THE AUDITORY FINGERPRINT:
MULTIDIMENSIONAL CHARACTERIZATION
OF INDIVIDUAL PITCH PERCEPTION
IN MUSICIANS

Jan Benner

MSc by Research Thesis 2017
THE AUDITORY FINGERPRINT:
MULTIDIMENSIONAL
CHARACTERIZATION
OF INDIVIDUAL PITCH PERCEPTION
IN MUSICIANS

Jan Benner

Master of Science by Research

University of Salford
School of Computing, Science and Engineering
Acoustics Research Centre

September 2017
A.3 Subject Questionnaire (German) ............................................. 52
A.4 Informed Consent (German) .................................................. 53
A.5 Ethical Approval (German) ..................................................... 54
A.6 Technical Manual - Headphones ........................................... 55
A.7 Technical Manual - Interface ................................................ 56
A.8 Conference Poster (ASA 2017) ............................................. 57

References .............................................................................. 58
List of Figures

1. Index of pitch perception, after Scheider et al. (2005).
2. AAT-test design, after Seither-Preisler et al. (2007).
3. Distribution of pitch perception index, after Ladd et al. (2013).
4. Distribution of subjects across University courses (school affinity).
5. Study population of the joint project, initially matched for musical aptitude (AMMA).
6. Study population of the joint project, initially screened for training intensity.
13. Statistical processing of psychoacoustic data based on PCA and Hierarchical cluster analysis.
14. Distribution of the overall played musical instruments (sorted clockwise) across subjects.
15. Frequency histograms, showing the overall distributions for age and the eight psychoacoustic parameters.
16. Scree-plot, showing the Eigenvalues of all extracted- and selected PCA components (color-coded).
17. Corresponding psychoacoustic parameters in component space after matrix rotation.
18. Individual distributions of component loads show high inter-individual variability.
19. Boxplots, showing mean, range and outliers of individual data for the four selected components.
20. Scree-plot, showing the consecutive agglomeration steps of cluster analysis and the selected 3-cluster solution.
<table>
<thead>
<tr>
<th>Page</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>Denrogram, showing the cluster analysis results for the selected 3-cluster solution (color-coded).</td>
</tr>
<tr>
<td>22</td>
<td>Cluster-specific distributions of school affinity and played musical instruments (sorted clockwise) across subjects.</td>
</tr>
<tr>
<td>23</td>
<td>Barplots, showing the averaged values (normalized) for parameters and components, with significant differences between clusters.</td>
</tr>
<tr>
<td>24</td>
<td>&quot;Auditory fingerprints&quot; averaged for each cluster, showing the combined multi-parametric results of all acquired data, including additional frequency-specific details of the Pitch Perception Preference Test.</td>
</tr>
<tr>
<td>25</td>
<td>Individual &quot;auditory fingerprints&quot;, showing the combined multi-parametric results of all acquired data for three exemplary cases, including additional frequency-specific details of the Pitch Perception Preference Test.</td>
</tr>
<tr>
<td>26</td>
<td>Three different modes of pitch perception preference, distinguished by the three specific subject clusters (colour-coded).</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Statistical tests for normal-distribution and homogeneity of variances.</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Communalities of psychoacoustic data based on the PCA results.</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>Eigenvalues and explained variances shown for each main component after PCA extraction.</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>Rotated component matrix after rotation, showing the component loads for each parameter.</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>Agglomeration schedule of cluster analysis, showing the resulting variance reduction for the selected 3-cluster solution.</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>Cluster-specific descriptive results calculated for all parameters and components.</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>Results from comparative statistical tests, showing significant differences between the 3 clusters.</td>
<td>35</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank my supervisor, Dr. William J. Davies, for his invaluable guidance, help and patience throughout my project. I would also like to thank my friends and colleagues Dr. Peter Schneider and Dr. Maria Blatow, for always fruitful discussions and their continual support, and Dr. Annemarie Seither-Preisler for her extensive statistical help and feedback on my work. I would also like to thank my friends and colleagues Dr. Julia Reinhardt, Dr. Thomas J. Re and Dr Manuela Utzinger, for their persistent help, correction reading and candy supply.
Declaration

All contents of this thesis are the work of the author alone, except where acknowledgement has been made. All verbatim extracts have been distinguished by quotation marks, and all sources of information have been specifically acknowledged. All of the data collection and analysis presented in this thesis are original works by the author, though select pieces of work were jointly researched. Aspects of the joint research project have previously been presented at the following conferences:


Aspects of the present thesis have previously been presented during a poster-session at the following conference:

- Benner, J. et al. (2017). *Differences in sound perception are reflected by individual auditory fingerprints in musicians*. ASA & EAA Conference, Boston. (Appendix 8)

Signed:

September, 2017
Abstract

Musicians have been repeatedly reported to show remarkable inter-individual differences in elementary hearing functions, sound perception mode, musical instrument preference, performance style, as well as musical abilities such as absolute- and relative pitch perception or auditory imagery (audiation). However, relevant literature in the field regarding perceptual and psychophysical aspects of sound and particularly pitch perception is highly contradictory, and subjective differences are mostly unconsidered. Moreover, it is largely unexplored how individual differences in (musical) pitch perception are related to further musical abilities and behavior. In the present work, “auditory fingerprints” were created basically based on a composite of five psychoacoustic hearing tests assessing subjective pitch perception in musicians. A total of 93 musicians, including 49 professionals and 44 amateurs, were individually measured for: (a) pitch perception preference (holistic vs. spectral mode), (b) relative pitch perception (musical interval recognition), (c) absolute pitch perception (“perfect pitch”), (d) frequency discrimination threshold (just noticeable difference), and (e) auditory imagery of tone-sequences. Overall, eight psychoacoustic parameters were extracted and analyzed using statistical methods. In addition, preferences for musical instruments were evaluated. At the individual level, the results show a high inter-individual variability across the eight psychoacoustic parameters, reflecting clear individual differences in pitch perception and related musical abilities. In addition, principal component analysis (PCA) revealed four different main components to sufficiently represent coherent aspects of the psychoacoustic data, namely: tonal musicality, pitch timbre preference, low-band sensitivity and high-band sensitivity. At the group level, multi-parametric cluster analyses revealed three sub-groups of subjects, showing significantly different results with respect to the underlying perceptive patterns. Consequently, at least three different modes of pitch perception are suggested, characterized by: 1. Pronounced analytic pattern recognition, focused on spectra / timbre and sensitive to single tones; 2. Pronounced holistic pattern recognition, focused on (missing) fundamental pitch and rather insensitive to single tones; 3. Less pronounced audiation and pitch detection abilities, linked to ambiguous multi-pitch sensations (“balanced mode”). Taken together, the findings suggest that individual “auditory fingerprints” extracted from psychoacoustic hearing tests, reflect remarkable inter-individual differences, but also typical patterns of perceiving (musical) pitch and sound. It can be concluded, that specific auditory characteristics are related to the individual musical (instrument) preference, style and performance of musicians, as well their learning abilities.
Chapter 1

Introduction and Literature Review

1.1 Ambiguities in Sound Perception

During a curious historic debate between 1840 and 1850, the German physicists G. S. Ohm and A. Seebeck discussed the pitch sensation of complex tones in “Annalen der Physik und Chemie”. Seebeck (1841) produced periodic sounds with the help of a mechanical siren, whilst he controlled the suppression of odd harmonics. The pitch, which he associated with the sound as a whole, always seemed to follow the fundamental component ($F_0$), even if its acoustic energy was weak. In contrast, Ohm (1843) reciprocated that Seebeck’s strong $F_0$ pitch sensation, in absence of acoustic power, had to be based on an illusion. He argued, that our ears perform a Fourier-like real-time frequency analysis. Thus each distinct pitch sensation originates from a corresponding sinusoidal wave (harmonic) of a complex tone - Ohm’s well-known “definition of tone”. Thereby the frequency of the lowest spectral component determines the pitch of the complex, while the other harmonics shape the timbre of a sound. After a further controversial debate (Seebeck, 1843; Ohm, 1844) Seebeck finally closed the story by confessing, that higher harmonics produced by the siren are audible but difficult to distinguish. He finally concluded, with a freely extended, however contradictory interpretation of Ohm’s law, that the sound as a whole may produce a distinct pitch corresponding to the periodicity of the sound pulses (Seebeck, 1844a; Seebeck, 1844b). Based on the colliding statements about their subjective self-sensations, they unintentionally reported highly individual perceptual differences in pitch sensation of complex tones for the first time in research history. Two decades later Ohm’s statement was picked up by the German physicist and physiologist H. von Helmholtz, who has been known for his well trained musical ears. In his work on “the sensation of tones as a physiological basis for the theory of music” (von Helmholtz, 1863) he pointed out characteristic subjective differences in sound perception. Basically he described two ways in which one may perceive a complex tone: a synthetic mode, whereby harmonics “fuse in the whole mass of musical sound” and an analytic mode, based on the separate perception of single harmonics. Consequently, he suggested that Seebeck (presumably being a synthetical listener) might not have been able to hear higher harmonics separately, as compared to Ohm (presumably being a analytical lis-
tener). Further, von Helmholtz discussed the difficulties of observing harmonics of complex tones, sensation of tone-color, as well as the challenges of subjective sound analysis and the role of attention, in a remarkably detailed manner. Two decades later the Finnish phonetician K. H. Pipping (1895) revived von Helmholtz’s hearing theory and introduced his idea of two possible modes of pitch perception while perceiving a complex tone (here “clang” refers to the German word “Klang” used by von Helmholtz):

“For a complex tone, we may direct our attention to the partials (tone pitch), but it is also possible to pay attention to the total impression of the sound (clang pitch); the latter pitch is not influenced by the absence of the fundamental, a group of partials being sufficient to perceive this pitch.” (Pipping, 1895; as cited in Plomp, 1967, p. 1528)

Here “tone pitch” may correspond to von Helmholtz’s descriptions of an analytic, and “clang pitch” to a synthetic mode.

Beginning of the 20th century electronic devices could then be used for sound generation and psychoacoustic measurements. The Hungarian physiologist and telephone engineer G. von Békésy proved the frequency dependent localization of auditory stimuli in the cochlea, as anticipated by von Helmholtz (Békésy, 1928). Afterwards the Dutch mathematician J. F. Schouten reanimated the historic debate of Seebeck and Ohm, by demonstrating that the pitch sensation of periodic pulses generated by an optical siren (1/20th of the repetition time of 1/200 s), devoid of the F0 component \( f_0 = 200 \text{Hz} \), is associated with the corresponding frequency of the missing fundamental (MF; Schouten, 1938). He concluded, that this specific sensation could not be explained by a nonlinear difference tone emerging physically at the auditory periphery. A conclusion which was in contrast to prior conclusions of von Helmholtz and the American physicist (and telephone engineer) Fletcher (1924). Thereby, Schouten claimed that Seebeck’s initial argument was basically correct. In further experiments Schouten observed the periodicity of the envelope pattern of harmonic clusters to be the same as the periodicity of the F0, even if the F0 component was absent (Schouten, 1940a; Schouten, 1940b). He suggested, that such pitch sensations would be caused by insufficient spectral resolution of the cochlea, based on neural detecting of periodic fluctuations in the envelope pattern of harmonic clusters. As an essential statement of Schouten’s “residue theory of pitch” the limited spectral resolving power of the inner ear may result in a sensation corresponding to the MF:

“The lower harmonics can be perceived individually and have almost the same pitch as when sounded separately. The higher harmonics, however, cannot be perceived separately but are perceived collectively as one component (the residue) with a pitch determined by the periodicity of the collective waveform, which is equal to that of the fundamental tone.” (Schouten, 1940b, p. 991)

Later on R. J. Ritsma (1962) observed a clear upper limit in the harmonic order \( h \) for the perception of a tonal residue, as well as an extending existence region towards cochlear-
resolved harmonics. Consequently, Schouten’s residue theory failed for an adequate explanation of pitch. Ritsma found that the existence region could be roughly described by values of $f < 5000\,\text{Hz}$ and harmonic order $h \leq 20$. Further, in the 1970’s Ritsma and R. Plomp independently postulated a dominance region for the strongest sensation of the $MF$, based on the perception of resolved harmonics of lower order $h < 7$ (Ritsma, 1967; Plomp, 1967). The following experiments of A. J. M. Houtsma and J. L. Goldstein (1972) indicated that the pitch of complex tones is mediated by a central processor that operates neural signals from spectrally resolved harmonics in the cochlea. In a parallel important key experiment by G. F. Smoorenburg (1970) two possible modes of hearing were described, based on at least 42 subjects (coworkers of his institute) in two repeated runs (one month follow-up). He investigated in which way complex tones consisting of just two harmonic components (here defined as $f_0/h, h$ with $f_0 =$ frequency of the $MF$ and $h =$ order of present harmonics) were perceived with respect to pitch. The first complex (complex-$\alpha$: $200/9, 10$) consisted of harmonics with frequencies $f_1 = 1800\,\text{Hz}$ and $f_2 = 2000\,\text{Hz}$, the second (complex-$\beta$: $250/7, 8$) accordingly $f_1 = 1750\,\text{Hz}$ and $f_2 = 2000\,\text{Hz}$. Each signal was produced by two sine-wave generators without mutual synchronization at a sensation level of $40\,\text{dB}_{SPL}$ presented binaurally with headphones. Presentation times of each signal were $160\,\text{ms}$, separated by a pause of equal duration. The two signals were repeated 25 times per run in a random order. In addition, a distinct noise band, masking simple tones of $200\,\text{Hz}$ and $250\,\text{Hz}$, was used in order to avoid noticeable difference tones ($f_2 - f_1$) according to Plomp (1965). Smoorenburg also was aware of possible combination tones of the type $f_1 - k(f_2 - f_1), f_1 < f_2$ according to Plomp and Goldstein (Plomp, 1965; Goldstein, 1967). After presenting the two signals successively (e.g. complex-$\alpha \rightarrow$ complex-$\beta$), subjects had to decide in a forced-choice task if the perceived tonal interval was either ascending or descending. The former should correspond to a $MF$ tracking cue, the latter to a tracking cue following the harmonic components itself:

“Taking into account the presence of combination tones, there is still the following unequivocal relation: if the pitch of the signal $f_1, f_2 = 1750\,\text{Hz}, 2000\,\text{Hz}$ is judged to be higher than the pitch of the signal $1800\,\text{Hz}, 2000\,\text{Hz}$, then the judgment must have been based upon the complex tone as a whole; if the pitch is judged to be lower, then the judgment must have been based upon one or more pitches of individual part-tones.” (Smoorenburg, 1970, p. 927)

Interestingly the subjects could be subdivided in two groups of equal size (bimodal distribution), pursuing one of the possible strategies consistently. Consequently, he concluded that analytic and holistic cues contributed to explain these two possible modes:

“We may infer that the perception of a pitch jump corresponding to the fundamental frequencies was based upon pitches of the complex tones perceived as a whole and that the judgments in opposite direction were based upon the pitches of individual part-tones or perhaps just upon timbre.” [...] “It merely
should demonstrate that there are important individual differences.” (Smoorenburg, 1970, p. 928)

Thus, Smoorenburg’s experiments were the first systematic explorations of individual differences based on ambiguous pitch sensations in response to MF complexes, according to the historic assumptions of von Helmholtz. That the perception of a residue differs between subjects was later confirmed in important experiments by E. Terhardt (1974), as well as E. De Boer (1976). This let De Boer to state:

“Now, when one listens to a tone, two attitudes are possible. One is listening to the sound as a whole, appreciating its pitch and perceiving the timbre as a characteristic quality of the entire sound. The other attitude is one of subjective analysis: one tries to break up the sound into constituent sounds (which happen to correspond to sinusoidal components), and the qualities of the sound as whole are lost.” (De Boer, 1976, p. 490)

However, due to the diverse conditions under which the experiments were performed (and, to large extent, personal preference) pitch terminology was inconsistently used and varied broadly, for example: residue pitch (Schouten, 1940b), low pitch (Thurlow, 1958), periodicity pitch (Licklider, 1951), time-separation pitch (Small, 1955), repetition pitch (Bilsen, 1966) and virtual pitch (Terhardt, 1974; Terhardt, 1979). Moreover, all these concepts could neither explain the manifold ambiguities in pitch perception, nor explain sufficiently which conditions decide whether analytic or holistic pitch cues were used. However, the underlying concept of pitch perception might be of a more multi-dimensional nature as it appears. Recent research in the field suggests the existence of different “axes” of sound perception. Consequently, musical pitch might not exist independently from, but could instead be understood as a distinct dimension of timbre (Warren, 2003; Halpern, 2004; Marozeau, 2007; Bizley, 2009; Ladd, 2013; Vurma, 2014; Allen, 2014; Schellenberg, 2015). In this context the Austrian composer A. Schönberg (1911) argued:

“I cannot readily admit that there is such a difference, as is usually expressed, between timbre and pitch. It is my opinion that the sound becomes noticeable through its timbre and one of its dimensions is pitch. In other words: the larger realm is the timbre, whereas the pitch is one of the smaller provinces. The pitch is nothing but timbre measured in one direction.” (Schönberg, 1911; as cited in Schneider 2009a, p. 319)

1.2 Individual Pitch Perception Mode

Today there is conjoint evidence for Smoorenburg’s basic conclusion of at least two different modes of pitch perception, suggesting a robust underlying phenomenon, even when experiments methodologically diverge substantially (Renken, 2004; Schneider, 2005a; Seither-Preisler, 2007; Schneider, 2009a; Ladd, 2013; Coffey, 2016). Further, similarities to the field
of soundscape research are apparent, as the basic observations from Smoorenburg and De Boer are reminiscent of two *modes of listening* proposed by W. W. Gaver (1993), who noted different ways of experience natural or technical sounds:

“ [...] it is possible to hear any sound in terms of its source (everyday listening) or in terms of its sensory qualities (musical listening). [...] both ways of experiencing the sound are valid.” (Gaver, 1993, p. 286)

In this context, recently W. J. Davies (2015) remarked the similarities between the research fields of soundscapes, music cognition and audio quality, as they deal with the same underlying perceptual and cognitive phenomenons. He also concluded that there exist different ways of perceiving a certain soundscape: either as a whole, or by “zooming” into a certain sound within it, or even into a specific component or feature of that sound. Moreover, many researchers frequently observed possible octave-shifts while listening to harmonic complex tones, corresponding to frequencies up to more than an octave above the (missing) \( F_0 \), as reported by Ritsma and Engel (1964), Terhardt (1972), Moore (1977) and Patterson (1990). Consistently, von Helmholtz (1863) had already noticed, that even professional musicians and acousticians experience *octave-ambiguities* frequently, whereas later Davis (1951) reported, that octave-ambiguities are the most common “errors” of musicians. This is in accordance with Schönberg (1911) who also noted the similarity of *octave-shifted* tones. In a later key experiment, Schneider et al. (2005a) performed a psychometric *Pitch Perception Preference Test* consisting of 162 different pairs of harmonic complex tones based on the concept of Smoorenburg’s experimental paradigm (1970) comparing the pitch sensation of two successively played complex tones (see chapter 2.2.1 for details). The highest harmonic component was always maintained constant in-between a pair of tones, in order to avoid changes in timbre. Combining *MF* complexes with additional complete complexes (including a physically present \( F_0 \)), allowed the detection of octave-shifted pitch perception, for example, one or two octaves above \( F_0 \). This case mainly occurred for harmonic complexes consisting of three components within a higher spectral range > 1000\( \text{Hz} \) and was suggested to show an additional third perceptual mode that is not part of the \( F_0 \) tracking. In this study Schneider et al. showed in particular that the perceived musical pitch of harmonic complex tones varied largely by up to three or four octaves, when the same sound was presented to different individuals \( (N = 420) \), including 181 music students, 125 professional-, 66 amateur-, and 48 nonmusicians. After correcting for octave-shifted percepts, the results showed that some subjects recognized dominantly the fundamental pitch \( (F_0, \text{holistic listeners}) \), whereas others perceived dominantly spectral aspects of the complex \( (SP, \text{spectral listeners}) \), irrespective of the musical state. The resulting broad bimodal distribution of subjects following either \( F_0 \) or \( SP \) pitch allowed for a classification of these two modes of pitch perception, in accordance with Smoorenburg’s concept (Figure 1). Thereby holistic listeners perceive pitch, *chroma*, and timbre as qualities of the entire sound, whereas spectral listeners tend to decompose the sound into groups of harmonics or certain spectral components. Overall, specifically musicians tend to have a more dominant prefer-
ence either towards extreme \(F0\) or \(SP\) perception. Using additional neuroimaging techniques, in a subgroup of 87 subjects, they observed the preferred mode of pitch perception to be reflected by structural and functional asymmetries of lateral *Heschl’s gyrus* - a anatomical structure known for containing areas of primary and secondary auditory cortex. Therefore, they concluded the existence of two specialized pitch centers, whereby (missing) \(F0\) pitch is extracted by the left- and spectral pitch by the right auditory cortex, which is in line with evidence regarding hemisphere-specific neural processing of temporal and spectral auditory information (Zatorre, 2001; Schneider, 2009a). In a second evaluation based on the same study, they further found evidence that individual preferences in pitch perception correspond to specific preferences for music instruments and musical performance style (Schneider, 2005b).

In a basically similar \(MF\)-task experiment Seither-Preisler et al. (2007) found complex tones with conflicting \(F0\) pitch, and parallel changes in timbre, to be heard differently by musicians and nonmusicians. In accordance with Schneider et al. they also reported a dichotomic distribution separating \(F0\) pitch listeners from \(SP\) pitch listeners. Furthermore, they also observed many subjects with inconsistent and inhomogeneous pitch perception leading to the assumption of “guessing” in such subjects, which might be also due to the conflicting interaction of parallel pitch- and timbre changes (Figure 2). They also suggested the results of Schneider et al. to be not bimodal, but with a bias towards \(F0\) responses, if octave-shifted percepts would be counted as \(F0\) percepts. In contrast to Schneider et al. they reported an effect of musical training leading to a more \(F0\) focused perception. They therefore concluded, that this perceptual bias may be related to intensity and starting age of musical training (Seither-Preisler, 2009).

In a recent study Ladd et al. (2013) confirmed previous observed individual differences in the perception during \(MF\)-tasks by comparing the overlapping results from Schneider et al. and Seither-Preisler et al. with results from their own different experiments. They performed
seven testings with \( N = 412 \) subjects, including 23 musicians, using an experimental MF-task similar to the concept of Schneider et al. By using many repetitions per testing and stimuli, they focused on re-test reliability, perceptual stability over time, and possible inconsistencies in the answering behavior. They found a distribution of \( F_0 \) and \( SP \) pitch perception with a bias towards \( F_0 \) percepts, due to the fact that about a quarter of subjects showed no consistent preference (Figure 3). This is in line with the observations of Seither-Preisler et al. Therefore, the results seem to diverge from the bimodal distribution of Schneider et al., but on closer inspection are explainable with respect to the effect of octave-ambiguities. However, Ladd et al. showed that despite divergent experimental details used in different MF-task based studies, the test reliability is high and the perceptual responses of subjects are remarkably consistent over time. Also, musicality and gender showed no influence, consistent with Schneider et al., but age showed a slight effect with older subjects more likely to give spectral responses. The latter effect might be rather caused by age related physical changes in the inner ear or even cortical alterations. Furthermore, they recognized that observed consistent intermediate responses (meaning neither extreme \( F_0 \) nor \( SP \) perception) can not be explained simply by the existence of just two types of listening. Therefore, they also suggested a possible additional mode of perception following the octave-ambiguity phenomenon of Schneider et al. They concluded that more than just two general modes of sound perception must exist and therefore the assignment of just one of two basic types of pitch perception to individuals would oversimplify the nature of underlying robust individual differences. Thus, it could be more reasonable to speak of two modes of pitch perception, that might be available in different combinations to different listeners, instead of speaking of two types of listeners themselves.

As the early psychoacoustic experiments investigated the basic perception of sinusoidal and complex tones - mostly based on self-sensation or including only few participants with unknown musical background - recent investigations focus on more realistic methodological conditions representative also of the average population. Researchers started to develop several methods of measuring auditory thresholds and pitch perception in order to investigate

![Figure 2. AAT-test design, after Seither-Preisler et al. (2007).](image)
specific psychoacoustic questions. Besides many studies mainly based on nonmusicians, recently musicians have become more and more an ideal model for exploring different dimensions of sound perception and auditory processing in detail, as musicians are known to reach extreme auditory thresholds and show pronounced perceptual effects. Consequently, several listening-tests were developed to focus on more specific musical skills, such as absolute- or relative pitch perception, as well as cognitive higher-level competences, such as auditory imagery (audiation) and musical aptitude.

1.3 Relative and Absolute Pitch Perception

Relative pitch (RP) perception is the ability to recognize more or less instantaneously the relative distance between two given tones, independent of their absolute localization in the tone space. While the concept of RP plays a crucial role in traditional professional music education systems, such as solmization methods (Choksy, 1999), novel musical education concepts based on evidence from neuroscience (Parn cott, 2002; Hodges, 2012) support the notion, that competences in RP perception are at the basis of professional musical expertise. While musically experienced listeners perform often better in interval recognition tasks compared to nonmusicians, musical education alone cannot explain excellent RP abilities and high performance in interval recognition alone (Denham, 2016). In fact, Thompson et al. (2012) revealed effects of intensity and non-spectral properties of sound on RP and interval size judgement. Their findings extend previous evidence, showing the influence of spectral attributes such as tonal context, timbre, and overall pitch height on RP (Krumhansl, 1979; Russo, 2005a; Russo, 2005b), as well as influences by acoustic attributes other than $F_0$ (Russo, 2005a; Russo, 2005b; Thompson, 2010). These findings are in line with a study by McDermott et al. (2010), which studied 122 musicians and 143 nonmusicians. They found evidence for individual differences in preferences for specific properties of RP and specific musical intervals, based on preferences for consonance, direction and harmonicity of spec-
Further, based on correlations with musical training data, they concluded that exposure to music intensifies preferences for harmonic frequencies because of their musical importance. They also noted the role of individual preferences for acoustical properties. Overall, these unique findings on RP perception suggest, that identical intervals might be perceived individually different. However, both the perceptual and the neural basis of RP processing is largely unexplored with respect to its multidimensional character, individual variability and the influence of musical training. In addition to RP perception, absolute pitch (AP) perception is the rare ability to recognize or produce the pitch of any given tone spontaneously without an external reference. The prevalence for AP abilities is estimated to be about 0.01% in the general population, but up to 7 – 32% in professional musicians (Baharloo, 1998; Gregersen, 1999). High AP performance also correlates with increased performance in musical dictation (Dooley, 2010) and is associated with a large auditory digit span (Deutsch, 2013). However, individuals may possess AP abilities in varying degrees. These range from absolute perception of specific notes, tonalities or instruments (partial AP) up to outstanding abilities including perception and production of any tone irrespective of the kind of sound (perfect AP). AP perception recently became a more popular research focus in relation to brain functions and development (Zatorre, 2003; Ross, 2005), perceptual and cognitive aspects (Vanzella, 2010; Elmer, 2015), distinct genetical traits (Zatorre, 2003; Athos, 2007; Theusch, 2009), as well as neuroanatomical correlates (Dohn, 2015). In particular, there is recent evidence for a multisensorial AP network in the brain’s right hemisphere, which appears to integrate primary auditory, sensory-motor and language related areas (Wengenroth, 2014). Further, there is evidence that AP performance is mostly independent of age, musical training, gender, or familiarity with specific stimuli (Ross, 2003; Ross, 2009; Jakubowski, 2016), overall suggesting an innate and stable ability that persists into adulthood. However, absolute pitch may not be as absolute as it seems with respect to quite plastic tone-labeling mechanisms. These may not necessarily be based on early musical experience but rather on the adopted cultural norms for different tunings in music (Hedger, 2013). However, across the scientific community RP perception is often assumed to be the prevailing perceptual “default mode” of musicians in the absence of AP perception (Schlaug, 1995; Pantev, 1998) irrespective of the fact, that excellent RP abilities may be as well rare among musicians. In fact, there is growing evidence for overlaps between AP and RP networks in regard to brain function and anatomy (Zatorre, 1998), electrophysiological correlates (Itoh, 2005), pitch memory (Schulze, 2009), music recognition (Creel, 2012), pitch matching of chords (McLachlan, 2013), global vs. local processing (Ziv, 2014), and correlated enhanced performance (Dooley, 2011). AP abilities also seem to accompany rare phenomenons such as synesthesia (Gregersen, 2013) or specific pathologies such as autism (Heaton, 1998; Mottron, 1999). There is further evidence for overlapping ethnicity effects: while AP abilities occur at a higher rate among East Asian musicians, similar ethnicity effects were found for RP abilities amongst Chinese and Korean populations. Despite Asian populations consistently outperform people from other origins in AP and RP tasks, this effect is not necessarily driven by previous musical or tone-language experience (Hove, 2010).
neuronal mechanisms behind RP and AP perception are still not clarified, both abilities play an important role for professional musicians and may have great influence to their individual auditory sensation, behavior, and performance. So far no standardized listening test for the gradual assessment of RP ability exists. However, a first version of a quantifiable computer-based “Interval Recognition Test” was recently developed by E. Hofmann (Music-Academy Basel), providing different task difficulties by considering the influences of interval size and interval direction (unpublished; see chapter 2.2.2 for details). A gradual AP perception test based on passive and active AP performance was developed by Wengenroth et al. (2014) (see chapter 2.2.3 for details).

1.4 Frequency Discrimination in Musicians

The ability to discriminate two nearby oscillating simple or complex tones is part of the elementary auditory discrimination functions, and can be assessed by measuring the just noticeable difference (jnd) in audiometric experiments. An excellent detection of small pitch differences up to < 2cent (1/50 semitone) might be essential for most professional musicians, as a key ability for tonal intonation. However, there is evidence for a high individual variability in musicians and nonmusicians. In particular, differences in frequency discrimination (FD) thresholds can generally differ about more than a factor of 10, as observed by H. Fastl A. Hesse (1984), in a experiment using six different test frequencies (125 – 4000Hz) and eight different tone durations (2 – 500ms). In a recent study, differences of FD values by a factor of 100 to 1000 have been observed, ranging from 1cent in some professional musicians up to 300cent (= 3 semitones) in some nonmusicians (Serrallach, 2016). Further, the ability to discriminate frequencies, tone-durations and tone-ramps was found to differ accordingly to the presence of auditory related disabilities, e.g. ADHD or Dyslexia. However, these differences might be partially based on changes over time, as a function of age, (musical) training and experience, or in response to environmental factors, as shown previously by C. K. Madsen et al. (1969) and M. F. Spiegel C. S. Watson (1984). Thus, it might not be surprising that musicians show in general smaller differences in auditory discrimination than nonmusicians (Fastl, 1984), whereas initial FD thresholds in musicians were observed to be about one-third the size of those in nonmusicians (Spiegel, 1984). In a more recent combined audiometric and electrophysiological study, Tervaniemi et al. (2005) used repeated complex tone stimuli (528/1 – 4; frequency changes: 0.8%, 2%, or 4%) during auditory ERP (event-related potentials) recording with EEG (electroencephalography), to focus on attended versus unattended pitch processing accuracy, in 13 musicians and 13 nonmusicians. They observed that, in general, musicians detected pitch changes generally faster and more accurately than nonmusicians. Further, musicians showed increased pitch discrimination accuracy not only for 0.8%, but also of 2% changes, when compared to the nonmusicians. Additional results of the EEG analyses suggested, that musical expertise affects merely attentive levels of FD processing but not necessarily pre-attentive levels. However, cognitive FD training can exert an influence on FD thresholds, and has been shown to not only be beneficial for musicians.
(Demany, 1985; Irvine, 2000; Delhommeau, 2002), but also for the treatment of tinnitus e.g. (Flor, 2004; Schneider, 2009b).

A reliable FD test was developed and provided by Stephan Ewert (University of Oldenburg) within a customizable audiometric “Alternative Forced Choice” (AFC) test package (see chapter 2.2.4 for details).

1.5 Auditory Imagery and Audiation

Auditory imagery is the ability to memorize sequences of sound while hearing them by the “inner ear”, a process also known as musical imagery if specifying the memorization of musical phrases. In one of the earliest experimental studies in this field, J.R. Bergan (1967) found relationships between music imagery, memory and pitch identification:

“One may define an auditory image as an auditory experience of realistic dimensions for which there is no apparent physical stimulus. By realistic dimensions is meant that the auditory image tends to be a replication of an auditory experience initiated in the environment. Thus, in its purest form there would be no experiential difference between an auditory image and an actual sound.” (Bergan, 1967, p. 99)

Accordingly, based on comprehensive research related to musical education, E. Gordon (1997) specifically described musical imagery as a process of “inner hearing” corresponding with his definition of audiation, meaning:

“Audiation is the ability to hear and comprehend music for which the sound is not physically present.” (Gordon, 1997, p. 46)

However, empirical findings suggest that auditory imagery is related to basic and higher auditory processing, by which the brain gives context to musical sounds in a mental scanning process, which can be understood as the musical equivalent of thinking in language or pictures (Halpern, 1988). Further, auditory imagery features aspects of pitch, timbre, loudness, as well as complex nonverbal auditory stimuli such as musical contour, melody, harmony, tempo, notational audiation, and environmental sounds. Not surprisingly, individual differences were observed with respect to perception and memory (detection, encoding, recall, mnemonic properties) considering relations with musical ability and experience (Hubbard, 2010). Concerning the relations between qualities in audiation and pitch perception, Bergan (1967) noted:

“The significance of the relationship between accuracy in pitch identification and musical imagery with respect to musicianship is that it suggests that the critical function of being able to make judgments concerning the pitch of sounds does depend on adequate internal representation of the sounds being judged.” (Bergan, 1965; as cited in Bergan, 1967, p. 99)
Recent neuroimaging studies in this field confirm, that mental auditory imagery for familiar melodies significantly induces activation in the frequency-responsive areas of primary auditory cortex (Oh, 2013). These findings point towards top-down pitch processing mechanisms in the auditory cortex similar to that used during the perception of external acoustical stimuli, demonstrating that auditory imagery functions are comparable to auditory perception (Vuvan, 2011). Furthermore, music-pedagogical evaluations confirm a strong relationship between performance in audiation and musical aptitude (Gordon, 1988; Gordon, 1998), while musical aptitude represents the potential to learn music, and is suggested to stabilize at the age of about nine years, prior to intensive musical education (Schneider, 2005a). On the other hand, there is also evidence for influence from musical training and musical experience on musical imagery abilities (Aleman, 2000), though auditory vividness and mental control have been shown to be more influential on pitch imagery performance than musical experience (Gelding, 2015). However, recent studies indicate considerable benefits for musicians compared to nonmusicians, with respect to increased musical notation performance (Brodsky, 2003), increased musical synchronization abilities (Pecenka, 2009), increased short term memory- (Williamson, 2010) and enhanced encoding of musical sequences (Brown, 2013), as well as increased musical skills linked to vividness of motoric imagery (Di Nuovo, 2016). Consequently, there is broad evidence that distinct audiation abilities can be interpreted as a core element of musicality reflected by Gordon’s “music learning theory” (Gordon, 2012) as a widely accepted model for music aptitude (Shuter-Dyson, 1999). Moreover, there is evidence for correlates with cortical structures in the brain, e.g. shown for (involuntary) musical imagery (Farrugia, 2015). In particular, anatomical size and electrophysiological source activity of distinct areas in auditory cortex show high correlations with the tonal raw score of Gordon’s standardized “Advanced Measures of Music Audiation” (AMMA) test (Gordon, 1989; see chapter 2.2.5 for details), revealing also large differences between musicians and nonmusicians (Schneider, 2002; Schneider, 2005a).

1.6 Research Questions and Objectives

Taken together, the research findings outlined above suggest the existence of large individual differences in musical pitch- and general sound perception on different perceptual and cognitive levels. The psychoacoustic assessment of basic and complex auditory parameters might reveal individual different patterns of sound perception, reflecting characteristic “auditory fingerprints” that may have suspected relevance for acoustical preferences and musical performance style. However, with regard to the particular relevance for (professional) musicians, there is so far no specific knowledge about the importance of individual auditory profiles in a musical context. Moreover, how distinct parameters of sound perception interact and how the individual interplay on the perceptual level can be reflected by different auditory profiles is unclear. Consequently, the aim of this work is to access the basis of auditory behavior on a perceptual level with a focus on music related pitch perception. Therefore it is necessary to quantify different elementary dimensions of individual sound perception.
by distinct auditory measurements. Exploring the parameters from corresponding psychoacoustic listening tests can reflect individual differences on multiple levels of musical hearing with a specific focus on pitch perception. Moreover, exploring differences and similarities between individual auditory fingerprints may help to classify characteristic patterns of sound perception with potential implications for musicians. Consequently, the research objectives of the present work are:

(i) Monitoring essential parameters of sound perception with a specific focus on the individual sensation of pitch

(ii) Identification of characteristic patterns of pitch perception on the individual and group level

(iii) Characterization of individual auditory profiles ("fingerprinting") and identification of common clusters of sound perception

(iv) Illustration of the inter-individual auditory variability and description of the characteristic sound perception modes with respect to musical abilities

Performing a selected composite of five psychoacoustic listening tests with 93 musicians (see chapter 2 for details) should provide a sufficient data base in order to investigate individual patterns of sound perception, with respect to:

(a) Pitch perception preference (holistic vs. spectral mode)

(b) Relative pitch perception (musical interval recognition)

(c) Absolute pitch perception ("perfect pitch")

(d) Frequency discrimination threshold (just noticeable difference)

(e) Auditory imagery of tone sequences (audiation)

Furthermore, besides the individual perceptual pattern analyses, group specific differences and similarities should be explored based on clustering- and principal component analyses (see chapter 2.3 for details), in order to delineate also typical patterns of auditory perception.
Chapter 2

Methods

2.1 Study Setting and Subjects

All experimental measurements and analyses were performed within the scope of a combined cross-sectional and longitudinal imaging study titled “Auditory neuroplasticity in the adult musical brain”, a SNF (Switzerland) and DFG (Germany) funded collaboration conducted by Dr. Maria Blatow (Swiss part) and Dr. Peter Schneider (German part). The main research aspects of this internationally funded research project included elementary auditory perception, musical abilities and corresponding neural correlates of hearing in the brains of musicians, with a focus on audio- and neuroplasticity. All subjects gave their informed consent to participate in the experiments approved by the local Ethics committee (see Appendix 4,5). The present work represents a subset of this larger project. Specifically, it is limited to the consideration of auditory perception and musical abilities, and does not utilize neuroimaging. All subjects were recruited, screened and measured by the author of the present work exclusively.

A population of 93 musicians (47 male / 46 female; mean age: 21.8 ± 2.6) including 49 musical students (28 Classical, 11 Jazz, 10 Early Music) was monitored at the beginning of their three year intensive University Bachelor program at the Music-Academy Basel (hereafter professionals). In addition, a control group of 44 hobby-musicians (medical students from the Medical Universities Basel and Heidelberg) were measured in parallel at the beginning of their respective University course (hereafter amateurs). For the present work, the distinction between professionals and amateurs will be named “state” and the deeper distinction between the specific University courses (professionals: Classical, Jazz, Early Music; amateurs: Medicine) will be named “school” (Figure 4).

All subjects were asked to complete an accompanying questionnaire (Appendix 3) by self-report, including the following queries:

- Name (anonymized by consecutive numbers), age, handedness.

- University program (Music, Medicine) and the respective major course for music students (Classical, Jazz, Early Music).
• Primary musical instrument played actively during life and to at least University start. In addition, the corresponding musical practice times [averageh/week] were acquired for three periods across the whole lifespan (childhood, adolescence, study/University start).

![Diagram of subject distribution across University courses](image)

Figure 4. Distribution of subjects across University courses (school affinity).

Both groups were initially matched for raw tonal score of AMMA test (Figure 5) as a level of elemental music aptitude, described in detail in chapter 2.2.5 (professionals: 32.8 ± 3.6, amateurs: 30.6 ± 3.4, inclusion border: 25). The AMMA score matching was suggested to guarantee a minimum requirement with respect to elemental musical abilities and aptitude reflecting an equal intrinsic musical potential for both, professional- and amateur musicians.

![AMMA raw tonal score distribution](image)

Figure 5. Study population of the joint project, initially matched for musical aptitude (AMMA).

Further, both groups exhibited a similar musical training intensity during childhood and early youth, but significantly diverged in this respect during later adolescence and early adulthood (figure 6), when they started their respective professional education (mean intensity [h/week]: professionals: 19.7 ± 7.5, amateurs: 6.4 ± 3.5) playing one or more instruments (including singing).
The above described AMMA score matching as well as the data on training intensity is based on a reduced subset of 30 professionals and 30 amateurs that were hand-picked from the total population of 93 subjects. That subset of 60 subjects was specifically determined and matched for the purpose of follow-up experiments within the scope of the longitudinal part of the comprehensive imaging study. Thus, data on training intensity and AMMA score matching is just mentioned descriptively at this point in order to reflect the initial study conditions. Consequently, the subset data was not used for further analyses in the present work, as it is not representative for the total population of investigated subjects.

### 2.2 Psychoacoustic Tests

Five psychoacoustic tests were performed computer-based in order to gradually quantify individual sound perception abilities based on: frequency discrimination thresholds, pitch perception preference, absolute- and relative pitch perception skills and audiation abilities. All subjects were tested individually in a isolated and silent test room at the research facilities of University Hospital Basel. The separate test sounds were exclusively presented through closed dynamic Sennheiser HDA-200 audiometric headphones (frequency response: 20 – 20000Hz; passive attenuation: 14dB at 125Hz to 44dB at 8kHz), connected to a mobile RME Fireface-400 (D/A conversion: 24bit, sampling rate: 48kHz) sound interface, in order to guarantee a solid sound production with a high signal-to-noise ratio and a nearly linear frequency response (see appendices 6, 7 for technical manuals). The average sound pressure
level was initially calibrated to $65\, dB_{SPL}$ measured at the headphone inside using a BrüelKjaer 2203 Sound Level Meter. Based on recommendations by Schneider et al. (personal communication) this presentation level is considered sufficient to avoid noticeable harmonic distortions or interfering combination tones. The background level in the test room was not specifically measured, as the used technical periphery provided appropriate test conditions and no specific loudness-sensitive tests were applied. Two psychoacoustic test sessions with equal total test times of approx. 45min respectively, were synchronized with two separate neuroimaging sessions on different days, in the course of the measurement schedule of the comprehensive study. The following test-sequences were applied for every subject in a consistent order:

**Session 1 (approx. 45min):**

(a) Pitch Perception Preference Test (Schneider, 2005a), approx. 20min

(b) Musical Interval Recognition Test (E. Hofmann, 2012, Music Academy Basel, unpublished), approx. 15min

(c) Quantitative Absolute Pitch Test (adapted version) (Wengenroth, 2014), approx. 10min

**Session 2 (approx. 45min):**

(d) Frequency Discrimination Threshold Test (part of AFC-Test package, S. Ewert, University Oldenburg), approx. 30min

(e) Advanced Measures of Music Audiation Test (AMMA) (Gordon, 1989), approx. 15min

### 2.2.1 Pitch Perception Preference Test

The dominant mode of pitch perception was determined using the Pitch Perception Preference Test, developed during 1995 to 1997 by P. Schneider and S. Bleeck (Schneider, 2005a) based on a ambiguous tone-interval paradigm with 162 tone pairs. Each tone-pair consists of two consecutive harmonic complex tones (tone duration: 500$ms$, ramp on-/offset: 10$ms$, inter-stimulus interval: 250$ms$). The test tones vary in number ($n = 2$ to 4), order ($h = 2$ to 16) and average spectral frequency ($f_{SPav}(1−6) = 0.25$ to $5.0kHz$) of the corresponding harmonics. The sequence of the stimuli was determined randomly. Certain harmonics which characterize timbre (e.g. upper partials) are deliberately kept constant within a tone-pair in order to minimize noticeable timbre changes. In a two-way forced choice task subjects were instructed to decide whether they perceive the second tone of a tone pair as higher or lower compared to the first tone. The perceived direction of the tone shift is upward (ascending interval) or downward (descending interval), depending on the subject’s dominant sound perception mode (spectral or holistic). Subjects were instructed, that in some cases tones might lack a clear pitch but instead groups of resolved partials with different pitches
could be heard, both ascending and descending. In those cases they were asked to judge the
direction of the pitch of the dominant percept. Accordingly pure “holistic listeners” (F0) ex-
clusively perceive the MF, whereas extreme “spectral listeners” (SP) exclusively perceive the
physically present harmonics and are incapable hearing a dominant MF. Depending on the
frequency range “intermediate listeners” might perceive holistic as well as spectral aspects
in a clearly balanced distribution, but in some cases ambiguously resulting in a conflicting
pitch sensation that may lead to inconsistencies. Two types of tone pairs were used (as shown
exemplarily in Figure 7): tone-pairs of type “A” consist of a physically present F0, while in
tone-pairs of type “B” the F0 is missing (= MF). Generally, in both types the 1st complex
tone (α) consists of different, but the 2nd (β) of equal harmonics. Further, in all tone-pairs of
type “B” the octave-shifted residue of the second tone is higher then the F0 of the first one.
In Figure 7 harmonics that are physical present are shown as continuous lines whereas har-
monics that are not part of the stimulus are shown as dotted lines. Components of harmonic
order h = 1 correspond to the related f₀ or MF.

Figure 7. Examplary tone-pairs of Pitch Perception Preference Test, after Schneider et al. (2005a).

All 162 tone-pairs were generated based on generation rules shown in Figure 8. A subgroup
of 54 tone-pairs refers to each condition of 2, 3 or 4 harmonic components (n) respectively.
All components of each of the 54 tone-pairs were defined by 9 different conditions, repre-
senting harmonic ranks between 2 and 16 with 6 different highest component frequencies
(f_{SP_{max}}(1−6) : 294, 523, 932, 1661, 2960, 5274Hz). Please see Appendix 2 for a complete
overview of all test components and their corresponding frequencies.

By logically matching tone-pairs of type A and B it is possible to analyze the proportion
of partially perceived octave-ambiguity (OA) and / or inconsistency (IC). The degree of OA
reflects the amount of perceiving a missing octave-shifted 1st or 2nd harmonic instead of the
related MF, whereas the degree of IC is considered to reflect the amount of inconsistent- or
random choices. Figure 7 shows the principle of this distinction: In both cases the answer
“up”, following complex α to β, can be interpreted as a dominant SP tracking. Accordingly,
the answer “down” can be interpreted as a dominant MF tracking. However, tracking the octave-shifted MF however leads to the answers “down” for the type “A” and “up” for the type “B” complex. Whereas an opposite behavior would lead to an interpretation of IC. In order to quantify the resulting perceptual behavior, an index of sound perception preference ($\delta_P$) was computed individually according to the number of given spectral (SP) and holistic (F0) answers, using the formula:

$$\delta_P = (SP - F0) / (SP + F0)$$

Based on these classifications separate sub-analyses for $\delta_P$, OA [%] and IC [%] were generated separately as a function of the six average spectral frequencies ($f_{SPav}(1−6) : 0.25, 0.45, 0.80, 1.5, 2.6, 5.0kHz$). Thereby the distance between two average frequencies corresponds to a musical interval of a major seventh (7+). In addition, the three parameters were averaged across $f_{SPav}(1−6)$ reflecting a representative cross-section of each parameter ($\delta_{Pav}, OAav$) for reasons of data simplification during most statistical analyses (see chapter 2.3). Detailed values of $f_{SPav}(1−6)$ and also IC values were only shown in chapter 3.2 in addition to $\delta_P$ and OA results, but were not used as raw input parameters for further statistical methods. For the present work, on the $\delta_P$ scale holistic listeners range between −1 to −0.34, spectral listeners between +0.34 to +1, and intermediate listeners between −0.33 to +0.33, following previous definitions by Schneider et al. (2005a). Proportionate amounts of OA perception as well as of potential IC were alike divided into “low” (0 to 25%), “medium” (26 to 50%), “high” (51 to 75%) and “extreme” (75 to 100%).

### 2.2.2 Musical Interval Recognition Test

Relative pitch (RP) perception abilities were determined and quantified using the “Musical Interval Recognition Test” (E. Hofmann, 2012, Music Academy Basel, unpublished). 72 tone-pairs were presented binaurally in a forced-choice task, based on the musically standard scale of chromatic halftone steps (tone duration: 500ms, ramp on-/offset: 20ms, inter-stimulus time: 50ms, interval recognition time: 3.9s). Subjects were instructed to determine the corresponding musical interval of each tone-pair and select the related interval number.

<table>
<thead>
<tr>
<th>type</th>
<th>rule</th>
<th>tone $\alpha$</th>
<th>harmonics in $\alpha$</th>
<th>tone $\beta$</th>
<th>harmonics in $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>“A”</td>
<td>2</td>
<td>$2 \rightarrow n+1$</td>
<td>n</td>
<td>$3 \rightarrow n+2$</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>$3 \rightarrow n+2$</td>
<td>n</td>
<td>$4 \rightarrow n+3$</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>$4 \rightarrow n+3$</td>
<td>n</td>
<td>$6 \rightarrow n+5$</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>$7 \rightarrow n+6$</td>
<td>n</td>
<td>$9 \rightarrow n+8$</td>
<td>n</td>
</tr>
<tr>
<td>“B”</td>
<td>1</td>
<td>$1 \rightarrow n$</td>
<td>n</td>
<td>$2 \rightarrow n+1$</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>$1 \rightarrow n$</td>
<td>n</td>
<td>$3 \rightarrow n+2$</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>$1 \rightarrow n+1$</td>
<td>n+1</td>
<td>$4 \rightarrow n+3$</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>$1 \rightarrow n+2$</td>
<td>n+2</td>
<td>$5 \rightarrow n+4$</td>
<td>n</td>
</tr>
</tbody>
</table>

Figure 8. Tone-pair generation rules for Pitch Perception Preference Test, after Schneider et al. (2005a).
on a chromatic octave scale (Figure 9), while intervals greater than the octave “8” should be projected back by starting from “1” (min. interval: prime “1” = 0 semitones; max. interval: duo-decime “12” = 19 semi-tones). The pause time between the tone-pairs is kept constant and short in order to force the subjects to decide spontaneously. For correct detected tones 1 point was accredited. The random choice score is 12 and the highest achievable score is 72 points. For the present work the scoring was divided into “poor” (0 to 24), “partial” (25 to 48) and “excellent” (49 to 72) RP perception abilities.

![Figure 9. Musical Interval Recognition Test, after E. Hofmann (2012, unpublished).](image)

### 2.2.3 Quantitative Absolute Pitch Test

Absolute pitch (AP) perception abilities were determined and quantified using “Quantitative Absolute Pitch Test” (Wengenroth, 2014) allowing the quantification of partial- and perfect AP perception. The test consists of 34 items with intermittent interference stimuli in a passive tone recognition task (adapted version of the original test, including also 7 active-singing items). The equally tempered test tones comprise 22 instrumental tones in high (\(N = 6\)), middle (\(N = 9\)) and low (\(N = 8\)) frequency range, as well as 5 vocal and 7 sine tones. In order to rule out any memory based interval recognition (related to RP) interference stimuli were inserted, consisting of 5 non-equal tempered sequential instrumental tones followed by 20s of glissando-like continuously distorted music pieces (Figure 10). Only the detection of the respective chroma was tested, independently of the musical octave-position. For correct determined tones 1 point, and for semi-tone errors 0.5 points were accredited. The random choice score is 7 and the highest achievable score is 34 points. For the present work the scoring was divided into “poor” (0 to 12), “partial” (12.5 to 23) and “excellent” (23.5 to 34) AP perception abilities.

![Figure 10. Quantitative Absolute Pitch Test, after Wengenroth et al. (2014).](image)
2.2.4 Frequency Discrimination Threshold Test

Elementary frequency discrimination (FD) was quantified using a customizable audiometric “Alternative Forced Choice” (AFC) test, which is part of the psychophysical-measurement software package for MATLAB provided by Stephan Ewert, University of Oldenburg (Ewert, 2012). During the adaptive, three-interval, 3AFC task subjects were instructed to determine the higher pitch in relation to the baseline-tones (Figure 11). In each phase of the experiment, individual FD thresholds (just noticeable difference) were estimated, while each trial consisted of three intervals indicated by lights. In two of the intervals the frequency of the stimulus was the same. In the third, selected at random, the aim frequency was higher. The subject’s task was to detect the interval containing the higher aim frequency. The baseline frequencies varied in three consecutive steps across the test \( f = 100, 500, 2500 \text{Hz} \); tone duration: 500ms; tone on-/offset: 20ms, inter-stimulus interval: 500ms). For each run the higher tone always started with a tone difference of 80cent above the baseline-tones. With every right answer the higher tone stepwise moved lower in frequency (steps: 6, 3 and 2cent progressively) until the individual FD threshold was reached (minimum measurable difference: 1.17cent). Test tones were presented monaurally for both ears separately. In addition, resulting monaural recorded data were post-hoc averaged between ears (FDav) for reasons of data simplification during most statistical analyses. Due to occasional technical issues with the test computer, FD results could not be recorded for two subjects with the numbers 36 and 52 (amateurs) because of data loss after the test performance. Consequently, these subjects were excluded from all data analyses requiring full datasets including FD threshold results, which affects also PCA and cluster analyses (see chapter 2.3).

![Figure 11. Frequency Discrimination Threshold Test, after S. Ewert (AFC; Ewert, 2012).](image)

2.2.5 Advanced Measure of Music Audiation

The “Advanced Measure of Music Audiation” (AMMA) test developed by E. Gordon (1989) consists of 30 pairs of fictitious short melodies played by piano. Every presented melody is repeated immediately, while the first melody is the reference but the second melody may have one of the following features: tonal change (10/30), rhythm change (10/30) or no change = identical melody (10/30). Subjects were instructed to compare every melody-pair and detect the respective feature in a three-way forced choice task (Figure 12). The “raw tonal test score” was calculated separately by evaluating the number of correct answers minus the number of false answers, plus a standardized baseline value of 20. The random choice score is 20 and the highest achievable score is 40 points. In previous work were AMMA test was
performed, non-musicians scored between 15 and 27, professional musicians between 25 and 40, and amateur musicians in an intermediate range between 17 and 35 (Schneider, 2002; Schneider, 2005a). As no non-musicians were included for the present work, the scoring was slightly adapted based on the inclusion border of 25, according to: “basic” (25 to 30), “high” (31 to 35) and “excellent” (36 to 40) audiation abilities. The additionally assessable “raw rhythm test score” was not evaluated for the present work, as it is not relevant in the context of pitch perception.

![Figure 12. Advanced Measure of Music Audiation, after E.E. Gordon (1989).](image)

### 2.3 Statistics and Data Analyses

All statistical analyses were calculated using SPSS Statistics software (IBM, 2017), performing the following methodical steps for data processing: Descriptive and explorative statistics, including frequency distributions, with classifications, tests for checking normal-distribution of data (Sapiro-Wilk test) and homogeneity of variances (Levene test) were used in order to initially mirror the data. In parallel, all psychoacoustic parameters were normalized using z-transformation for better comparison of the respective values that originate from different test scales and as a required input condition for following factor analyses. The standardized psychoacoustic parameters were then assessed individually as well as group-wise, in order to reflect individual auditory patterns. Due to significant deviation of the parameters from normal distribution, non-parametric tests were used for all further statistical analyses. Non-parametric correlations between parameters were calculated using Spearman’s Rho. Correlations for nominal categorial data (state, school, gender, musical instruments) were not available, as they were not providing ordinal ranks. Non-parametric comparisons between groups (e.g. between subject clusters) were performed using Kruskal-Wallis test and chi-square test. Pairwise comparisons were performed post-hoc using Mann-Whitney test, and corrected for multiple comparison using Bonferroni correction. Principal component analysis (PCA) was calculated based on an integrated correlation matrix in parameter space for data reduction, but most importantly in order to detect potential main components explaining the variance in the data. A rotated variable solution was applied using Varimax-rotation with Kaiser-Normalization, for better interpretability of resulting component loads. Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (reasonable values > 0.5) and Bartlett’s Test of Sphericity (reasonable significance < 0.01) was assessed, in order to evaluate the performance quality of the PCA procedure. Qualification of parameters was verified
by assessing the respective communalities after PCA extraction (reasonable values > 0.5),
with respect to explanation of variance. Input parameters for PCA included normalized
parameters acquired from the psychoacoustic measurements: AMMA tonal score, AP score,
RP score, \( \delta_p \) and OA values, as well as FD thresholds separately for the three FD test
frequencies. All further metadata such as age, gender, state/school, were excluded from PCA in
order to avoid interferences with the psychoacoustic parameters. Relevant main components
were determined based on detection of Eigenvalues greater than 1. Determination of relevant
parameters represented by resulting components were based on the rotated component matrix,
by detecting the highest component loads (= correlation of variables with components)
per parameter and extracted component respectively (reasonable values > 0.5). The factor
strength per component was further assessed for each subject respectively, in order to re-
fect individual auditory patterns based on the components (Figure 13). Hierarchical cluster
analysis was calculated in order to identify similar patterns of normalized psychoacoustic pa-
rameters in-between subject space. The common agglomerative Ward-linkage method was
chosen as appropriate clustering method, as it detects the most appropriate homogeneous
clusters in the data and is not prone to chain-formations. A reasonable amount of clusters
was picked based on detection of the largest initial variance reduction differences and by
matching the agglomeration steps with reasonable cluster patterns visualized in the addi-
tional hierarchical dendrogram. The amount of variance reduction was calculated by deter-
mining the differences between the respective variances of the corresponding agglomeration
coefficients. Based on this, the largest variance-reduction steps were determined to validate
the appropriate selection of clusters. See methods in Pfurtscheller (2017) for reference. The
generated dendrogram was reviewed for overall homogeneity of subject distribution in in-
dividual clusters and results were qualitatively screened to be within expectations (Figure
13).

Figure 13. Statistical processing of psychoacoustic data based on PCA and Hierarchical cluster analysis.
Chapter 3

Results

3.1 Statistical results

3.1.1 Explorative Results

The frequency distributions of age and handedness, primary musical instrument and psychoacoustic parameters were explored across all 93 subjects and the corresponding mean and standard deviations (mean ± SD) were calculated for all parameters respectively. All subjects range between 18 and 29 years of age with a model value of approx. 21 (mean: 21.8 ± 2.6), including 79 right-handers (84%), 6 left-handers (6.5%) and 8 ambidextrous subjects (8.5%). Overall 17 different modern as well as historic musical instruments (including voice) were played actively by the subjects. These can be assigned to at least five different instrument categories, namely: plucked, strings, woodwinds, brass, voice; (Figure 14). The most frequently observed main instruments observed across all subjects were: piano / keys (28.6%), voice (13.2%), violin (11.0%) and guitar / lute (9.9%).

![Figure 14. Distribution of the overall played musical instruments (sorted clockwise) across subjects.](image)

The psychoacoustic test results were likewise initially explored for all 93 subjects with respect to frequency distribution in the respective test scale, which is different for each listening
test (see chapter 2.2 for details). The FD thresholds were explored for only 91 subjects, as the corresponding data from two amateur subjects was not available (see chapter 2.2.4 for details). Overall, the separate results of all eight psychoacoustic parameters delineated in Figure 15 show a high individual variability across the characteristic distribution patterns, as reflected by the respective value ranges and standard deviations. In detail, the respective parameters show:

(a) Mean index of pitch preference ($\delta_{Pav}$, averaged across $f_{SPav(1-6)}$) is distributed asymmetrically in the range between $-1$ and $+0.97$ (mean: $-0.53 \pm 0.49$). A clear majority of 74.2% of subjects are $F0$ listeners with a peak around $-0.9$. Further 8.6% of subjects are spectral- and 17.2% rather intermediate listeners. Mean octave-ambiguity values ($OAav$, averaged across $f_{SPav(1-6)}$) are distributed in the range between 0 and 46.3% (mean: $15.98 \pm 10.85$). A clear majority of 79.6% of subjects show low OAav perception, with a maximum around 8. Further 20.4% of subjects show rather medium OAav perception.

(b) Relative perception (RP) scores are distributed homogeneously in the range between 0 and 69 (mean: $32.39 \pm 18.21$). A majority of 43.0% of subjects show partial RP perception abilities. Further 20.4% are excellent RP possessors, whereas 36.6% are rather poor RP possessors, with a model value around 11.

(c) Absolute perception (AP) scores are distributed asymmetrically in the range between 1 and 32.5 with a model value around 7, which corresponds to the random-choice score (mean: $9.27 \pm 0.65$). A clear majority of 77.4% of subjects show poor AP perception abilities. Further 18.3% are partial-, whereas 4.3% are excellent AP possessors.

(d) Mean frequency discrimination thresholds ($FDav$, averaged across monaural FDs) are distributed in the range from 22.83 to 113.83 cent (mean: $56.65 \pm 20.09$) for 100 Hz, from 2.83 to 40.17 cent (mean: $10.47 \pm 6.25$) for 500 Hz and from 2.17 to 26.50 (mean: $7.35 \pm 3.83$) for 2500 Hz test tones. Furthermore, subjects show decreased FD abilities (up to 1.14 semi-tones) related to increased data variance for the lower frequency bands (100 and 500 Hz). Additional monaural FD results show no significant FD threshold differences between left (FDle) and right ears (FDre), but remarkable inter-individual differences ranging from 1.17 – 21.17 cent at 2500 Hz up to 18.00 – 119.17 cent at 100 Hz. Overall, monaural results are reflected adequately by the above shown “pseudo-binaural” averaged FD values, used for the further performed statistical methods. Moreover, two specific outliers were identified to be siblings (no twins), both playing piano at the same school of Music-Academy (subjects 64 and 63). They show remarkable poor FD abilities, while demonstrating extremely exotic, but similar FD patterns for all three tested frequency bands. However, they appeared in different clusters (see section 3.1.3), due to differences in other parameters.

(e) The tonal raw scores of AMMA test are distributed in the range between 25 (corresponding to the inclusion border) and 40 (corresponding to the maximum score; mean score: 25.
31.72 ± 3.73). A majority of 45.2% of subjects shows high audiation abilities, with a peak around a score of 33. Further 24.7% show excellent, whereas 30.1% show rather basic audiation abilities.

In parallel, all psychoacoustic parameters were normalized on a common scale by z-transformation, as described in chapter 2.3, for better comparison between the test results for the purpose of auditory pattern analyses (chapter 3.2) and as required input condition for subsequent factor analyses (chapter 3.1.3). The resulting standardized z-score values range on a scale from −5 to +5 around a fixed mean of 0 and a fixed variance / standard deviation of 1. Further, all parameters were tested for conformity with normal-distribution, as well as for homogeneity of variances (Table 1). These findings were then used to select the appropriate statistical test for analysis (parametric vs. non-parametric). The results show a significant deviation of all
parameters from normal distribution, as well as inhomogeneous variances for most of the parameters. Consequently, specific non-parametric tests were used for all further correlation- and comparative analyses, as they were expected to fit more appropriately to the data. Bivariate correlations between all parameters were performed using Spearman’s Rho test, showing the following relevant and highly significant (** \( p < 0.01 \)) correlations for:

- \( \delta_{Pav} \) and \( OA_{av} \): \( r = 0.64^{**} \)
- \( RP \) and \( AP \): \( r = 0.46^{**} \)
- \( RP \) and \( AMMA \): \( r = 0.43^{**} \)

<table>
<thead>
<tr>
<th>Test of Normal-Distribution Conformity</th>
<th>Test of Homogeneity of Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>Shapiro-Wilk Statistic</td>
</tr>
<tr>
<td>age</td>
<td>0.903</td>
</tr>
<tr>
<td>gender</td>
<td>0.636</td>
</tr>
<tr>
<td>handedness</td>
<td>0.438</td>
</tr>
<tr>
<td>state</td>
<td>0.634</td>
</tr>
<tr>
<td>school</td>
<td>0.749</td>
</tr>
<tr>
<td>AMMA</td>
<td>0.969</td>
</tr>
<tr>
<td>AP</td>
<td>0.776</td>
</tr>
<tr>
<td>RP</td>
<td>0.964</td>
</tr>
<tr>
<td>( \delta_{P_{av}} )</td>
<td>0.820</td>
</tr>
<tr>
<td>( OA_{av} )</td>
<td>0.935</td>
</tr>
<tr>
<td>( FD_{av100} )</td>
<td>0.959</td>
</tr>
<tr>
<td>( FD_{av500} )</td>
<td>0.768</td>
</tr>
<tr>
<td>( FD_{av2500} )</td>
<td>0.825</td>
</tr>
</tbody>
</table>

Table 1. Statistical tests for normal-distribution and homogeneity of variances.

Taken together, the explorative results show the preferred mode of pitch perception, represented by \( \delta_{P} \), to correlate with the amount of OA perception, indicating that higher degrees of octave-shifted perception might be more frequent in spectral listeners. Furthermore, medium positive correlations appear between RP and AP perception abilities, as well as audiation abilities, indicating certain relations between the respective scores. No further relevant correlations were found for other parameters such as FD thresholds. Moreover, correlations for gender, state, school and musical instruments were not available (see chapter 2.3 for details).

### 3.1.2 Principal Component Analysis

Principal Component Analysis (PCA) was applied (see chapter 2.3 for details) on the psychoacoustic results in an attempt to determine the potential main components explaining the data variance. Only psychoacoustic parameters were included as input variables (normalized) for PCA processing, irrespective of further meta-data such as state, school, age, gender, handedness or musical instrument. First, the resulting performance quality of the PCA procedure was verified by the following criteria:

- KMO Measure of Sampling Adequacy: 0.549
• Bartlett’s Test of Sphericity: 91.7 (approx. chi-square); $p < 0.001$ (significance level)

• Communalities for each parameter after extraction: > 0.5 (see details in Table 2)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial</th>
<th>Extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMMA</td>
<td>1.000</td>
<td>.583</td>
</tr>
<tr>
<td>AP</td>
<td>1.000</td>
<td>.688</td>
</tr>
<tr>
<td>RP</td>
<td>1.000</td>
<td>.775</td>
</tr>
<tr>
<td>$\delta_{Pav}$</td>
<td>1.000</td>
<td>.676</td>
</tr>
<tr>
<td>OAv</td>
<td>1.000</td>
<td>.716</td>
</tr>
<tr>
<td>FDa100</td>
<td>1.000</td>
<td>.743</td>
</tr>
<tr>
<td>FDa500</td>
<td>1.000</td>
<td>.703</td>
</tr>
<tr>
<td>FDa2500</td>
<td>1.000</td>
<td>.805</td>
</tr>
</tbody>
</table>

Table 2. Communalities of psychoacoustic data based on the PCA results.

The PCA processing identified eight components after rotation. Only four of these, those with Eigenvalues greater than 1 were considered main components (Table 3). The four selected components are marked in the scree-plot (Figure 16). Inspecting the “rotation sums of squared loadings” shows, that in total 71.1% of data variance are explained sufficiently by the selected four components after extraction (Table 3). In particular, one quarter of the data variance can be attributed solely to component-1.

<table>
<thead>
<tr>
<th>Total Variance Explained</th>
<th>Initial Eigenvalues</th>
<th>Rotation Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>2.053</td>
<td>25.665</td>
</tr>
<tr>
<td>2</td>
<td>1.372</td>
<td>17.156</td>
</tr>
<tr>
<td>3</td>
<td>1.215</td>
<td>15.191</td>
</tr>
<tr>
<td>4</td>
<td>1.049</td>
<td>13.111</td>
</tr>
<tr>
<td>5</td>
<td>.804</td>
<td>10.047</td>
</tr>
<tr>
<td>6</td>
<td>.599</td>
<td>7.489</td>
</tr>
<tr>
<td>7</td>
<td>.528</td>
<td>6.601</td>
</tr>
<tr>
<td>8</td>
<td>.379</td>
<td>4.738</td>
</tr>
</tbody>
</table>

Table 3. Eigenvalues and explained variances shown for each main component after PCA extraction.

The represented psychoacoustic parameters show a clear orthogonal distinction, as illustrated for the first three component dimensions in the rotated component space (Figure 17). After assessing the respective component loads across all parameters for each extracted component (Table 4), the parameters with the highest loads were determined. The resulting components show the following specific representations of parameters:

• Component-1: AMMA / AP / RP

• Component-2: $\delta_{Pav}$ / OAv

• Component-3: FDa100 / FDa500

• Component-4: FDa2500
### Rotated Component Matrix

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMMA</td>
<td>0.659</td>
<td>0.050</td>
<td>-0.171</td>
<td>0.344</td>
</tr>
<tr>
<td>AP</td>
<td>0.733</td>
<td>-0.032</td>
<td>0.033</td>
<td>-0.385</td>
</tr>
<tr>
<td>RP</td>
<td>0.864</td>
<td>0.154</td>
<td>-0.024</td>
<td>-0.070</td>
</tr>
<tr>
<td>$\delta_{av}$</td>
<td>0.223</td>
<td>0.784</td>
<td>-0.064</td>
<td>0.087</td>
</tr>
<tr>
<td>OAav</td>
<td>-0.058</td>
<td>0.833</td>
<td>0.020</td>
<td>-0.134</td>
</tr>
<tr>
<td>FDav100</td>
<td>-0.157</td>
<td>0.113</td>
<td>0.833</td>
<td>-0.109</td>
</tr>
<tr>
<td>FDav500</td>
<td>0.083</td>
<td>-0.234</td>
<td>0.712</td>
<td>0.366</td>
</tr>
<tr>
<td>FDav2500</td>
<td>-0.086</td>
<td>-0.040</td>
<td>0.096</td>
<td>0.887</td>
</tr>
</tbody>
</table>

Rotation converged in 5 iterations.

Table 4. Rotated component matrix after rotation, showing the component loads for each parameter.

![Scree-plot](image)

Figure 16. Scree-plot, showing the Eigenvalues of all extracted- and selected PCA components (color-coded).

Consequently, component-1 shows a combined representation of AMMA, AP, and RP scores, indicating a close interplay of musical key abilities (hereafter named “tonal musicality”). Component-2 represents $\delta_P$ in conjunction with the degree of OA (hereafter named “pitch timbre preference”). Component-3 represents the FD thresholds at low frequencies based on the lower band 100 and 500Hz test frequencies (hereafter named “low-band sensitivity”), whereas component-4 independently forms a distinct factor representing FD thresholds at 2500Hz (hereafter named “high-band sensitivity”). Furthermore, the respective component loads were then determined for each subject individually as delineated in Figure 18 on a normalized scale, showing a high overall variability across subjects for all components. Certain outliers were detected by inspecting additional boxplots for each component (Figure 19), showing relatively high (dots) or extreme (stars) values in the component space: Case 26 shows an outstanding high musicality value in the “tonal musicality” component, whereas in the FD based components some cases show a relatively poor sensitivity for lower (cases 22, 31) or higher (cases 64, 63; sisters) tones. However, no outliers appeared for the “pitch focus” component.
In summary, calculating PCA as a combined result across the eight raw psychoacoustic test parameters was verified to sufficiently well characterise the dataset. This resulted an extraction of four main components which show to cumulatively explain over 70% of the data’s variance. Further, the extracted components represent inherent associations between parameters which are clearly separated in the rotated component space and therefore can be reasonably assigned to distinct perceptional factors, as further discussed in chapter 4.

Figure 17. Corresponding psychoacoustic parameters in component space after matrix rotation.

Figure 18. Individual distributions of component loads show high inter-individual variability.
3.1.3 Hierarchical Cluster Analysis

Hierarchical cluster analysis was applied to identify similar perceptual patterns across subject space, based solely on the psychoacoustic raw parameters only (likewise PCA input). First, the output from the Ward-linkage clustering procedure was inspected based on the resulting agglomeration-schedule, which summarizes the coefficients of the consecutive clustering steps (Figure 20). In order to measure the quality of the clustering procedure, the consecutive reduction of variance was calculated for each cluster step, based on the corresponding coefficients per agglomeration stage (Table 5, here only shown for the highest 5 cluster steps). The resulting cumulative variance reduction was determined to be 16.87% for a 2-cluster, and 27.1% for a 3-cluster solution, in reference to an initial no-cluster solution. These were found to be the largest variance-reduction steps, in comparison to further smaller reduction steps to cluster 4 and following clusters (see 2.3 for details).
Further, a 3-cluster solution was found to likewise match appropriately with the visualization of the 3-cluster level in the additional dendrogram (Figure 21), when graphically detecting three homogeneous clusters of equal size in subject space (yellow lines). Consequently, based on a final subjective decision, the 3-cluster solution was selected. The selected 3-cluster solution provides sufficient group sizes for following inter-group comparisons and in order to explore general perceptual patterns. The resulting cluster membership of each subject to its corresponding cluster was evaluated and color coded (cluster-1: red, cluster-2: green, cluster-3: blue) in the dendrogram and additional data table (Appendix 1).

The resulting cluster-specific values are listed in Table 6 with respect to: N of subjects, mean, SD, SEM, confidence intervals and minimum / maximum values. The resulting group sizes per cluster are:

(i) Cluster-1: N = 35
(ii) Cluster-2: N = 29
(iii) Cluster-3: N = 27

Hence, the results show, the group sizes to be sufficiently balanced and the distribution of subjects to not be significantly different between clusters (chi-square test: \( p > 0.56 \)). Moreover, results of state, school and musical instruments (Figure 22) show the following proportions of frequencies for each cluster:

(i) Cluster-1: Mainly professionals (approx. 91%) including the majority of all three musical disciplines (Classical 46%, Jazz 26%, Early Music 20%), playing rather strings...
(total 31%), woodwinds (total 9%), historic instruments such as cembalo (6%) and including most of the singers (20%).

(ii) Cluster-2: Mixed group of likewise amateurs (approx. 59%) and professionals (approx. 41%), including mainly classical musicians (31%), playing mainly piano (66%) and strings (total 17%).

(iii) Cluster-3: Mainly amateurs (approx. 81%), playing mainly brass (total 15%), guitars (15%) and woodwinds (total 11%).

Furthermore, potential differences between the clusters were tested by comparing means using Kruskal-Wallis and Mann-Whitney tests (see chapter 2.3 for details). The comparative results listed in Table 7 show the clusters to generally differ:

- Highly significant at the level of \( **p < 0.0003 \) for state, school, AMMA, RP, AP, \( \delta_{Pav} \), OAav, FDav100, tonal musicality and pitch timbre preference.

- Significant at the level of \( **p < 0.003 \) for FDav500 and low-band sensitivity.

All further parameters show no general significant differences between clusters, such as: gender, handedness, musical instrument, FDav2500 and high-band sensitivity. Moreover, pairwise comparisons between clusters show the following differences:

(i) Cluster-1 vs. -2:

- Highly significant at the level of \( **p < 0.0003 \) for state, RP, \( \delta_{Pav} \), OAav, FDav100, tonal musicality and pitch timbre preference.

- Significant at the level of \( **p < 0.003 \) for school, AP, FDav500 and low-band sensitivity.
Table 6. Cluster-specific descriptive results calculated for all parameters and components.
(ii) Cluster-1 vs. -3:

- Highly significant at the level of $***p < 0.0003$ for state, school, AMMA, RP, AP, tonal musicality and pitch timbre preference.
- Significant at the level of $**p < 0.003$ for $\delta_{\text{Pav}}$.

(iii) Cluster-2 vs. -3:

- Highly significant at the level of $***p < 0.0003$ for AMMA, $\delta_{\text{Pav}}$ and tonal musicality.
- Significant at the level of $**p < 0.003$ for RP, OAav, FDav100 and pitch timbre preference.

Consequently, evaluating the results of all psychoacoustic parameters and component loads (Figure 23) between clusters, shows clearly distinctive multi-parametric patterns for each cluster (error-bars: SD).
Evaluating the composed multi-parametric profiles shows the following cluster-specific features:

(i) Cluster-1: Mainly professionals (approx. 91%) including the majority of all three musical disciplines (Classical 46%, Jazz 26%, Early Music 20%), playing mainly Strings (total 31%), Woodwinds (total 9%), historic instruments such as cembalo (6%) and including most of the singers (20%). By comparison, this cluster shows the overall highest values across all psychoacoustic parameters and the corresponding components. Both, AP and RP abilities are remarkably high. $\delta_P$ values tend towards $+1$, indicating mainly SP listeners, in accordance with high OA values suggesting a focus on timbre. FD values are clearly over average indicating remarkably increased FD abilities or high sensitivity in low and high frequency bands.

(ii) Cluster-2: Mixed group of likewise amateurs (approx. 59%) and professionals (approx. 41%), including mainly classical musicians (31%), playing mainly piano (66%) and strings (total 17%). Only AMMA values are relatively high above average in this cluster similar to cluster-1. Both, AP and RP abilities are rather spreaded around 0. Thus, the tonal musicality component is in total balanced out at approx. 0 in normalized component space. However, all further psychoacoustic parameters and components show overall relatively low values. $\delta_P$ and OA show the lowest values by comparison, indicating primarily F0 listeners without a specific focus on timbre. Mean FD values are significantly under average indicating significantly lower FD abilities, according to a lower sensitivity in low and high frequency bands.

(iii) Cluster-3: Mainly amateurs (approx. 81%), playing primarily brass (total 15%), guitars (15%) and woodwinds (total 11%). By comparison, AMMA, AP and RP show relatively low values, reflected by a clearly under average tonal musicality component. $\delta_P$ values spread around 0, indicating mainly intermediate listeners. However, OA is remarkably high in this cluster, indicating a focus on timbre, followed by above average
FDav100 and FDav500 thresholds, in accordance with the corresponding components. Overall, this cluster shows the most ambivalent contrasts between parameters and components.

Taken together, three separate clusters of subjects were found solely based on specific features of their psychoacoustic test results. In combination with the post-hoc assessed proportions of respective school affinities and musical instrument preferences, the resulting multi-parametric features lead to distinctive characteristics attributable to each cluster. The comparative group results show clearly, that most psychoacoustic parameters and the corresponding main components were found to be represented significantly different between the clusters. Consequently, these results provide the basis for investigating cumulative perceptual and behavioral patterns, as elucidated in the following chapter.

3.2 Auditory Fingerprints

3.2.1 Characteristic Group Profiles

The Individual multi-parametric data was collectively investigated for each selected cluster, in order to combine the perceptual and behavioral results. In addition, more detailed data was assessed specifically for $\delta_P$ in dependence with OA and IC as well as monaural data of the separate FD thresholds. The multi-dimensional patterns were then evaluated specifically for each cluster, also based on visualization. The cluster-specific group results are cumulated in Figure 24, summarizing the following features (from left to right panels):

(a) Dendrogram from the hierarchical cluster analysis, color coded based on the selected three subject clusters.

(b) Proportions [%] of school (reflecting also state) and played musical instruments (reduced to the instrument categories: plucked, strings, woodwinds, brass, voice).

(c) Main components from PCA as well as the underlying psychoacoustic raw parameters, delineated separately by boxplots in normalized space (demonstrating mean values, percentiles as well as certain outliers).

(d) Frequency specific sub-analyses of Pitch Perception Preference test (see chapter 2.2.1), showing $\delta_P^{av}$, OAav and ICav, as a function of the six defined average spectral frequencies $f_{SPav}(1−6)$ [kHz]. Related proportions of OAav and ICav [%] are assigned to and marked in the $\delta_P^{av}$ subplot, if certain threshold borders (dotted lines) are exceeded (OAav: $^0<10,^*25,^{**}35,^{***}45\%$; ICav: $^110,^{**}20,^{***}30\%$).

(e) Monaural FDav thresholds for left- and right ears and each test frequency separately.
Figure 24. "Auditory fingerprints" averaged for each cluster, showing the combined multi-parametric results of the Pitch Perception Preference Test.
Bringing together the multi-parametric features of each cluster, basically described in chapter 3.1.3 and listed in Table 6, with the frequency depending details of pitch perception preference results depicted in the panels d) and e), leads to prominent characteristic profiles that are overall attributable to the corresponding individuals as follows:

(i) Subjects in cluster-1 show predominantly: Professional musical state with increased preference for mostly strings, exotic/renaissance instruments and singing (b). Strong tonal musicality resulting of high audiation related to high absolute- and relative pitch identification abilities (c). Rather analytical pattern recognition of complex tones focused on single sound details (c, d). Differentiated perception of harmonics (more spectral shift of \( \delta_P \)) and preference for sound spectra with prominent timbre (increased OAs, mainly at higher frequency bands) with slight ICs at low frequency bands (d). Extreme sensitivity to pitch differences and highly selective frequency detection abilities, particularly for lower tones (e).

(ii) Subjects in cluster-2 show predominantly: Mixed musical state with increased preference for mainly plucked instruments such as piano (b). High tonal musicality resulting of likewise strong audiation related to proficient absolute- and relative pitch identification abilities (c). Rather holistic pattern recognition of complex tones (c, d). Highly “completing” sound perception focused on pitch sensations of the \( F_0 \) (solid holistic shift of \( \delta_P \) across all frequency bands) and preference for rather clean sound spectra with unobtrusive timbre (decreased OAs, except for the top frequency) with likewise ICs (over 20% at 0.25kHz) predominant at the lowest frequency bands (d). Less sensitive to pitch differences and comparatively indefinite frequency detection abilities, particularly for lower tones (e).

(iii) Subjects in cluster-3 show predominantly: Amateur musical state with increased preference for mainly woodwind and brass-, but also plucked instruments such as guitars (b). Lower tonal musicality resulting of rather under-average audiation-, and uncertain absolute- and relative pitch identification abilities (c). Rather ambiguous multiple-pitch sensations (c, d). Rather “balanced” perception of harmonics (slight spectral shift of \( \delta_P \), mostly at the outer frequency bands) and preference for sound spectra with prominent timbre (increased OAs, mainly at higher frequency bands) with likewise fewest ICs at low frequency bands (d). Highly sensitive to pitch differences and proficient selective frequency detection abilities, most likely for lower tones (e).

Moreover, certain outliers were detected by inspecting additional boxplots for both, components and psychoacoustic parameters (panels c, d), identifying extreme values in both directions of normalized space. However, the overall most extreme outliers appeared for cluster-2 (cases 64, 31, 22). Taken together, the described multi-parametric conjunction of perceptual and behavioral results demonstrate clearly the existence of distinct profiles, that may lead to the suggestion of characteristic “auditory fingerprints”.
3.2.2 Exemplary Individual Cases

Individual results were further evaluated for the selection of three exemplary cases, representing the specific multi-parametric patterns of each cluster. The corresponding individual profiles are delineated in Figure 25, basically referencing to the summarized results described for the cluster’s average in the last section. The threshold borders for Pitch Perception Preference details shown in panel (d) were adapted to the individual cases, for reasons of more extreme values: OA: \( O < 10,^{*}25,^{**}50,^{***}75\%; \) IC: \( ^{*}10,^{**}20,^{***}30\% \).

The three selected individuals show the following profiled features (described by raw parameters in addition to z-scores in the figure):

(i) Case 6: Professional male musician assigned to cluster-1 (a), with affinity to Classical school of Music-Academy, predominantly playing cello (b). Strong tonal musicality (2.25) based on high audiation (raw AMMA score = 31.0), related to excellent absolute- (raw AP score = 30.5) and relative (raw RP score = 58) pitch identification abilities (c). Extreme analytical pattern recognition (pitch timbre preference: \( \delta_{PAV} = 0.64, OA_{AV} = 11.1\% \)) of complex tones focused on single sound details (c). Clearly differentiated perception of harmonics, with extreme spectral shift of \( \delta_{P} \) towards \( +1 \) predominantly in center frequency bands, preference for sound spectra with prominent timbre, with increased OAs mainly at higher frequency bands (\( *2.6kHz \) and \( ***5.0kHz \)), contrasted by a clear non octave-shifted focus at low frequency bands (predominantly \( 0.8kHz \)) and slight ICs (ICav = 4.2\%) at low frequency bands (d). Extreme sensitivity to pitch differences (low-band sensitivity: \( -0.12, \) high-band sensitivity: \( 0.83 \)) with extremely selective FD abilities (FDav100 = 40.67\cent, FDav500 = 12.50\cent, FDav2500 = 5.67\cent)(c), particularly for lower tones of the left ear (FDle100 = 58.50\cent, FDle500 = 11.83\cent, FDle2500 = 5.67\cent), but also right ear (FDre100 = 22.83\cent, FDre500 = 13.17\cent, FDre2500 = 5.67\cent)(e).

(ii) Case 51: Amateur female musician assigned to cluster-2 (a), with affinity to medical school, predominantly playing piano (b). Average tonal musicality (−0.61) based on high audiation (raw AMMA score = 32.0), related to poor absolute- (raw AP score = 4.0) and relative (raw RP score = 17) pitch identification abilities (c). Extreme holistic pattern recognition (pitch timbre preference: \( -1.11, OA_{AV} = 7.9\% \)) of complex tones (c). Highly “completing” sound perception focused on pitch sensations of the \( F0 \) (solid holistic shift of \( \delta_{P} \) towards \( -1 \) across most frequency bands) and preference for rather clean sound spectra with unobtrusive timbre (overall no OAs, except for the top frequency \( 5.0kHz \)), but mostly clear non octave-shifted perception (predominantly \( 0.25kHz, 0.45kHz, 0.8kHz, 2.6kHz \)) with moderate ICs (ICav = 6.9\%) mainly at the lowest frequency band (\( *0.25kHz \))(d). Rather insensitive to pitch differences (low-band sensitivity: \( -0.82, \) high-band sensitivity: \( -0.26 \)) with less selective FD abilities (FDav100 = 79.17\cent, FDav500 = 13.83\cent, FDav2500 = 7.83\cent)(c), particularly for lower tones of the left ear (FDle100 =
Figure 25. Individual "auditory fingerprints", showing the combined multi-parametric results off all acquired data for three exemplary cases, including additional frequency-specific details of the Pitch Perception Preference Test.
83.83\text{cent}, \text{FDle}500 = 13, 17\text{cent}, \text{FDle}2500 = 3, 17\text{cent}), and also right ear (\text{FDre}100 = 74.50\text{cent}, \text{FDre}500 = 14.50\text{cent}, \text{FDre}2500 = 12.50\text{cent})(e).

(iii) Case 95: Amateur male musician assigned to cluster-3 (a), with affinity to medical school, predominantly playing acoustic-guitar (b). Under average tonal musicality (−1.48) based on basic audiation (raw AMMA score = 26.0), related to poor absolute- (raw AP score = 5.5) and relative (raw RP score = 12) pitch identification abilities (c). Rather ambiguous multiple-pitch sensations (pitch timbre preference: 1.51, $\delta_{pav} = 0.20$, OAav = 30.6%) with “balanced” perception of harmonics (spectral shift of $\delta_{p}$ towards +1 in center frequency bands, but more balanced around 0) and preference for sound spectra with prominent timbre (increased OAs, mainly at higher frequency bands: $^*0.25kHz$, $^*0.45kHz$, $^*0.8kHz$, $^*2.6kHz$ and $^{**}5.0kHz$) and increased ICs (ICav = 14.6%) at all frequency bands ($^*0.25kHz$, $^*0.45kHz$, $^{**}1.5kHz$, $^*2.6kHz$)(d). Proficient sensitivity to pitch differences (low-band sensitivity: −0.01, high-band sensitivity: 0.95) with average to high FD abilities (FDav100 = 64.83\text{cent}, FDav500 = 6.17\text{cent}, FDav2500 = 4.17\text{cent})(c), particularly for higher tones of the left ear (\text{FDle}100 = 66.50\text{cent}, \text{FDle}500 = 6.50\text{cent}, \text{FDle}2500 = 5.83\text{cent}), but also for the right ear (\text{FDre}100 = 63.17\text{cent}, \text{FDre}500 = 5.83\text{cent}, \text{FDre}2500 = 2.50\text{cent})(e).

Taken together, the three different exemplary cases show individual pronounced conjunctions of perceptual and behavioral results, which are representative for their corresponding cluster. Inspecting individual characteristic profiles may lead to the suggestion of typical individual “auditory fingerprints”, as further discussed in chapter 4.
Chapter 4

Discussion

In summary, the results show remarkable inter-individual differences in musicians on both levels, pitch perception and musical abilities. Furthermore, multi-parametric patterns of sound perception were reflected by characteristic “auditory fingerprints” at the individual and group level: At the individual level, the results show a high inter-individual variability across the eight psychoacoustic parameters, reflecting clear individual differences in pitch perception and related musical abilities. In addition, principal component analysis (PCA) revealed four different main components to sufficiently represent coherent aspects of the psychoacoustic data. Specifically, these components are: tonal musicality, pitch timbre preference, low-band sensitivity and high-band sensitivity. On the group level, multi-parametric cluster analyses revealed three sub-groups of subjects, showing significantly different results with respect to the underlying perceptional patterns. The described multi-parametric conjunction of perceptual and behavioral results demonstrate clearly the existence of distinct profiles, that may lead to the suggestion of characteristic “auditory fingerprints”. Consequently, at least three different modes of pitch and sound perception are suggested, characterized by:

(i) Pronounced analytic pattern recognition, focused on spectrum / timbre and highly sensitive to single tones.

(ii) Pronounced holistic pattern recognition, focused on (missing) fundamental pitch and rather less sensitive to single tones.

(iii) Less pronounced audiation and pitch detection abilities, linked to ambiguous multi-pitch sensations (“balanced mode”), with preference for timbre.

The results of pitch perception preferences ($\delta_p$) clearly reflect the high variability between subjects, confirming the early ambiguous-pitch observations of Smoorenburg (1970) and thus indicating a clear dichotomic distribution separating $F_0$ pitch listeners (holistic mode) from $SP$ pitch listeners (spectral mode). In the present sample, a clear majority of approx. 74% of subjects corresponded to a rather holistic mode, while a minority of only approx. 9% of subjects belonged instead to a rather spectral mode. The remaining approx. 17% were part of a third, intermediate mode. These results are basically in line with similar observations in musicians and nonmusicians reported by Schneider et al. (2005a, 2005b), Seither-Price...
et al. (2007) and Ladd et al. (2013). However, the averaged index of pitch preference $\delta_{pav}$ was clearly asymmetric shifted towards the holistic perception ($\delta_{pav} = -0.53$), which is in contrast to the reported balanced bimodal distribution of $\delta_p$ shown by Schneider et al. (2005a). As discussed previously by Seither-Preisler et al. and Ladd et al., this divergence is assumed to be explained by the applied correction for octave-shifted $F0$ percepts applied by Schneider et al., based on the assumption that they are not part of $F0$ perception. However, this correction has been taken into account in the present work. Therefore, the observed predominance of holistic perception might rather be related to the subject characteristics of this sample. Earlier, other samples have been observed that demonstrated inversely a strong dominance of spectral pitch perception, e.g. the musicians of the Royal Liverpool Philharmonic Orchestra (RLPO; Schneider, 2005b). In the present work musicians perceived in average about 20% octave-ambiguities (OAs) for different frequency bands when listening to complex tones. These OA sensations might be related to a specific focus on timbre or tone-color of complex sounds, as reported rarely in early psychoacoustic studies by Terhardt (1972) and Patterson (1973). Thus, the observed octave-ambiguity phenomenon points to the existence of a third, timbre-related dimension of pitch perception. Further, it can be assumed that the relative frequency of OAs are related to the respective mode of pitch perception, as the highest values of OAs of about 45% were found for rather SP listeners (mainly assigned to cluster-1 and -3 shown in Figure 26). Accordingly, a correlation of $r = 0.64**$ between $\delta_p$ and OA was found. The interaction between $\delta_p$ and OAs was further found to be expressed by a common component, interpreted as pitch timbre preference. Interestingly, intermediate pitch listeners, who are mainly represented by a separate sub-group mainly appearing in cluster-3, demonstrate an independent perceptual pattern, not solely explainable by neither one of the two basic modes (SP or $F0$), nor by noticeable inconsistencies (ICs) in their answer behavior. However, the interpretation of that third, intermediate pitch perception mode is quite challenging and limited for the present work, also due to the observed bias towards holistic pitch perception. A larger pool of subjects and selective listening-test adaptions would be necessary to further investigate the perceptual continuum in-between, or orthogonal to the two basic modes, which is in accordance to the suggestions of Ladd et al. (2013).

The results of relative pitch (RP) and absolute pitch (AP) perception abilities likewise underline the high variability between subjects. The results on RP perception show a broad-scaled distribution of RP abilities across the totality of musicians, revealing about 20% excellent, 43.0% partial and 37% rather poor RP possessors. In parallel, cluster analyses revealed RP abilities to be the strongest discriminant between all selected sub-groups of musicians, showing excellent RP possessors to appear mainly in cluster-1. These findings basically confirm recent observations on individual RP differences by McDermott et al. (2010), who also noted also underlying individual preferences for specific acoustical properties. These findings clearly demonstrate, that the ability to identify musical intervals differs strongly even between individual musicians, suggesting RP perception not to be the everybody’s “default-mode”. Furthermore, RP scores showed to be positively correlated with AP scores.
Figure 26. Three different modes of pitch perception preference, distinguished by the three specific subject clusters (colour-coded).

\( r = 0.46^{**} \), underlining the recently growing evidence for overlaps between AP and RP networks in regard to brain function and anatomy (Zatorre, 1998), electrophysiological correlates (Itoh, 2005), pitch matching of chords (McLachlan, 2013), global vs. local processing (Ziv, 2014), and perceptual performance (Dooley, 2011). Moreover, RP scores showed to be positively correlated also with audiation abilities \( r = 0.43^{**} \), measured by the raw tonal AMMA score, underpinning reports of overlaps between RP perception and pitch memory (Schulze, 2009) or music recognition (Creel, 2012). In the present work, about 18% of musicians showed partial and at least about 4% excellent AP abilities. The remaining majority of about 77% of musicians demonstrated mainly poor AP perception abilities, ranging around the random-choice level. Furthermore, AP scores appeared to be highest in cluster-1, which was found to include most of the professional musicians. These findings are in line with previous reports, estimating AP perception to be present for about 0.01% individuals across the general population, but for about 7 – 32% across professional musicians (Baharloo, 1998; Gregersen, 1999). However, while AP abilities are more and more in the scope of recent research projects, RP perception abilities are rarely investigated. Ultimately both, the perceptual and the neural basis of RP as well as AP processing are largely unexplored with respect to their multidimensional character, potential overlaps, individual variability and underlying mechanisms.

The results of frequency discrimination (FD) thresholds as well emphasize the high variability between musicians. Overall, the measured individual monaural FD thresholds showed remarkable inter-individual differences ranging from 1.17 at high-, up to 119.17 cent at low tested frequencies, according to differences of a factor of 100 or more. Furthermore, different patterns in FD were found to be expressed by separate main components, interpreted as low-band sensitivity (for 100 and 500 Hz bands) and high-band sensitivity (for the 2500 Hz band). However, musicians with excellent FD thresholds, according to the smallest just noticeable
differences (jnds) between the tested frequencies, were found to appear in cluster-1 and -3, particularly showing better sensitivity to lower frequencies. These findings are consistent with reported huge inter-individual differences in FD by a factor of 100 up to 1000, ranging from 1cent in musicians up to 300cent (= 3 semitones) in some nonmusicians (Serrallach, 2016). Interestingly, two specific outliers were identified to be siblings (no twins), both playing piano at the same Music class (subjects 64 and 63). They showed both, remarkable poor FD abilities, while demonstrating extremely exotic, but similar FD patterns for all three tested frequency bands. Such cases are very rare but of high interest, and may contribute to disentangle the underlying factors, such as innate predispositions, early maturational factors, or synchronized environmental factors of musical training or education.

The results of audiation, based on investigating the raw tonal AMMA score, confirm the high variability between musicians. Overall, audiation abilities appear to differ remarkably in-between musicians, showing a broad distribution of excellent (25%), high (45%) and rather basic (30%) tonal AMMA scores. These findings are consistent with previous reports of high inter-individual variability in audiation abilities by (Schneider, 2002; Schneider, 2005a). Furthermore, the observed relations between RP, AP and audiation abilities, expressed by a common component interpreted as tonal musicality, were overall found to be strongest in professional musicians represented mainly by cluster-1. This is reasonable with regard to the broad evidence showing distinct audiation abilities to be interpretable as a core element of musicality, as reflected by Gordon’s “music learning theory” (Gordon, 2012). Moreover, audiation scores were widely accepted as a model for music aptitude (Shuter-Dyson, 1999), suggesting advanced audiation abilities to be a innate key factor for excellent musical practice. Furthermore, the overall increased component values shown for cluster-1 subjects, suggest auditory imagery processes to be related to several pitch perception and discrimination abilities. Overall, it can be concluded that auditory imagery processes can preserve spectral and temporal properties of auditory stimuli and enhance auditory (pitch) discrimination while interfering with auditory detection, by involving brain areas used during auditory perception (Hubbard, 2010). Consequently, a important role of auditory imagery abilities in music performance can be assumed, while individual differences in audiation may be a source of variation in expressive performance excellence (Keller, 2012).

The variability of preferred musical instruments is likewise very high, in total 17 different instruments. The diversity of played musical instruments clearly reflects the high variability of musical styles and preferences, also reflecting the three different musical disciplines (schools) provided by the Music-Academy, Basel. Interestingly, musical instruments as well as the affinity to the different schools appeared to distribute clearly different between the three selected clusters of subjects. As the applied cluster analyses were solely based on psychoacoustic parameters from the respective listening tests, the results suggest strong relations between sound perception and musical preferences.

Finally, it should be mentioned, that the neural basis of auditory processing, and particularly pitch coding, is currently in the scope of various studies. There is more and more evidence that musicians differ from nonmusicians with respect to their individual auditory and cor-
responding neurological properties (Pantev, 2001; Tervaniemi, 2005; Seither-Preisler, 2007; Boh, 2011; Wengenroth, 2014; Di Nuovo, 2016). Moreover, recent studies provide some support for nurture effects, showing a role for musical experience, but also indicate high individual variability for preferring acoustic properties (McDermott, 2010). Despite the fact that the source for these individual differences in sound perception is still not disclosed, several experimental findings in the neuroscience field point towards a corresponding underlying neuroanatomical (Rousseau, 1996; Schneider, 2005a; Wengenroth, 2014), neurophysiological (Winkler, 1997; Johnsrude, 2000; Patel, 2001; Schneider, 2002; Coffey, 2016), training dependent (Schulte, 2002; Bengtsson, 2005; Klein, 2016), as well as genetical and environmental factors (Drayna, 2001; Mosing, 2013; Mosing, 2014; Butkovic, 2015; Oikkonen, 2016). However, the exact perceptual mechanisms of pitch and sound perception are of high complexity and therefore part of an ongoing debate in the field (Zatorre, 2001; Patterson, 2002; Schneider, 2009a; Bizley, 2010; Moore, 2014; Plack, 2014).
Musicians have been variously reported to show remarkable inter-individual differences in elementary hearing functions, sound perception mode, musical instrument preference, performance style, as well as musical related abilities such as absolute- and relative pitch perception and auditory imagery (audiation). However, relevant literature in the field regarding perceptual and psychophysical aspects of sound and particularly pitch perception is highly contradictory, and subjective differences are, for the most part, not considered. Moreover, it leaves largely unexplored the manner in which individual differences in musical pitch perception are related to further musical abilities and behavior. Taken together, the findings of the present work suggest that individual “auditory fingerprints” extracted solely from psychoacoustic hearing tests, suggest partially inherent modes of perception. The corresponding inter-individual differences could be classified into a limited set of common patterns of perceiving pitch and sound. It can be assumed, that specific auditory characteristics are related individual preferences for musical instruments, musical performance style, as well as musical learning strategies. However, the extent to which these inter-individual perceptual differences are linked to musical training or represent innate properties related to musical aptitude remains unexplored. By examining elementary and complex auditory processing in children in a longitudinal design, Seither-Preisler et al. (2014) found evidence for relatively strong influences of predispositional factors, as compared to the influence of environmental factors and musical achievement. Nevertheless, the observed three clusters in the present work are a solid base for investigating larger samples in future works. Using a consequent longitudinal approach should help to shed light on the link between perception, neuroanatomy and training. Moreover, it would be of interest to inspect the related individual preferences for specific musical instruments and performance styles. The outcome of the present work may have relevance to musicians in several ways. First, if musicians better understand their specific auditory abilities it might help them to improve or adapt their musical techniques and training. Alternatively, these findings might lead to novel applications to the specific needs of the individual listener, be that a professional or amateur musician, or a patient with specific auditory needs (e.g. hearing impairments, noise sensitivity). Finally, these findings could contribute to our overall understanding of the variety of pitch and sound perception and musical preferences across individuals.
## Appendix A

### A.1 Data table

<table>
<thead>
<tr>
<th>SEX</th>
<th>SUBJECT DATA</th>
<th>INSTRUMENTS</th>
<th>AD JKT</th>
<th>JA T</th>
<th>KEL</th>
<th>MJR</th>
<th>EJR</th>
<th>JSR</th>
<th>FLA</th>
<th>HAR</th>
<th>H-AMAG</th>
<th>R-AMAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>age 14; sex 1; blood type A</td>
<td>Gitarre</td>
<td>16</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>age 15; sex 2; blood type B</td>
<td>Saxophon</td>
<td>17</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>14</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>age 16; sex 3; blood type O</td>
<td>Violine</td>
<td>18</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>15</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>M</td>
<td>age 17; sex 4; blood type AB</td>
<td>Posa</td>
<td>19</td>
<td>8</td>
<td>10</td>
<td>5</td>
<td>16</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SUBJECT DATA</th>
<th>INSTRUMENTS</th>
<th>AD JKT</th>
<th>JA T</th>
<th>KEL</th>
<th>MJR</th>
<th>EJR</th>
<th>JSR</th>
<th>FLA</th>
<th>HAR</th>
<th>H-AMAG</th>
<th>R-AMAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>age 18; sex 5; blood type A</td>
<td>Flute</td>
<td>20</td>
<td>9</td>
<td>11</td>
<td>6</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>M</td>
<td>age 19; sex 6; blood type B</td>
<td>Clarinet</td>
<td>21</td>
<td>10</td>
<td>12</td>
<td>7</td>
<td>18</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>F</td>
<td>age 20; sex 7; blood type O</td>
<td>Trompete</td>
<td>22</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>M</td>
<td>age 21; sex 8; blood type AB</td>
<td>Corno</td>
<td>23</td>
<td>12</td>
<td>14</td>
<td>9</td>
<td>20</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>
A.3 Subject Questionnaire (German)

MUSIK & GEHIRN
FORSCHUNGSprojekt
Fragebogen zur musikalischen Aktivität

Name ____________________________ Datum __________________

Alter ______________ Händigkeit ______________________

Universität ______________________ Schwerpunktfach ________

Aktives Musizieren

Übungszeiten zu Lebensabschnitten: Stunden pro Tag (h/T) oder pro Woche (h/W)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Passives Musizieren

Stunden pro Tag (h/T) oder pro Woche (h/W)
Mentales Musizieren / Üben aktuell: ca. ______
Konzentriertes Musikhören aktuell: ca. ______

Tinnitus / Geräuschempfindlichkeit

▪ Nein

▪ Ja, Tinnitus und zwar: □ ständig □ gelegentlich

▪ Ja, Geräuschempfindlichkeit und zwar: □ ständig □ gelegentlich

▪ Ich habe Interesse an einem ausführlichen Fragebogen zur genauen Tinnitus-Analyse

Musik & Gehirn Forschungsprojekt
Kontakt: jan.benner@unibas.ch
### A.4 Informed Consent (German)

Schriftliche Einverständniserklärung der Versuchsperson zur Teilnahme an einer wissenschaftlichen Studie

**Version 2 vom 01.12.2011**

- Bitte lesen Sie dieses Formular sorgfältig durch.
- Bitte fragen Sie, wenn Sie etwas nicht verstehen oder wissen möchten

<table>
<thead>
<tr>
<th>Nummer der Studie:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Titel der Studie:</strong></td>
<td>Auditorische Neuroplastizität im erwachsenen Gehirn</td>
</tr>
<tr>
<td><strong>Sponsor (vollständige Adresse):</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Ort der Studie:</strong></td>
<td>Abt. Neuroradiologie, Universitätsspital Basel</td>
</tr>
<tr>
<td><strong>Prüfärztin/Prüfarzt</strong></td>
<td>Dr. med. Maria Blatow</td>
</tr>
<tr>
<td><strong>Versuchsperson</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Name und Vorname:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Geburtsdatum:</strong></td>
<td></td>
</tr>
</tbody>
</table>

- Ich wurde vom unterzeichnenden Arzt/Wissenschaftler mündlich und schriftlich über die Ziele, den Ablauf der Studie, über die zu erwartenden Wirkungen, über mögliche Vor- und Nachteile sowie über eventuelle Risiken informiert.
- Ich hatte genügend Zeit, um meine Entscheidung zu treffen.
- Ich bin darüber informiert, dass eine Versicherung Schäden deckt, falls solche im Rahmen der Studie auftreten.
- Ich weiss, dass meine persönlichen Daten nur in anonymisierter Form an aussenstehende Institutionen zu Forschungszwecken weitergegeben werden. Ich bin einverstanden, dass die zuständigen Fachleute des Studienauftraggebers, der Behörden und der Kantonalen Ethikkommission zu Prüf- und Kontrollzwecken in meine Originaldaten Einsicht nehmen dürfen, jedoch unter strikter Einhaltung der Vertraulichkeit.
- Ich nehme an dieser Studie freiwillig teil. Ich kann jederzeit und ohne Angabe von Gründen meine Zustimmung zur Teilnahme widerrufen, ohne dass mir deswegen Nachteile entstehen.

**Ort, Datum** | **Unterschrift der Versuchsperson**
---|---


**Ort, Datum** | **Unterschrift der Prüfärztin/des Prüfarztes**
---|---

**Version 2 vom 01.12.2011**
A.5 Ethical Approval (German)

Beschlussmitteilung der Ethikkommission beider Basel


Titel des Forschungsprojektes

Auditivische Neuroplastizität im erwachsenen Gehirn

Ref.Nr. EK: 324/11

Prüferin

Name, Vorname, Titel: Blatow, Maria, Dr. med.
Funktion: Oberärztin, Abt. Diagnostische und Interventionelle Neuroradiologie
Adresse: Universitätsspital, 4031 Basel


X normales Verfahren □ vereinfachtes Verfahren □ Nachbegutachtung

Die Ethikkommission kommt zu folgendem Beschluss:

X A positiv
□ B positiv mit Bemerkungen (siehe Seite 2ff)
□ C mit Auflagen (siehe Seite 2ff)
□ Nachbegutachtung durch Ethikkommission notwendig □
□ schriftliche Mitteilung an Ethikkommission ausreichend □
□ D negativ (mit Begründung und Erläuterung für die Neubeurteilung) (siehe Seite 2ff)
□ E Nicht-Eintreten (mit Begründung) (siehe Seite 2ff)

Der Beschluss gilt auch für die im "Antrag auf Begutachtung" gemeldeten weiteren Prüferinnen im Zuständigkeitsbereich der Ethikkommission.

Pro Memoria: Pflichten des/der verantwortlichen Prüfer/in
- Geprüfte Produkte und Vergleichsprodukte (Arzneimittel und Medizinprodukte) müssen - zur Sicherstellung der Qualität und der Sicherheit - fachgerecht hergestellt, evaluiert und eingesetzt werden.
- Meldepflicht bei:
  a) schwerwiegenden unerwünschten Ereignissen (serious adverse events) unverzüglich
  b) neuen Erkenntnissen, die während des Versuchs verfügbar werden und die Sicherheit der Versuchspersonen sowie die Weiterführung des Versuchs beeinflussen können
  c) Änderung des Protokolls (Versuchsplans)
  d) Ende oder Abbruch der Studie
- Zwischenbericht: einmal pro Jahr
- Meldungs- oder Bewilligungspflicht von Studien bei Swissmedic bzw. anderen Bundes- oder kantonalen Behörden - sofern erforderlich (bei sponsorsierten Studien ist dies die Pflicht des Sponsors)
- Schlussbericht

Für die Ethikkommission:

Ort, Datum: Basel, 05. Januar 2012
Name(n): Prof. M. Kränzlin
                 Prof. A. P. Perruchoud

Unterschrift(en):
HDA 200 Audiometric Headphone

Closed dynamic headphones designed for extended high frequency testing.

**Features**

- Excellent passive attenuation (based on Peltor™ Ear Defenders)
- Very high quality sound reproduction
- Convenient single sided cable
- Padded headband and additional adjustable/removable cushions for increased comfort
- Soft, replaceable circumaural ear pads
- Color coded ear cups, right (red) left (blue)

**Technical Data**

- **Frequency response:**<20 Hz to >20,000 Hz
- **PTB calibrated:** see table
- **Transducer principle:** dynamic, closed
- **Nominal impedance:** 40 Ohm
- **Characteristic SPL:** 100 dB at 1 kHz, 1 mW
- **Max permanent load:** 500 mW
- **Coupling:** circumaural
- **Caliper pressure:** 10 N
- **Weight (with cable):** 330 g
- **Cable:** approx. 3 m, single-sided, open-ended
- **Connection:**
  - yellow + L
  - black − L
  - red + R
  - white − R

**HDA 200 Frequency Response Test Conditions**

- All measurements are done on a calibrated coupler B&K 4153 (artificial ear) with the standard cone YJ0304 above the adapter plate, type DB 0843.
- The pressure of the headband shall be 10N ± 1N.
- The RMS input voltage to the headphone is 0.5 V.
- The measurements are done with steady state sine wave signals.
- The output impedance of the signal source shall be <1 Ohm.
- Climatic conditions:
  - Temperature: T=20°C
  - Humidity: H=50%rel
  - Atmospheric pressure: P=approx. 100kPa

**Standard SPLs (dB 20µPa) @ .5 Vrms**

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Passive attenuation (db)</th>
</tr>
</thead>
<tbody>
<tr>
<td>125</td>
<td>112.5</td>
</tr>
<tr>
<td>250</td>
<td>113.0</td>
</tr>
<tr>
<td>500</td>
<td>112.0</td>
</tr>
<tr>
<td>750</td>
<td>111.0</td>
</tr>
<tr>
<td>1,000</td>
<td>108.5</td>
</tr>
<tr>
<td>2,000</td>
<td>104.0</td>
</tr>
<tr>
<td>3,000</td>
<td>104.0</td>
</tr>
<tr>
<td>4,000</td>
<td>104.0</td>
</tr>
<tr>
<td>5,000</td>
<td>106.5</td>
</tr>
<tr>
<td>6,000</td>
<td>107.5</td>
</tr>
<tr>
<td>8,000</td>
<td>105.5</td>
</tr>
<tr>
<td>9,000</td>
<td>105.0</td>
</tr>
<tr>
<td>10,000</td>
<td>102.5</td>
</tr>
<tr>
<td>11,200</td>
<td>102.0</td>
</tr>
<tr>
<td>12,500</td>
<td>103.0</td>
</tr>
<tr>
<td>14,000</td>
<td>98.5</td>
</tr>
<tr>
<td>16,000</td>
<td>100.0</td>
</tr>
</tbody>
</table>

All data are influenced by temperature, humidity and static pressure.

Sennheiser Electronic Corporation  One Enterprise Drive  PO Box 987, Old Lyme, CT 06371
Tel: 860-434-9190  Fax: 860-434-1759  Web site: www.sennheiserusa.com
A.7 Technical Manual - Interface

Tech Specs

Input AD: 6 x 1/4" TRS (4 x Line, 2 x Line/Instrument), 2 x XLR/TRS Combo connector (2 x Mic/Line), all servo-balanced.
Output DA: 6 x 1/4" TRS, servo-balanced, IC-coupled signal path. 1 x 1/4" TRS unbalanced.

Input Digital: 1 x ADAT optical or SPDIF optical; SPDIF coaxial (AES/EBU compatible)
Output Digital: 1 x ADAT optical or SPDIF optical; SPDIF coaxial (AES/EBU compatible)

MIDI: 2 x MIDI I/O via breakout cable (4 x 5-pin DIN jacks), for 32 channels low-jitter hi-speed MIDI

Dynamic range AD: 110 dB RMS unweighted, 113 dBA
THD AD: < -100 dB (< 0.001 %)
THD+N AD: < -96 dB (< 0.0012 %)
Crosstalk AD: > 110 dB

Dynamic range DA: 110 dB RMS unweighted, 113 dBA (unmuted)
THD DA: < -100 dB (0.001 %)
THD+N DA: < -96 dB (0.0015 %)
Crosstalk DA: > 110 dB

Input/Output level for 0 dBFS @ Hi Gain: +19 dBu
Input/Output level for 0 dBFS @ +4 dBu: +13 dBu
Input/Output level for 0 dBFS @ -10 dBV: +2 dBV

Sample rate internally: 32, 44.1, 48, 64, 88.2 kHz, 96 kHz, 128, 176.4, 192 kHz
Sample rate externally: 28 kHz - 200 kHz

Frequency response AD/DA, -0.1 dB: 5 Hz - 20.4 kHz (sf 44.1 kHz)
Frequency response AD/DA, -0.5 dB: 1 Hz - 43.1 kHz (sf 48 kHz)
Frequency response AD/DA, -1 dB: 1 Hz - 80 kHz (sf 96 kHz)
A.8 Conference Poster (ASA 2017)

Differences in sound perception are reflected by individual auditory fingerprints in musicians

Jan Berner1, Julia Reinhardt1, Elke Hofmann2, Christoph Steppich2, Peter Schecker2, Maria Blaton3, William I. Davies1

1 Department of Radiology, Division of Neuroradiology, University of Basel Hospital, Switzerland
2 Department of Neuroanatomy, University of Basel Medical School, Germany
3 Department of Psychology, University of Basel Research Centre, Basel University Hospital, Switzerland
4 Related Physics, University of Basel, Switzerland

Abstract
Musicians have been reported to show significant inter-individual differences in elementary hearing functions, sound perception mode, musical instrument preference, performance style, as well as more complex musical abilities like absolute- and relative pitch perception, and auditory imagery. However, it remains unexplored how individual elementary hearing functions and corresponding musical abilities are connected to and to what extent they reflect individual differences in the musical behavior of musicians.

Methods
Using 5 separate hearing tests, all subjects were individually tested for: musical imagery (AMMA tonal score, Gordon, 1998), absolute pitch perception (AP, Mengareth et al., 2014), relative pitch perception (RP, unpublished), pitch perception preference (S & O; Schröder et al., 2005), and frequency discrimination threshold (FD; Ewert and Dau, 2006).

Results
On the group level the assessed raw parameters (a) and calculated main components (b) were found to be represented significantly different in the three characteristic clusters of subjects (p < .001 for AMMA, RP, AP, SI; OE, RD, FD).

Conclusions
Taken together, the findings suggest that inter-individual differences in pitch & sound perception exist and are reflected by characteristic auditory fingerprints on the individual and group level, which may be at the base of the specific musical preferences, style and performance of musicians.

References
References


Delhommeau, K., Micheyl, C., Jouvent, R., & Collet, L. (2002). Transfer of learning


Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., . . . Rupp,


