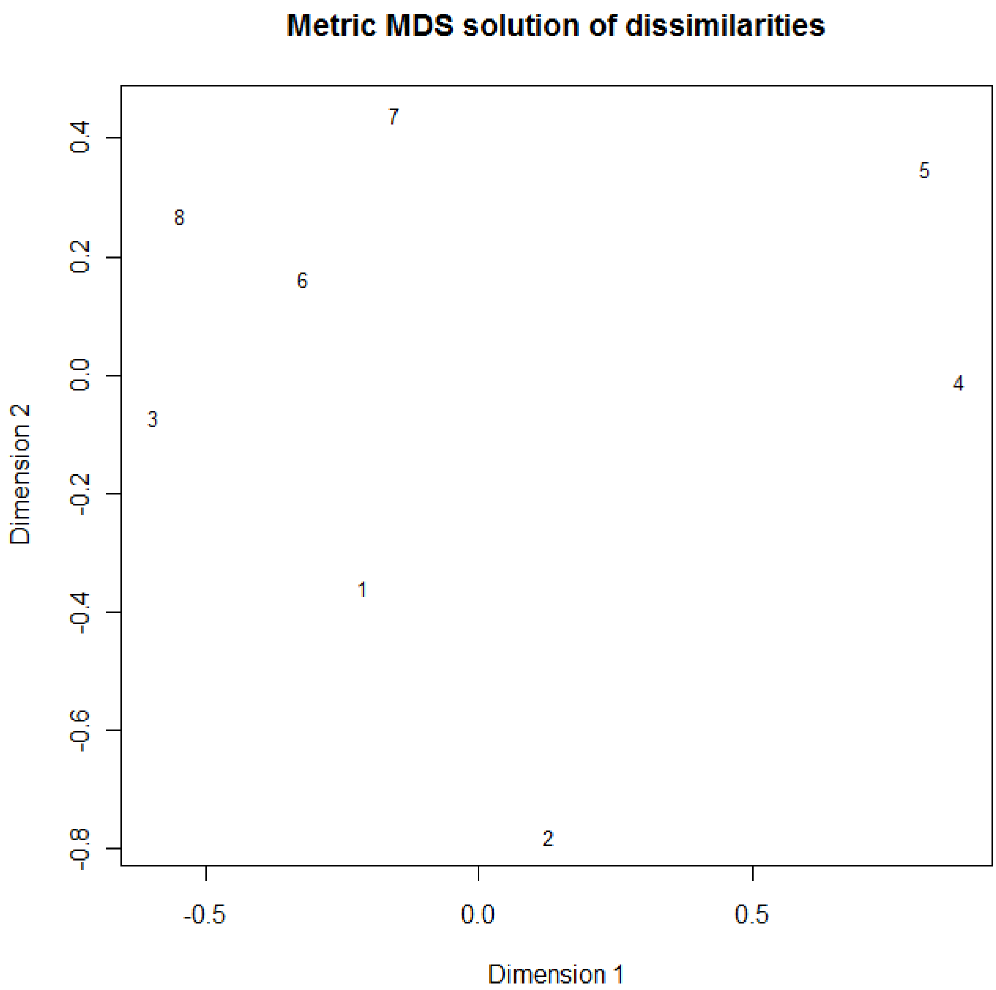

Figure 8.40: *Stimulus space in two dimensions for participant no. 8. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.015.*

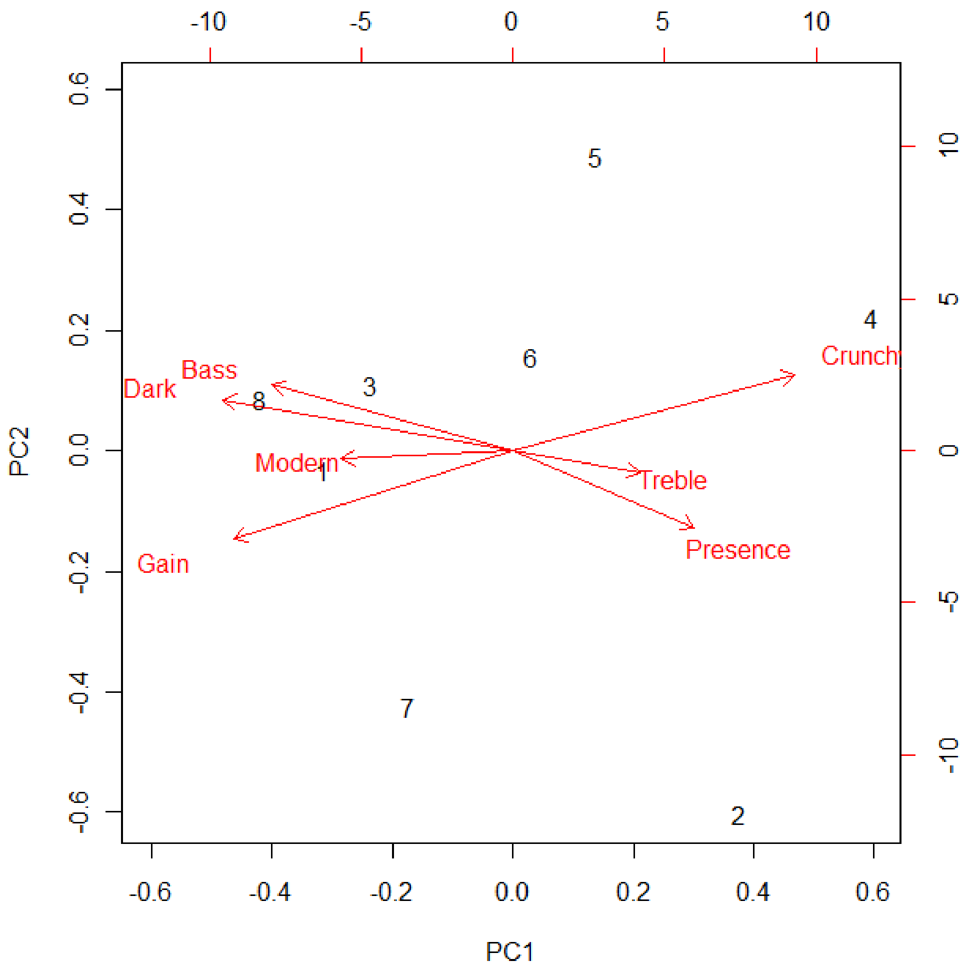

Figure 8.41: *Biplot of the first two principal components for participant no. 9. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 96.53 % of the variance observed in the data.*

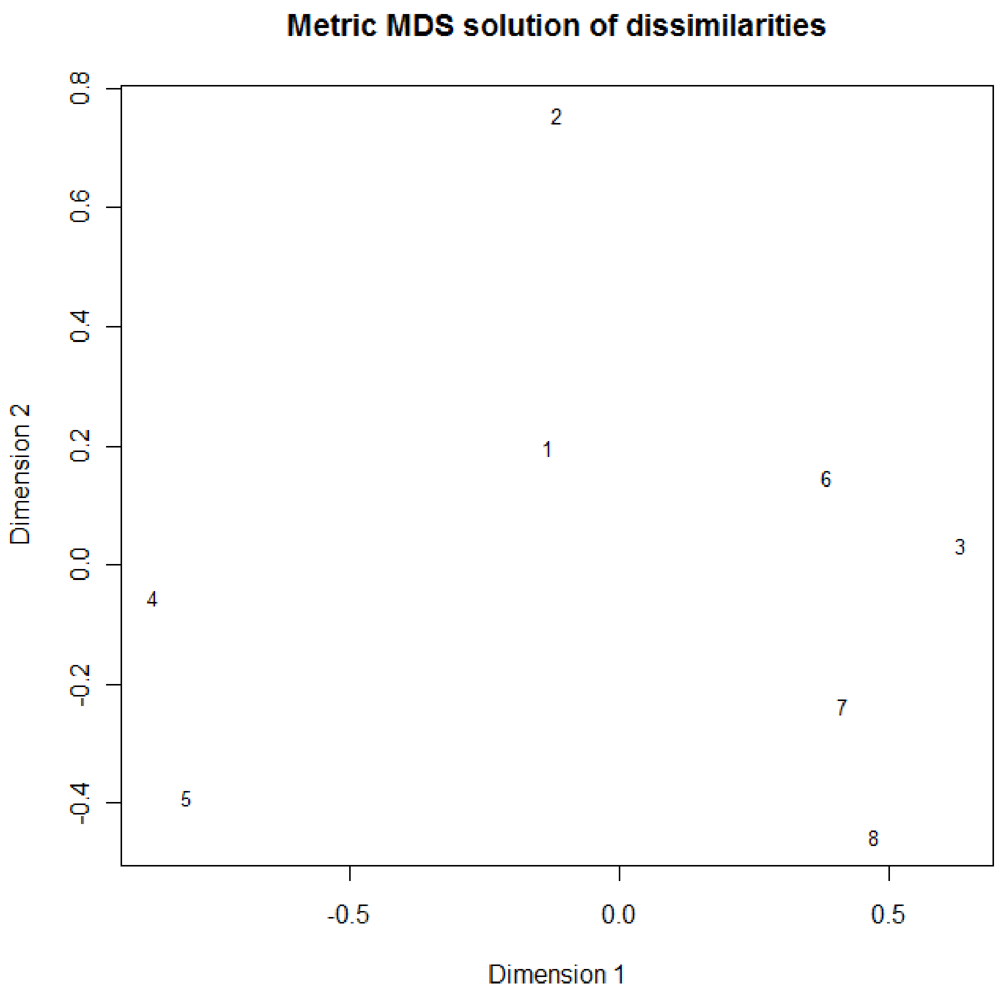

Figure 8.42: *Stimulus space in two dimensions for participant no. 9. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.023.*

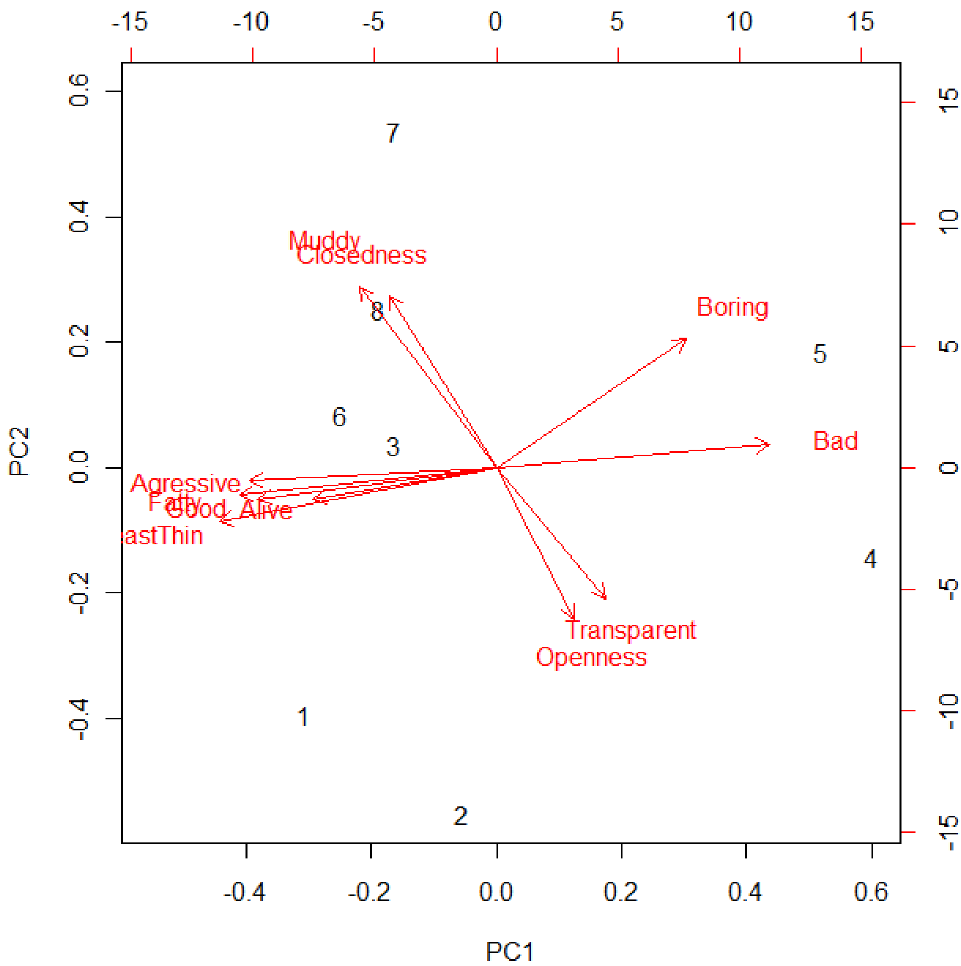

Figure 8.43: *Biplot of the first two principal components for participant no. 10. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 89.56 % of the variance observed in the data.*

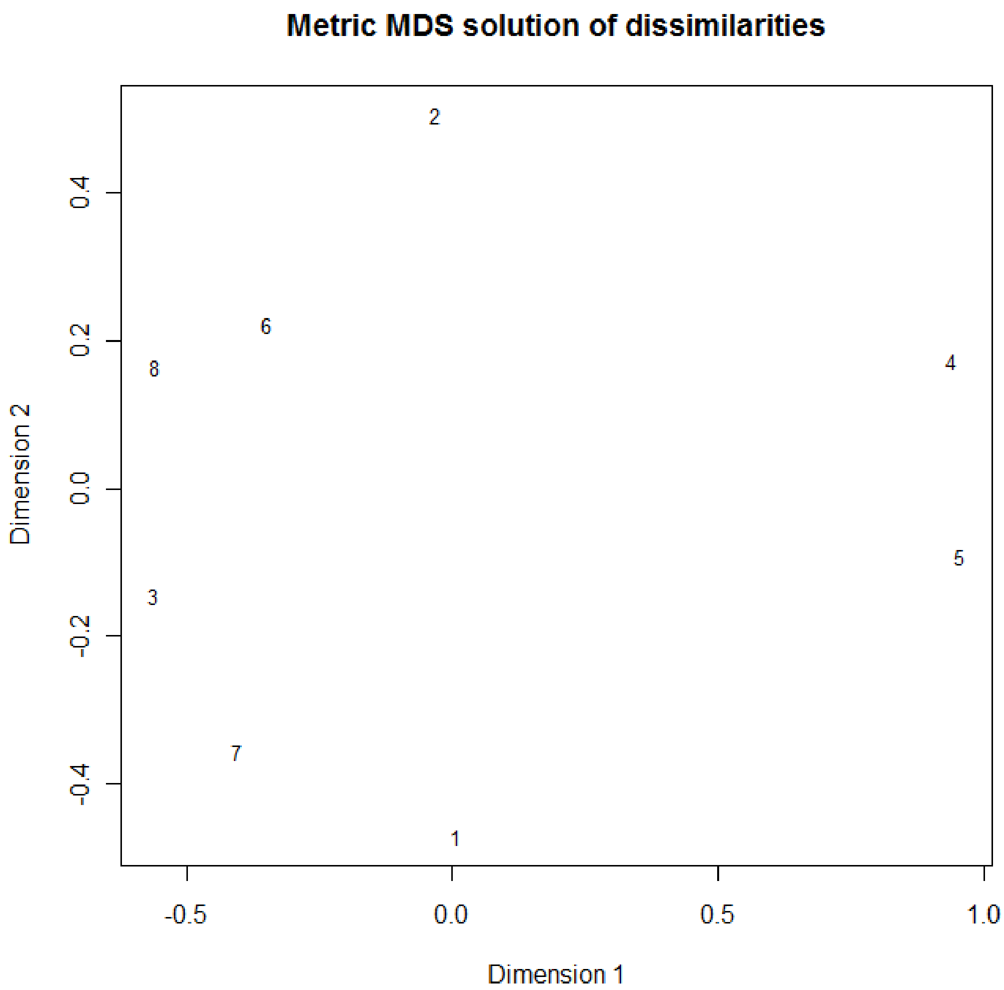

Figure 8.44: *Stimulus space in two dimensions for participant no. 10. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.012.*

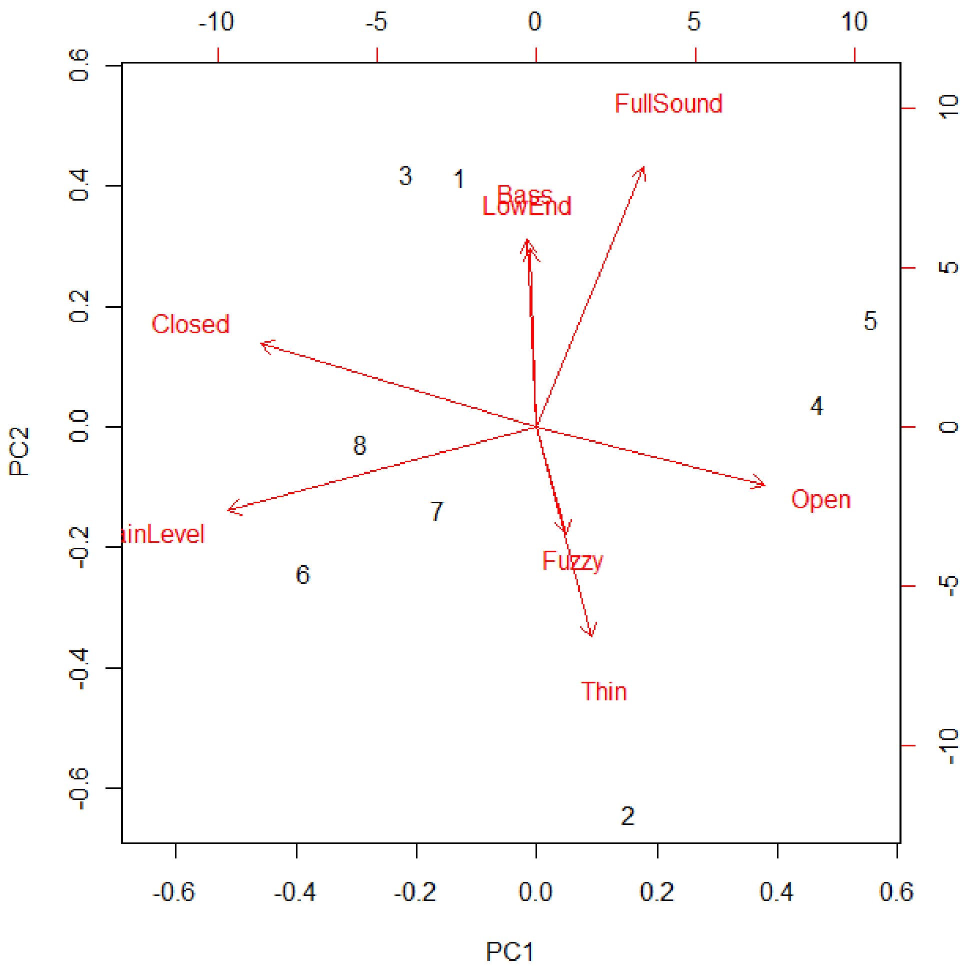

Figure 8.45: *Biplot of the first two principal components for participant no. 11. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 80.69 % of the variance observed in the data.*

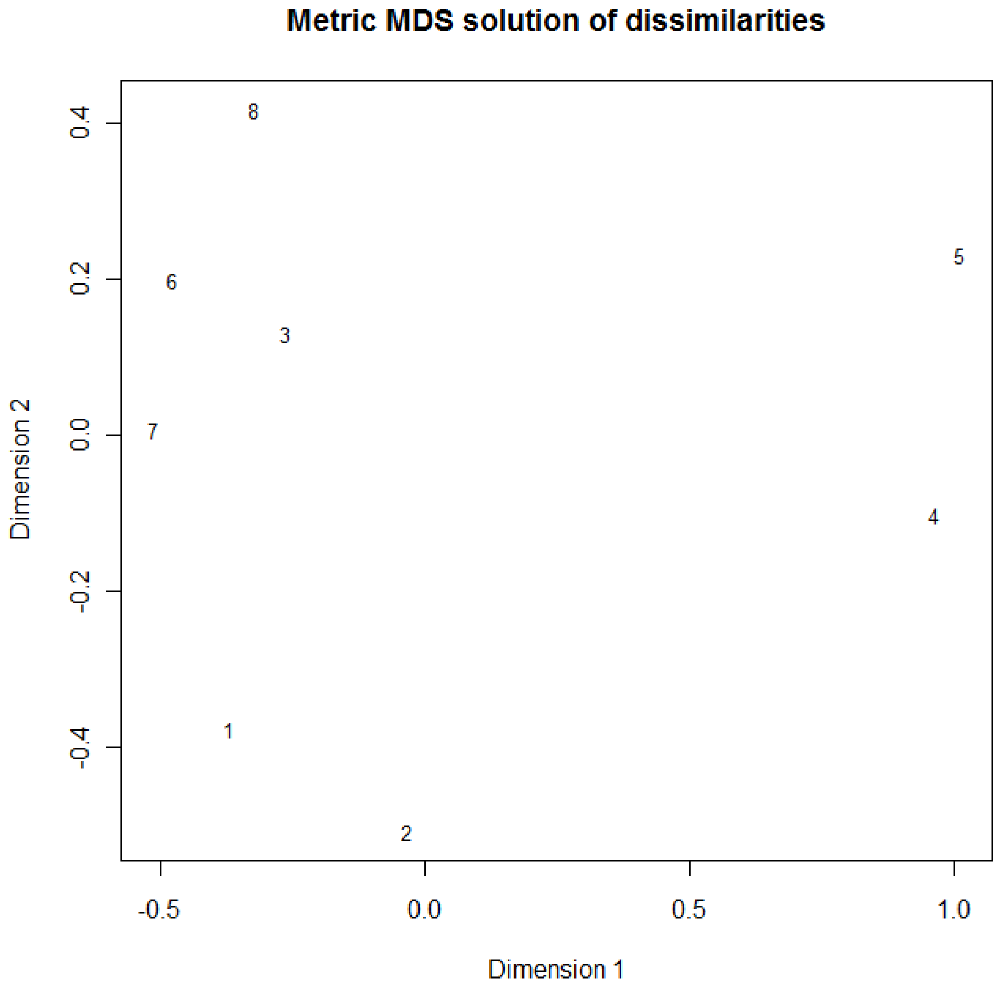

Figure 8.46: *Stimulus space in two dimensions for participant no. 11. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.015.*

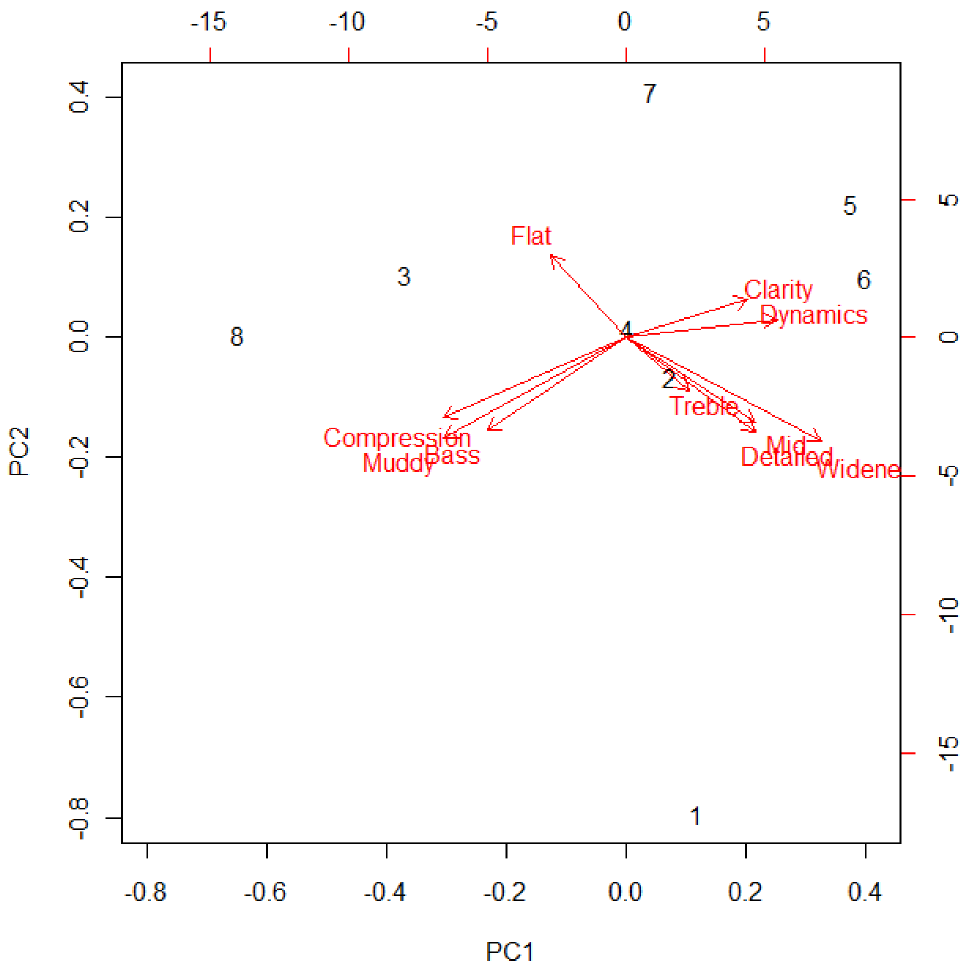

Figure 8.47: *Biplot of the first two principal components for participant no. 12. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 80.20 % of the variance observed in the data.*

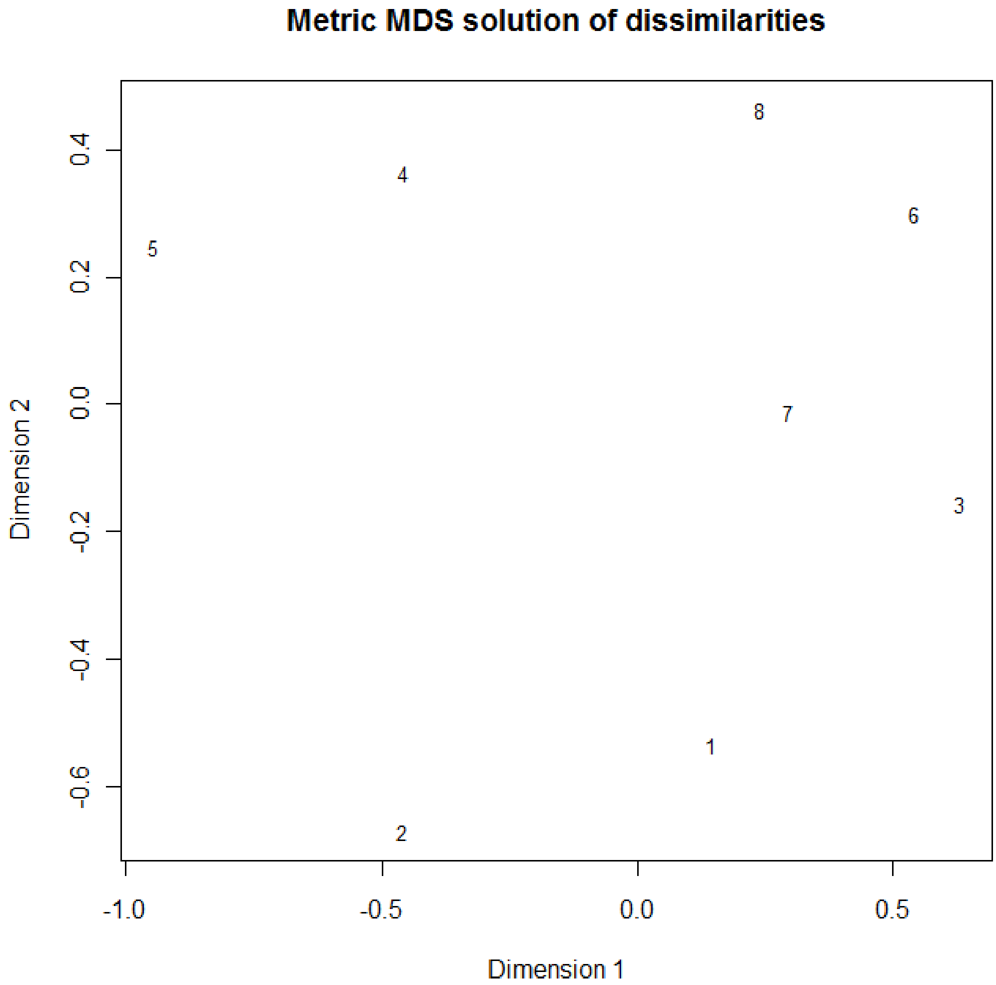

Figure 8.48: *Stimulus space in two dimensions for participant no. 12. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.019.*

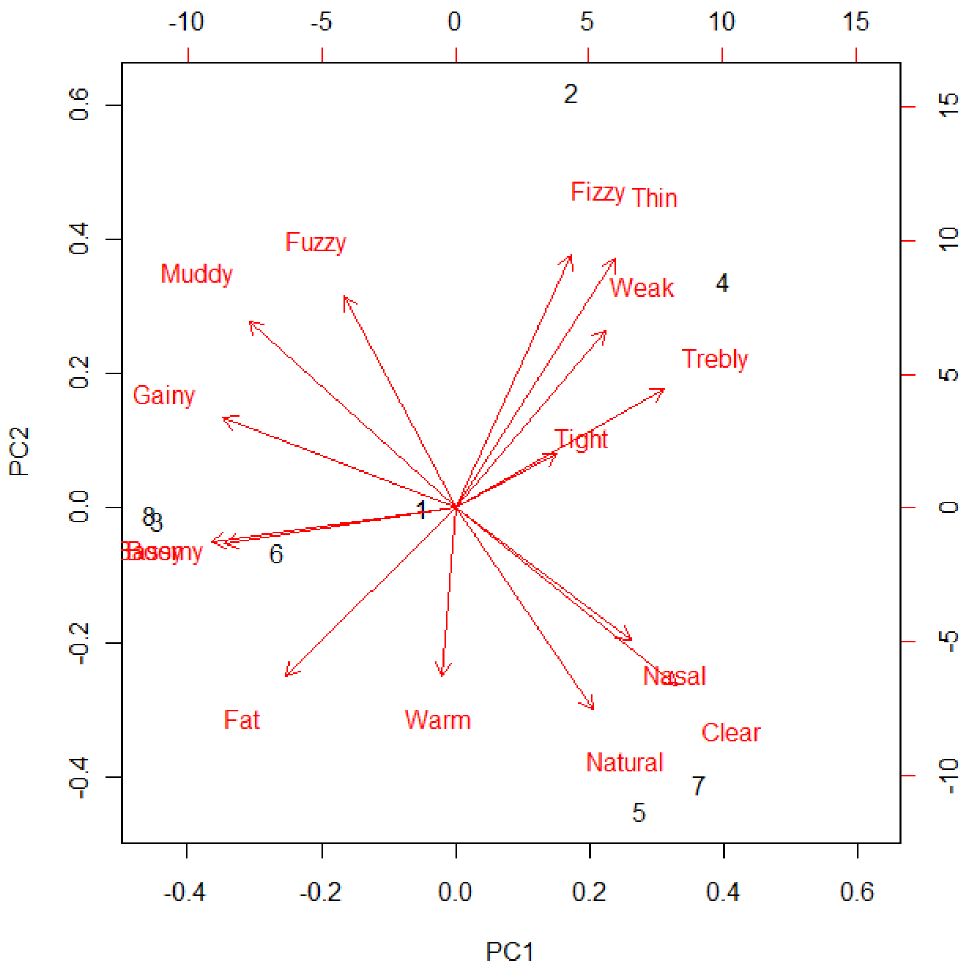

Figure 8.49: *Biplot of the first two principal components for participant no. 13. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 86.18 % of the variance observed in the data.*

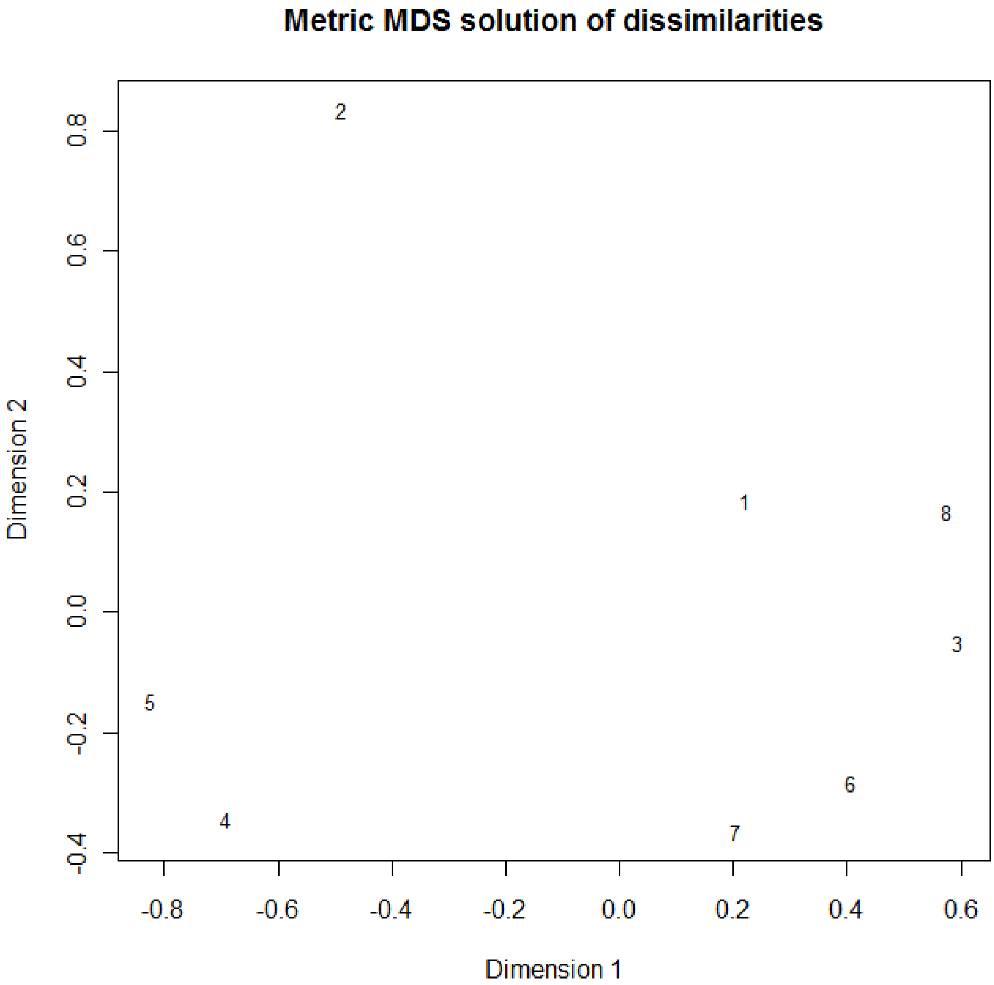

Figure 8.50: *Stimulus space in two dimensions for participant no. 13. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.013.*

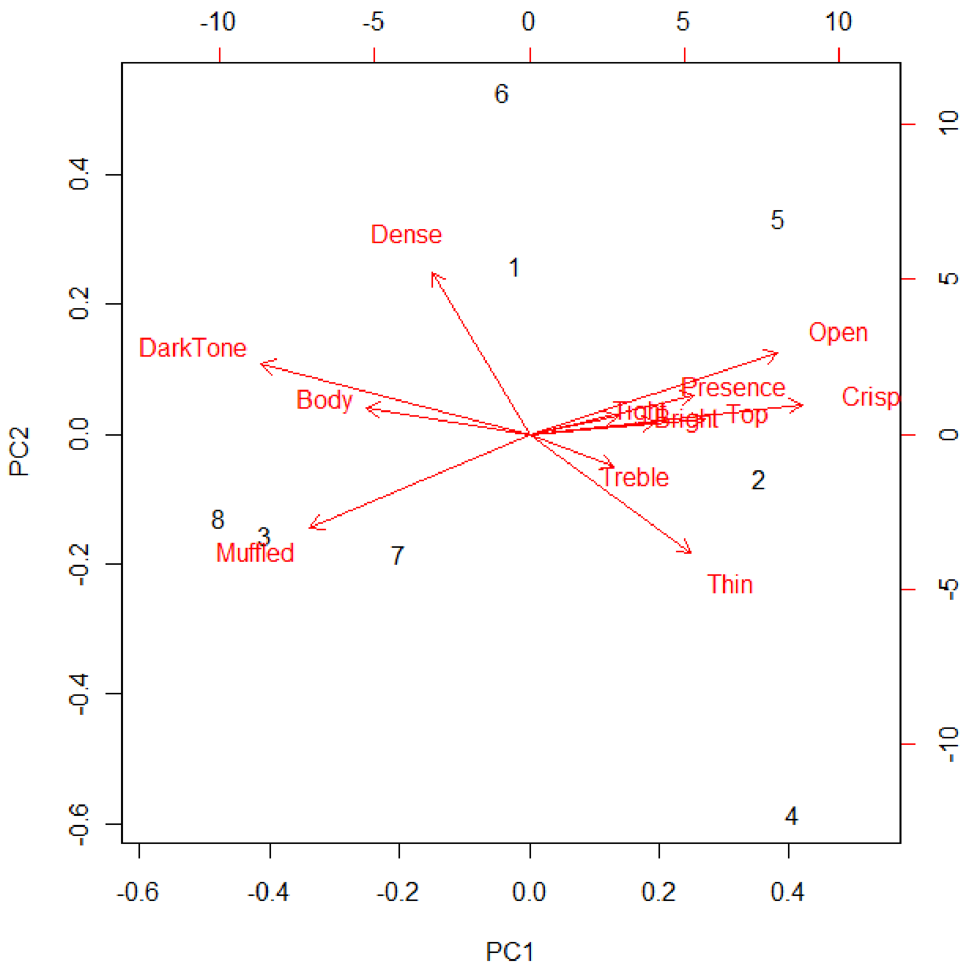

Figure 8.51: *Biplot of the first two principal components for participant no. 14. The biplot is solely generated of individual attribute scaling data. The attribute rating data for "muddiness" was removed due to a a missing response for amplifier 4. The first two principal components can explain 83.35 % of the variance observed in the data.*

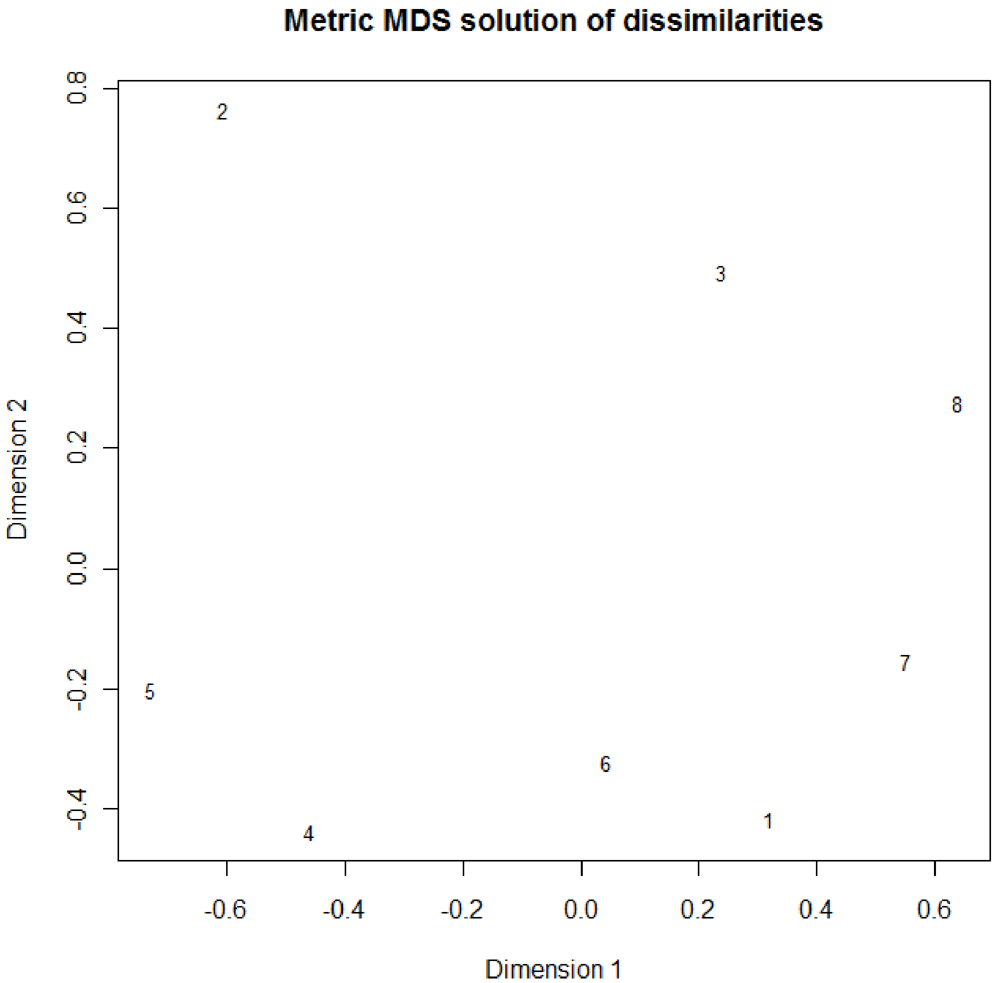

Figure 8.52: *Stimulus space in two dimensions for participant no. 14. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.021.*

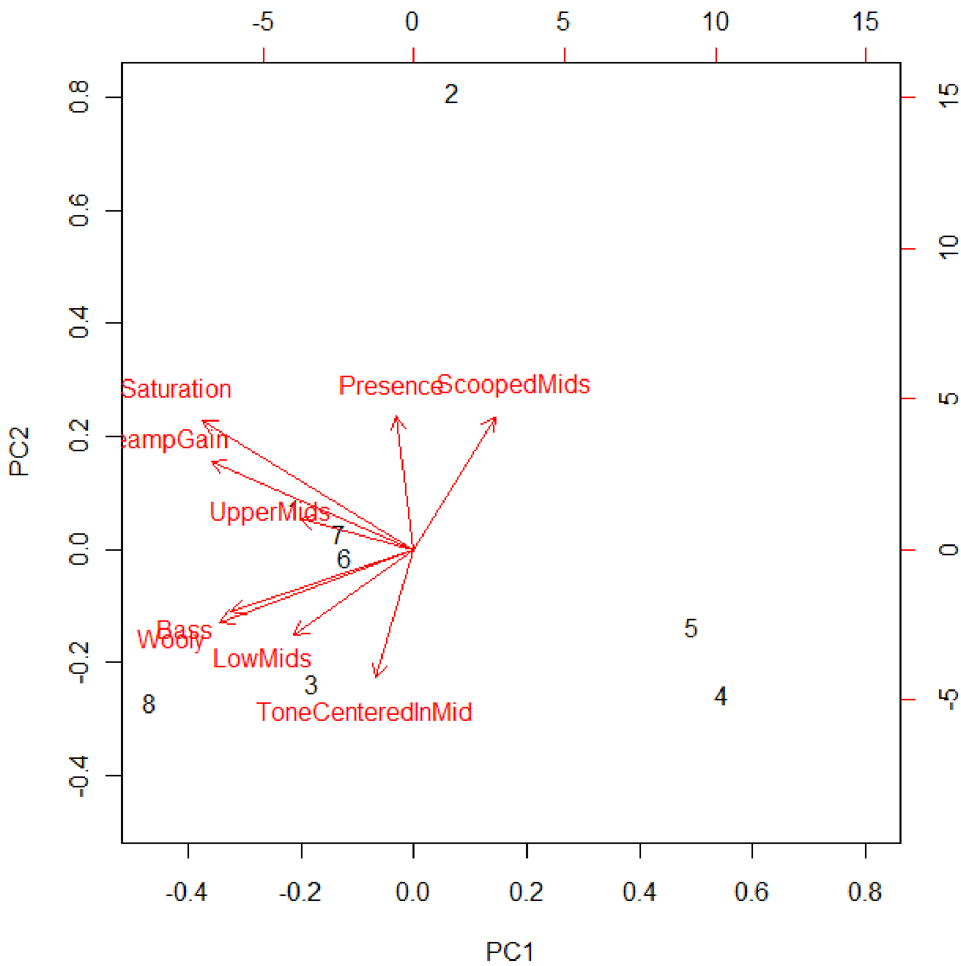

Figure 8.53: *Biplot of the first two principal components for participant no. 15. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 86.77 % of the variance observed in the data.*

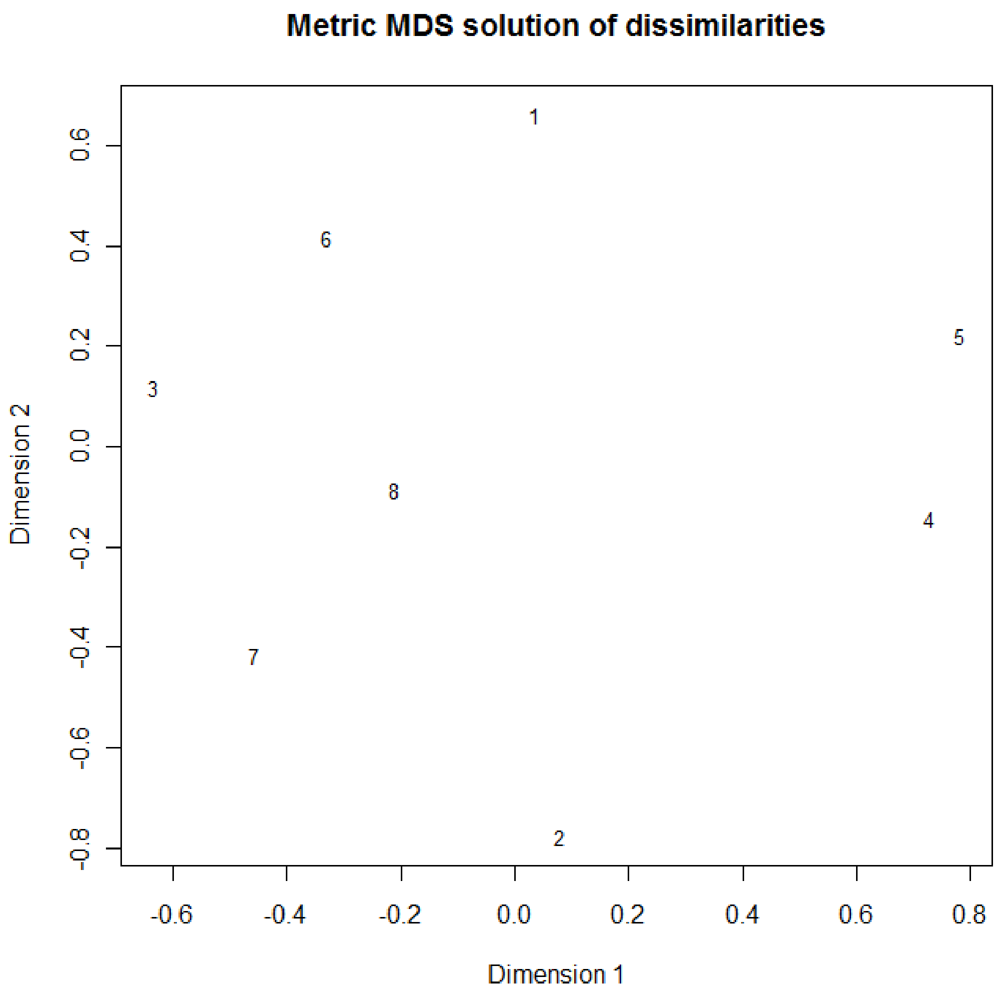

Figure 8.54: *Stimulus space in two dimensions for participant no. 15. This multidimensional scaling plot is solely generated of individual dissimilarity data. Amplifier 1 is positioned behind the M in "UpperMids". The metric stress measure of the solution is 0.033.*

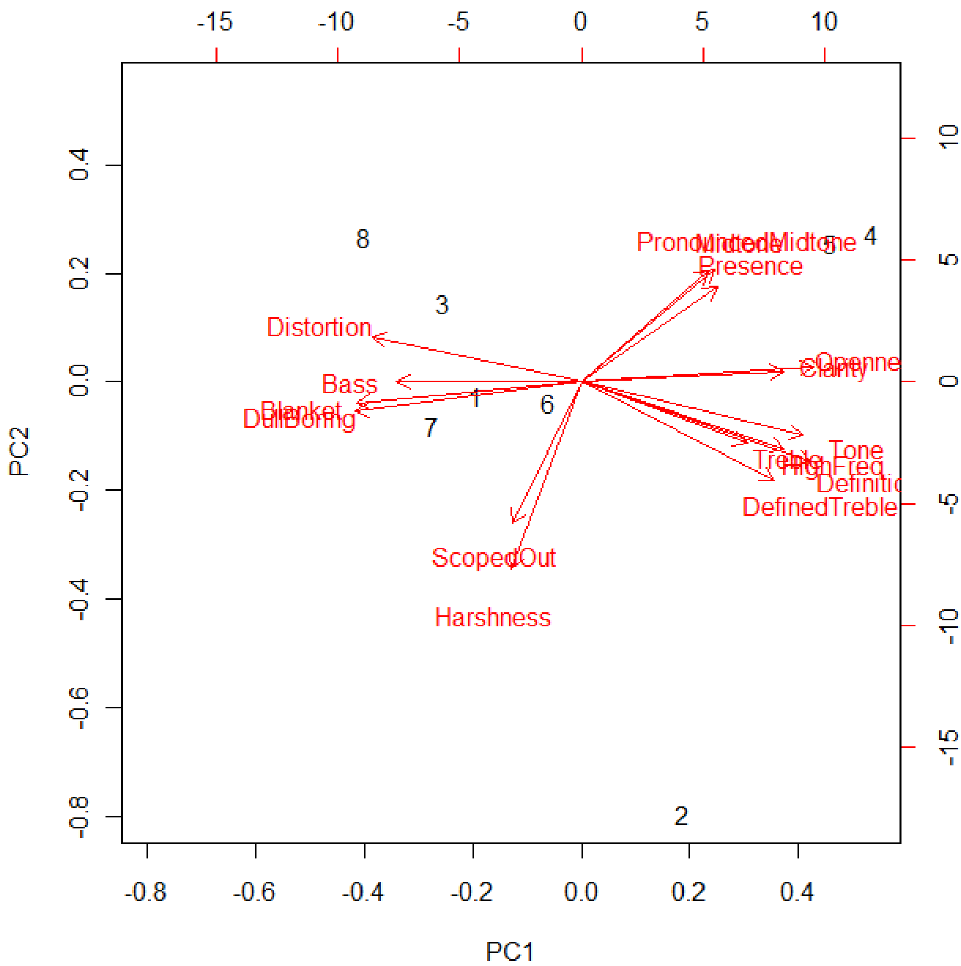

Figure 8.55: *Biplot of the first two principal components for participant no. 16. The biplot is solely generated of individual attribute scaling data. The first two principal components can explain 91.31 % of the variance observed in the data.*

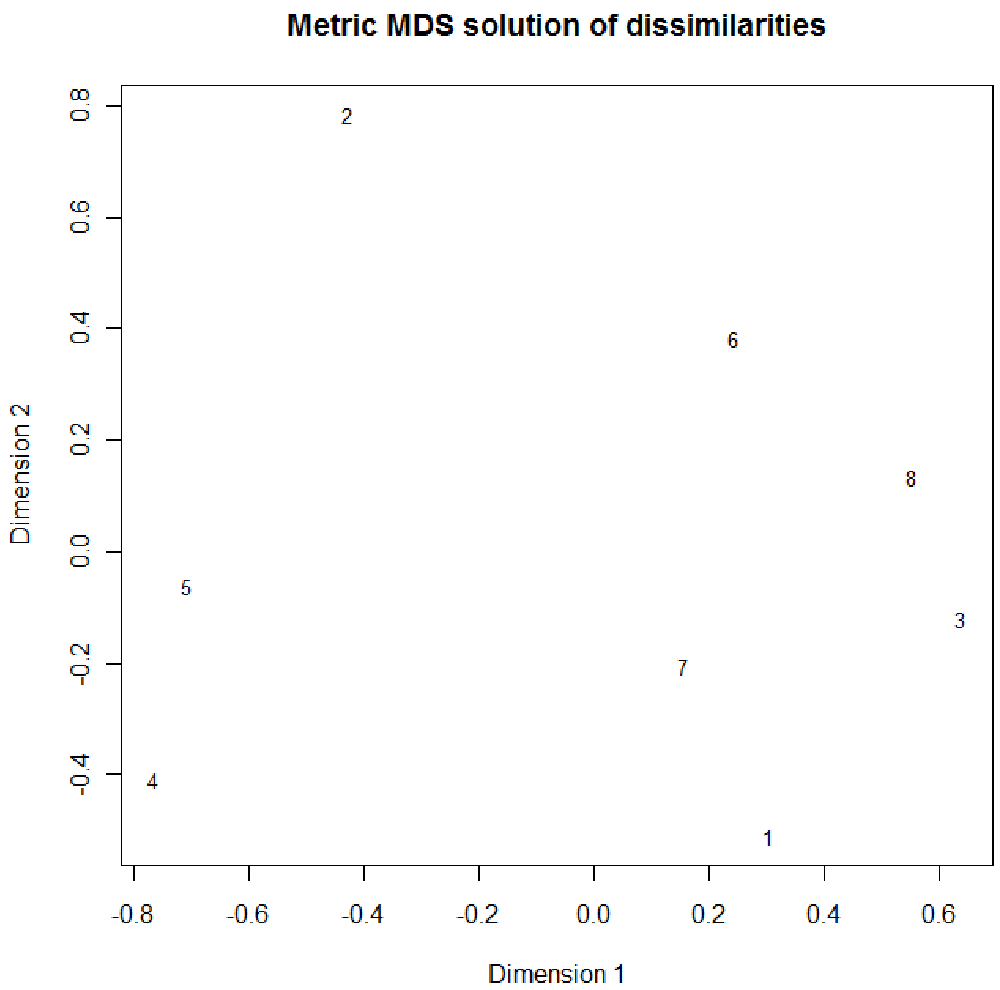

Figure 8.56: *Stimulus space in two dimensions for participant no. 16. This multidimensional scaling plot is solely generated of individual dissimilarity data. The metric stress measure of the solution is 0.020.*