

Cognitive and neural determinants of music appreciation and aesthetics

Eva Istók



Cognitive Brain Research Unit, Cognitive Science, Institute of Behavioural Sciences
University of Helsinki, Finland

Finnish Centre of Excellence in Interdisciplinary Music Research
University of Jyväskylä, Finland

Academic dissertation to be publicly discussed,
by due permission of the Faculty of Behavioural Sciences
at the University of Helsinki Main Building, Fabianinkatu 33
on the 22nd of August, 2013, at 12 o'clock

University of Helsinki
Institute of Behavioural Sciences
Studies in Psychology 88: 2013

Supervisors: Professor Mari Tervaniemi
Cognitive Brain Research Unit
Cognitive Science
Institute of Behavioural Sciences
University of Helsinki
Helsinki, Finland
&
Finnish Centre of Excellence in Interdisciplinary Music Research
University of Jyväskylä
Jyväskylä, Finland

Professor Uwe Seifert
Musikwissenschaftliches Institut
Philosophische Fakultät
Universität zu Köln
Köln, Germany

Reviewers: Professor Barbara Tillmann
Lyon Neuroscience Research Center
Team Auditory Cognition and Psychoacoustics
INSERM U1028 - CNRS UMR5292
University Claude Bernard Lyon 1
Lyon, France

Professor Marcel Zentner
Department of Psychology
University of York
York, United Kingdom

Opponent: Professor Eckart Altenmüller
Institute of Music Physiology and Musicians' Medicine
University of Music, Drama and Media
Hannover, Germany

ISSN-L 1798-842X
ISSN 1798-842X
ISBN 978-952-10-8667-0 (paperback)
ISBN 978-952-10-8668-7 (PDF)
<http://ethesis.helsinki.fi>
Unigrafia
Helsinki 2013

Contents

Abstract	5
Tiivistelmä	6
Acknowledgements	7
List of original publications	9
Abbreviations	10
1 Introduction	11
1.1 Music aesthetics from a neuropsychological perspective	11
1.1.1 Aesthetics and beauty	14
1.1.2 Music preference, liking, and evaluative judgments	15
1.1.3 Music emotions	17
1.1.4 Expressivity in music and processing fluency	18
1.2 Music emotions, preference, and beauty in childhood	19
1.3 Indices of aesthetic processing: Event-related potentials	21
1.3.1 Stimulus feature encoding and stimulus identification: N1 and P2	22
1.3.2 Expectation, attention, and target detection: N2b and P3a	23
1.3.3 Evaluative impression formation: Early negativity and ERAN	24
1.3.4 Affective stimulus processing: Late positive potential (LPP)	25
2 Aims of the thesis	27
3 Methods	28
3.1 Behavioural experiments	28
3.1.1 Participants	28
3.1.2 Stimuli	28
3.1.3 Experimental setup and data analysis	30
3.1.3.1 Free verbal association task	30
3.1.3.2 Child-adapted rating scales	31

3.2 ERP recordings	32
3.2.1 Participants	32
3.2.2 Stimuli and experimental procedures	33
3.2.3 Data acquisition and analysis	35
4 Results and discussions	37
4.1 Study I: Aesthetic responses to music	37
4.2 Study II: Aesthetic judgments in school-aged children	38
4.3 Study III: Music genre preference and evaluative judgments	42
4.4 Study IV: Expressive timing and phrase boundary processing	44
5 General discussion	48
5.1 The nature of aesthetic experiences in music	48
5.2 Music preference and aesthetic judgments in young children	49
5.3 Neural correlates of aesthetic processing	50
5.4 A model of music appreciation and aesthetics	52
6 Conclusions	55
References	57

Abstract

Music appreciation and aesthetics have been subject to philosophical considerations and empirical investigations for decades. These scientific endeavours have been complicated by the fact that multifarious biological and cultural factors interact in all musical actions.

The present thesis made use of behavioural methods and the event-related-potential technique (ERP) to investigate cognitive and neural determinants of music appreciation and aesthetics. At first, the structure and content of the knowledge-based concept of music aesthetics was examined on the basis of hundreds of verbal descriptions. In a second step, the developmental stage of school-aged children was investigated with regard to their recognition of music emotions, personal music preferences, and their concept of beauty. The event-related potential (ERP) method served to study whether long-term music preferences affect evaluative and cognitive judgments of music and whether expressive timing facilitates the neural processing of musical phrases.

The results suggest that one's experience of beauty is at the core of aesthetic responses to music among both music experts and laymen. In line, by the age of six years, children have already developed a mental concept of beauty that resembles that of adults. They are able to recognise music emotions and to express their own music preferences. These early music preferences may become stable over a lifetime and later seem to modulate neural correlates of judgment processes during active music listening. Expressive timing, in turn, determines the neural processing of musical phrases even when these are unattended. In addition, expressive timing seems to facilitate the sequential organisation of melodies by directing listeners' attention towards key musical events.

The concept of beauty, music preference, music emotions, and the action of successfully structuring musical information constitute major elements of music appreciation. Thus, this thesis deepens our understanding of the conceptual and neural underpinnings of pleasurable musical experiences and allows extending and transferring current general models of aesthetic processing to music-specific demands.

Tiivistelmä

Musiikkimieltymyksiä ja –estetiikkaa on lähestytty tieteessä sekä filosofian että kokeellisen tutkimuksen avulla. Koska nykyäsityksen mukaan kaikki musiikkiin liittyvä toiminta perustuu sekä biologisiin että opittuihin tekijöihin, on alan tutkimus suurten haasteiden edessä.

Tässä väitöskirjassa musiikkimieltymyksiä ja esteettisiä arvioita tutkittiin hyödyntämällä useita lähestymistapoja. Ensimmäiseksi selvitettiin musiikin esteettisten arvioiden käsittehierarkiaa vapaan assosiaation menetelmällä, jossa aikuiset koehenkilöt listasivat musiikkiin liittyviä adjektiiveja. Seuraavaksi kouluikäisten lasten musiikkiemootioita, –mieltymyksiä ja kauneuskäsitteitä tutkittiin innovatiivisella pelinomaisella tutkimusalustalla. Lisäksi kahdessa tutkimuksessa aivojen tapahtumasidonnaisia jännitevasteita käytettiin osoittamaan, kuinka eri musiikkityyleihin kohdistuvat musiikkimieltymykset (”fanittaminen”) sekä musiikissa käytetyt tunnelmaisun keinot vaikuttavat musiikin hermostolliseen käsittelyyn.

Tulosten mukaan kauneuden kokemus on olennaista musiikin kuulijoiden esteettisissä arvioissa riippumatta heidän musiikillisen harjaantuneisuutensa määrästä. Vastaavasti jo 6-vuotiailla on hallussaan samankaltainen kauneuden käsite kuin aikuisilla. Vakiinnuttuaan musiikkimieltymykset ilmenevät musiikin hermostollisessa käsittelyssä eriytyen hienovaraisesti kuuntelutehtävän ohjeistuksen/tavoitteiden? suhteen. Musiikilliseen ilmaisuun liittyvien musiikin ajallisten muutosten hermostollinen käsittely on automatisoitunutta, joskin musiikkiin kohdistettu tietoinen tarkkaavaisuus voi musiikin käsittelyä vielä tehostaa.

Väitöskirjan tulosten mukaan kauneuden käsite, musiikkimieltymykset sekä –emootiot sekä musiikin ajallisen ja ilmaisullisen rakenteen hahmottaminen ovat keskeisiä musiikillisen toiminnan elementtejä, joiden käsitteellistä ja hermostollista prosessointia voidaan monimetodisella tutkimusotteella luotettavasti selvittää.

Acknowledgements

The present thesis was carried out at the Cognitive Brain Research Unit (CBRU), Institute of Behavioural Sciences, University of Helsinki. It was made possible by the financial support from the Centre for International Mobility (CIMO), the German Academic Exchange Service (DAAD), the European Commission (FP6-BrainTuning), and the Centre of Excellence in Interdisciplinary Music Research (CoE, Academy of Finland).

I express my deepest gratitude to my supervisor, Prof. Mari Tervaniemi, for her support, her guidance and patience and especially for her trust in my work. I further wish to thank her for her openness and for giving me the opportunity to become part of an international BrainTuning network thereby getting insights into academics that go beyond science and research. I am also grateful for the time spent with her lovely children, Lasse and Elsa. My thesis would have not been possible without Prof. Risto Näätänen, Prof. Teija Kujala, and Prof. Petri Toiviainen who all created a perfect working environment at CBRU and the CoE. I also wish to thank Prof. Uwe Seifert for encouraging me to enter the fields of psychology and cognitive neuroscience.

I am also grateful to my co-authors Doc. Elvira Brattico, Doc. Minna Huotilainen, Prof. Thomas Jacobsen, Prof. Anders Friberg, Sirke Nieminen, Kaisu Krohn, Aileen Ritter, and Mira Müller. My work has significantly benefitted from their excellent scientific experiences and their will to share it. I wish to thank also Prof. Barbara Tillmann and Prof. Marcel Zentner who helped improving the present thesis by giving precious comments and suggestions.

Many colleagues and friends at the CBRU and the CoE not only contributed to my work but also established a welcoming atmosphere and lovely surrounding which made my time in Finland an unforgettable experience. Thank you, Eino Partanen, Tuomas Teinonen, Teppo Särkämö, Riia Kivimäki, Lilli Kimppa, Sari Ylinen, Satu Pakarinen, Vesa Putkinen, Alexander Sorokin, Anke Sambeth, Kristina Relander-Syrjänen, Maria Mittag, Alina Leminen, Miia Seppänen, Ritva Torppa, Tuomas Eerola, Jaakko Erkkilä, Suvi Saarikallio, Jörg Fachner, Marc Thompson, Geoff Luck, Vinoo Alluri, Jonna Vuoskoski, Esa Ala-Ruona, Rafael Ferrer Flores, Marko Punkanen, Birgitta Burger, Anemone van Zijl, Olivier Lartillot, and Markku Pöyhönen. I wish to thank especially Tommi Makkonen, Miika Leminen, Piiu Lehmus, Marja Junnonaho, and Anna Rämä for

always being there when help was needed. I also wish to thank all members of the BrainTuning project for three exciting years.

My sincere gratitude goes towards my friends Gina Zarkada and Vasilis Chousionis. Thank you so much for your warmth and liveliness. Very special thanks also go to my friends and companions Veerle Simoens and Clemens Maidhof. I also wish to express my gratitude to my mother, my father, and my sister for their patience and trust, not only during the PhD years.

With love and gratitude, I wish to thank my husband Kai for all his love and understanding. Thank you for always taking things with a twinkle in your eye. Finally, the acknowledgement would be incomplete without mentioning my daughters Mathilde and her little sister. Life has been wonderful and will be even more astonishing with the two of you.

Hamburg, February 2013

Eva Istók

List of original publications

This thesis is based on the following publications:

- I** Istók, E., Brattico, E., Jacobsen, T., Krohn, K., Müller, M., & Tervaniemi, M. (2009). Aesthetic responses to music: A questionnaire study. *Musicae Scientiae*, 13(2), 183–206.
- II** Nieminen, S., Istók, E., Brattico, E., & Tervaniemi, M. (2012). The development of the aesthetic experience of music: preference, emotions, and beauty. *Musicae Scientiae*, 16(3), 372–391.
- III** Istók, E., Brattico, E., Jacobsen, T., Ritter, A., & Tervaniemi, M. (2013). ‘I love Rock ‘n’ Roll’ – Music genre preference modulates brain responses to music. *Biological Psychology*, 92(2), 142–151.
- IV** Istók, E., Friberg, A., Huotilainen, M., & Tervaniemi, M. (2013). Expressive timing facilitates the neural processing of phrase boundaries in music: Evidence from event-related potentials. *PLoS One*. 8(1), e55150. doi: 10.1371/journal.pone.0055150.

The publications are referred to in the text by their roman numerals. The articles are reprinted with the kind permission of the copyright holders.

Abbreviations

ANOVA	analysis of variance
BAEP	brain stem auditory evoked potentials
dB	decibel
EEG	electroencephalogram
ERP	event-related potential
HM-music	heavy metal music
HMM-fans	heavy metal music fans
ISI	inter-stimulus interval
IOI	inter-onset interval
LA-music	Latin-American music
LAm-fans	Latin-American music fans
ms	milliseconds
MLAEP	middle latency auditory evoked potentials
mOFC	medial orbito-frontal cortex
LLAEP	long latency auditory evoked potentials
LPP	late positive potential
μV	microvolt

1 Introduction

1.1 Music aesthetics from a neuropsychological perspective

Music appreciation is a complex phenomenon that comprises a host of biological and cultural factors that in combination contribute to why we listen to music and why music plays such an important role in human life. Music has been subject to systematic empirical investigations of aesthetic perception already since the second half of the 19th century. First and foremost, these early beginnings were due to the new understanding of auditory physiology as provided by the work of Hermann von Helmholtz (1821–1894).

The advent of empirical aesthetics as a scientific discipline is commonly dated back to Gustav Theodor Fechner (1801–1887). Empirical aesthetics is considered one of the oldest branches in psychology. In his *Vorschule der Ästhetik* (1867), Fechner aimed to establish an ‘aesthetics from below’ that was concerned with the systematic description of rules which may account for the influence of stimulus properties on the resulting experience of liking and pleasantness. Fechner thereby took a mainly descriptive approach without seeking to explain the psychological mechanisms underlying aesthetic experiences.

Several decades later, Daniel E. Berlyne (1924–1977) introduced his *new experimental aesthetics* which primarily aimed at investigating the effects of stimulus properties on physiological and emotional reactions (Allesch, 2006). Central to his approach was the *arousal potential* of aesthetic stimuli that is determined by their inherent properties such as complexity and novelty. The relationship between the hedonic value of these so-called collative variables and the arousal elicited in the perceiver was expressed in terms of an inverted U-shaped function. That is, increasing *arousal potential* is positively correlated with an increase in hedonic experience. When the *arousal potential* reaches a certain level, the hedonic value decreases and may result in even negative experiences.

As a result of the cognitive revolution beginning in the 1960s, the study of emotions and affective phenomena in response to aesthetic objects were neglected until recently. Nowadays, the interest in aesthetics from psychological and neurobiological perspectives is growing and new models and theories of aesthetic appreciation and its neural underpinnings are currently being developed (e.g., Leder, Belke, Oeberst, & Augustin, 2004; Ramachandran & Hirstein, 1999; Reber, Schwarz, & Winkielman, 2004; Silvia, 2005). Following Berlyne’s tradition to account for emotional experiences resulting from

the perception of aesthetic stimuli, Leder et al. (2004) proposed a model of aesthetic processing that rests on several stages of information-processing (Figure 1).

Accordingly, aesthetic experiences may be the result of an automatic and deliberate integration of several related cognitive and affective processes. Although the model primarily focuses on modern visual art, the authors suggest that it may be applicable also to other domains such as music. Hence, when listening to music, an initial perceptual analysis informs about acoustic properties such as intensity, pitch, and timbre. This information is then implicitly integrated into memory and compared against previous experiences. That is, listeners automatically appraise whether the music is familiar or not and whether it is prototypical for a broad music genre such as classical music or heavy metal music. Indeed, listeners seem to be able to recognize music genres of sound examples of 250 ms duration (Gjerdingen & Perrott, 2008) and to determine whether they are familiar with a song already after 500 ms of music presentation (Filipic, Tillmann, & Bigand, 2010).

The classification of music according to its content and specific style may become conscious and explicit at the following stage of the model. The categorization of music into genres can now be verbalized. At the stage of explicit classification, music expertise and preferences may determine the way music is processed. For instance, jazz musicians may have a more elaborate classification system regarding the various sub-genres of jazz music than rock musicians have. Similarly, heavy metal music fans may process heavy metal music differently than Latin-American music fans.

According to the model, all of the proposed cognitive processing stages are accompanied by affective responses that are eventually evaluated and lead to two types of outcome: aesthetic judgments and aesthetic emotions. Aesthetic judgments are understood as more cognitive way of judging, for instance, music in terms of its quality whereas aesthetic emotion may lead listeners to turn on the volume and to intensify their emotional experience or to switch the radio channel to avoid listening to music that is displeasing. Although the model roughly distinguishes between aesthetic judgments and aesthetic emotions, a clear description of the quality of aesthetic emotions is not provided. What exactly is an aesthetic emotion? How does it feel like? Is it an unspecific positive response or could it be best describe as an experience of beauty?

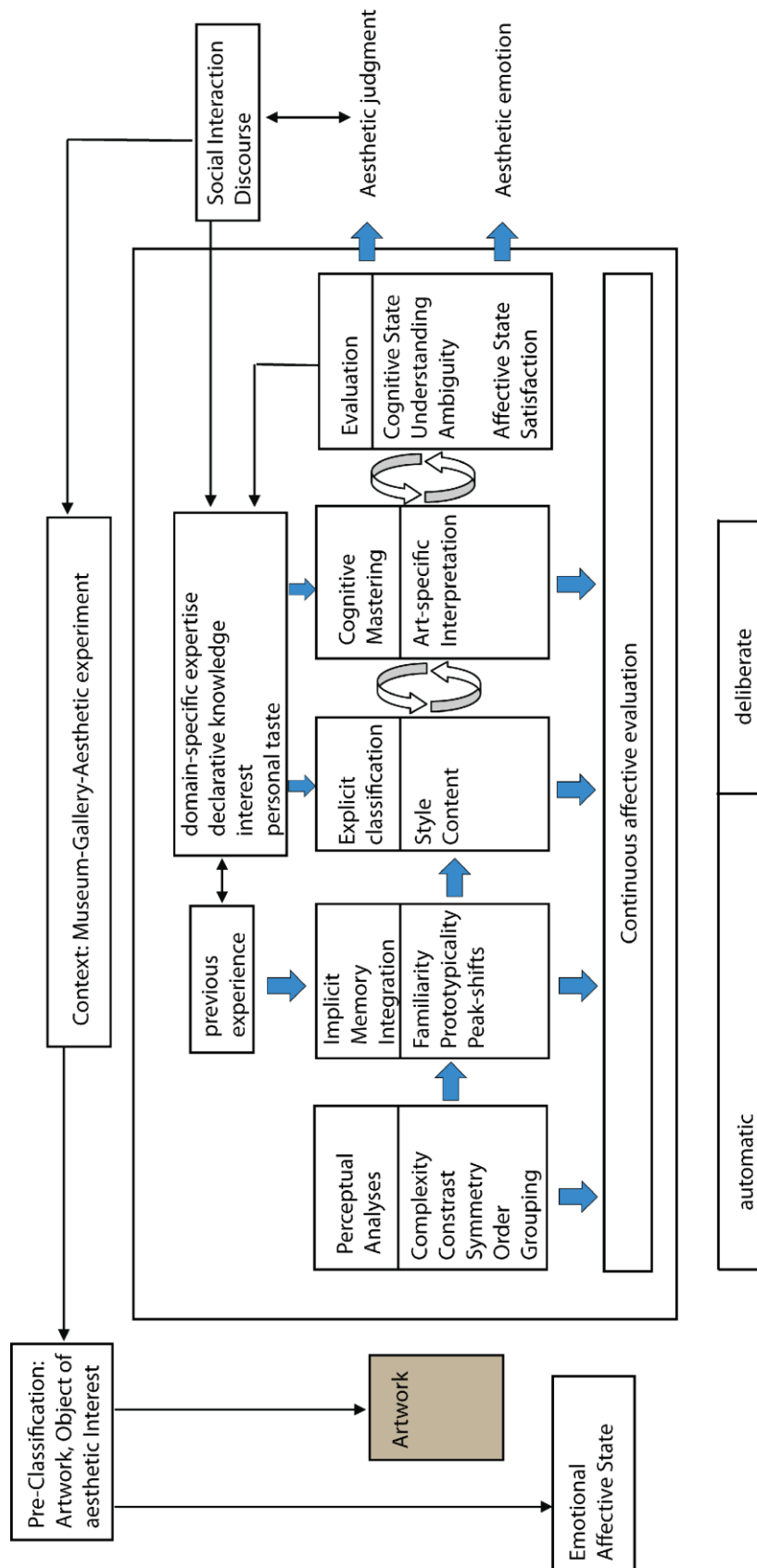


Figure 1. A model of aesthetic experience and aesthetic judgments. Adapted from Leder et al. (2004).

In addition, the model does not take into account possible influences of the emotional content expressed by an aesthetic object on the aesthetic outcome processes. Especially in music, aesthetic response may be dependent on whether the music expresses, for instance, happiness or sadness (for a review, see Hunter & Schellenberg, 2010). The model also does not make any assumptions about when the different processing stages develop during childhood. Young children at the age of 3 are able to actively choose music they like (Lamont, 2008). Are these liking judgments of aesthetic nature?

Another challenge is posed by the fact that music is a dynamically evolving sound continuum. That is, the analysis of the complexity or the sequential grouping structure of music requires the on-line processing and integration of successive musical events into meaningful units over time. Although the initial encoding of genre-specific acoustic properties such as distorted guitar sounds may account for rapid evaluative judgments, musical structural features such as harmony, rhythm and tonality certainly contribute to why music may be aesthetically pleasing or not. Beyond musical structural features, subtle modulations of the micro-structure of a music performance that render music expressive may affect the aesthetic processing of music.

The experience of beauty, music emotions, preferences, and expressivity in music may substantially contribute to music appreciation and are therefore examined in more detail in the following.

1.1.1 Aesthetics and beauty

The study of aesthetics has traditionally been associated with beauty. As shown by Jacobsen, Buchta, Kohler, & Schröger (2004), the link between the two concepts is not a factitious one. Instead, it is vivid also in the everyday understanding of aesthetics of young adults. Asking several hundreds of participants to list terms describing the aesthetics of objects, the authors found that *beautiful* was named by 92 % of the participants. In contrast, only 42 % of the volunteers listed the adjective *ugly*. All other terms were less frequently named. The results convincingly demonstrate that the concept of beauty is central to the investigation of aesthetic processes and art appreciation.

Notwithstanding, up to date there is no consensus about the origin and emergence of beauty. What is beauty? The objectivist view considers beauty as an attribute that is determined by object qualities and, hence, independent of the perceiver (Tatarkiewicz,

1963). In contrast, the subjectivist perspective posits that ‘beauty lies in the eye of the beholder’. Beauty is regarded as subjective experience that depends on various psychological functions and factors, for instance, individual preferences, expertise, or personality traits (Tatarkiewicz, 1963). As such, the experience of beauty is prone to huge inter-individual differences.

An intermediate account of beauty was provided by Reber et al. (2004). The authors suggest that the experience of beauty is on the one hand determined by objective stimulus features. However, the processing of these stimulus features to a great extent depends on the perceiver. Familiarity, preferences, personality, and expertise may affect the way in which object properties such as hue and saturation in colour perception, or consonance and dissonance in music are processed. The more fluently stimulus properties can be processed the more positive is going to be the resulting aesthetic pleasure. Accordingly, beauty is “*a pleasurable subjective experience that is directed toward an object and not mediated by intervening reasoning*” (Reber et al., 2004; p. 365).

A brain-based theory of beauty was recently proposed by Ishizu and Zeki (2011). They found that the medial orbito-frontal cortex (mOFC) was activated when participants experienced the beauty of both music and paintings suggesting that there is a domain-general faculty of beauty. Corroborating findings, including however only visual stimuli, were provided by Jacobsen, Schubotz, Höfel, and von Cramon (2006) and Hideaki and Semir (2004). As evident, psychological and neuroscientific approaches regard beauty as an inherently human experience rather than an abstract concept that is subject to theoretical considerations only. Beauty, in this regard, is not limited to a handful of music experts. Instead, it is conceived as an ubiquitous experience that similar to emotions may occur on a daily basis, sometimes mildly and some times more intense.

1.1.2 Music preference, liking, and evaluative judgments

Music preferences commonly are treated as affective states (Scherer & Zentner, 2001) because they are strongly connected with valence, that is, with positive or negative experiences resulting, for example, from listening to preferred compared to non-preferred music. Importantly, music preferences begin to develop during early childhood when parents and siblings greatly influence the choice of music. Music related activities gain special importance in adolescence (North, Hargreaves, & O'Neill, 2000; Saarikallio &

Erkkilä, 2007; Schwartz & Fouts, 2003). At that time, music preferences consolidate but still may be plastic and affected by peer groups until they become more stable during adulthood (Mulder, Ter Bogt, Raaijmakers, Gabhainn, & Sikkema, 2010).

Music preferences have a strong connection to familiarity resulting from the repeated exposure to the preferred music genre (North & Hargreaves, 2008; Witvliet & Vrana, 2007). In this sense, a specific form of expertise may evolve that does not necessarily require active musical training.

Music preference may also be linked to personality traits (e.g., Rentfrow & Gosling, 2003; Zweigenhaft, 2008). For instance, preference for intense and rebellious music such as heavy metal or rock seems to positively correlate with openness to new experiences (Rentfrow & Gosling, 2003). Further, Schäfer and Sedlmeier (2009) found that physiological arousal and the use of music for communication and self-reflection strongly determine music preferences.

Music preferences are often quantified using liking and disliking judgments. However, there is a need to distinguish between these two affective responses. Evaluative judgments in terms of liking or disliking judgments refer to a more immediate affective response to music. According to Zajonc (1980), evaluative judgements are fast, automatic and holistic and they may result in spontaneous behaviour of approaching or avoiding an aesthetic object such as skipping the radio channel or increasing the sound volume. In contrast to aesthetic judgments described as outcome of aesthetic processing in the model by Leder et al. (2004), evaluative judgments as understood by Zajonc (1980) do not need cognitive mediation. Empirical evidence regarding the rapidness and automaticity of liking judgments during aesthetic processing is scarce. Yet, recent findings suggest that the evaluative categorization of visual aesthetic objects happens within 600 to 1000 ms after stimulus onset (for a review, see Cela-Conde, Agnati, Huston, Mora, & Nadal, 2011; see also 1.4.).

Importantly, evaluative judgments are presumably closely tied to individual long-term music preference. That is, as also suggested by Leder et al. (2004), listening to music may automatically trigger memory structures that are not limited to explicit knowledge about, for instance, a certain music genre, but may also entail emotions implicitly associated with the music.

1.1.3 Music emotions

The study of music emotions has significantly increased during the last two decades (for a review, see Eerola & Vuoskoski, 2013). It has, for instance, been argued that one of the most important reasons for listening to music is that emotions may be evoked (for an overview, see Sloboda 2010). In fact, adolescents use music for self-display (North et al., 2000) and to regulate their moods and emotions (North et al., 2000; Saarikallio & Erkkilä, 2007). Thus, an investigation of music appreciation and aesthetics needs to consider the emotional content of music.

Two models of emotions have dominated research on music emotions: basic emotion models and dimensional models such as the circumplex model of emotions (Russel, 1980). The basic emotion model assumes a limited number of discrete emotions, such as happiness, sadness, anger, fear, and disgust. Basic emotions are believed to be innate and universal (Ekman, 1992). This assumption is supported by findings showing that adults consistently and effortlessly judge music as happy or sad (Peretz, Gagnon, & Bouchard, 1998). The circumplex model of emotions, in contrast, suggests that emotions may be best described in terms of their valence and their arousal. Both dimensions are conceived of a continuum from positive to negative (valence) and arousing to sleepy (arousal) along which emotions may be classified. For instance, happiness would be characterized by medium arousal and positive valence.

The applicability of the different emotion models to music emotions has repeatedly been discussed along with the question whether music induces or only expresses emotions (Gabrielsson, 2001/2002). Konečni (2008) suggests that emotions elicited during music activities might be best described as states of *being moved* or *aesthetic awe*. Aiming to compile emotions that best describe affective states induced by music, Zentner, Grandjean, and Scherer (2008) found that music emotions require a more nuanced description than available in current emotion models. Specifically, adjectives such as *dreamy*, *nostalgic*, *sentimental* or *touched* were among the most frequent terms used to denote music-induced emotions.

How do emotions expressed by music relate to aesthetic responses and liking judgements? North and Hargreaves (1997) showed that both liking ratings and the rated arousal potential of pop music excerpts predicted emotions expressed by music. In line with the circumplex model of emotions, music excerpts that receive high mean scores for liking and low mean scores for arousal were at the same time rated as more relaxing. The

study provides strong evidence for a close connection between emotions expressed by music and liking for music. Supporting this assumption, Thoma, Ryf, Mohiyeddini, Ehlert, and Nater (2012) found that the choice of music in everyday-listening situations is determined by the emotional content of the music. As a consequence, research in music aesthetics and appreciation needs to carefully take into account the emotion expressed or induced by music.

1.1.4 Expressivity in music and processing fluency

An investigation of music appreciation and aesthetics would not be complete taking into account only music structural features, music emotions, and listener characteristics (e.g., music preferences, personality). Instead, a great deal of why music is so deeply rooted in human culture is the communicative role it plays between the music performer and the listener. On a large scale, music performers may express and transmit specific emotions to their audience. Music performers may also influence the way listeners organize musical sounds by slightly varying acoustic features, for example, tempo, timbre, or intensity.

The term expressivity or musical expression entails both the communication of emotions as well as small-scale variations in acoustic parameters that support, for instance, the perception of musical groups (Gabrielsson, 1987) or metrical accents (Sloboda, 1983). Juslin (2003) has identified three more components of musical expression. First, ‘random variability’ that is due to limitations in perceptual-motor skills. In fact, these fluctuations contribute to the lively character of performances. Second, ‘biological motion’ refers to the transformation of specific human motion patterns to music. For instance, Friberg and Sundberg (1999) showed that the specific timing pattern of final *ritardandi* (i.e., the slowing down at the end of a musical phrase) is similar to the deceleration of runners. Third, performers may intentionally increase the level of unexpectedness by delaying a certain musical event. As already Meyer (1956) noted, such a delay creates tension which, in turn, may contribute to the affective character of music.

Importantly, expression is regarded as a function of both the acoustic properties of a performance and the way these properties are perceived by the listener (Juslin, 2003). That is, for musical communication to be successful there must be a match between performers’ expression and listeners’ perceptual abilities. In fact, such an agreement between

performer and listener may account for the greater success of some performers over others.

As suggested by Juslin (2003), there is a complex relationship between performance expression and music appreciation and aesthetics. One putative mechanism that could explain why expressively played music is preferred to deadpan music (i.e., music that lacks any expression) (Juslin & Laukka, 2003) is the concept of processing fluency as introduced by Reber et al. (2004). Music performers may, for instance, slow down the tempo and reduce the intensity of the played notes at the end of a musical phrase. These cues inform about the closure of a musical group and at the same may prepare the listener for the occurrence of a significant musical event, such as the beginning of a new phrase. The extraction of temporal and loudness information may lead to the allocation of attention resources that, in turn, facilitate the processing of the music's grouping structure. The successful processing of musical structural information eventually may result in a pleasurable experience.

1.2 Music emotions, preference, and beauty in childhood

The recognition of basic emotions in music such as happiness or sadness is effortless and greatly consistent among adults (Peretz et al., 1998). Tempo and mode thereby provide essential cues to determine whether music sounds sad or happy. Slow tempo and minor mode are associated with sadness whereas music played with fast tempo and composed in major mode is commonly considered happy (for a review, see Juslin & Laukka, 2003). At what age do children recognize basic emotions in music?

It is conceivable that the extraction of temporal information and the following interpretation in regard to emotion has a biological foundation because tempo is regarded as a domain-general signal property. For instance, fast tempo is commonly associated with heightened arousal (Trehub, Hannon, & Schachner, 2010). Given the diversity of musical systems all over the world, it can be assumed that modes are music-specific and culturally shaped. That is, to being able to recognize music emotions, presumably children first need to learn the cultural conventions in regard to the link between, for instance, major mode and happiness.

Several studies have shown that young children are able to assign basic emotion labels to music at the age of 3 to 5 years. To do so, children predominantly rely on tempo

information and only later, at the age of 6 to 9 years, make use of mode cues (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001). Kastner and Crowder (1990), however, found that even 3-year-olds are capable of distinguishing happy and sad music based on mode information. Interestingly, the ability to associate happiness with the major mode seems to emerge earlier in development than the ability to link minor mode with sadness (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Gerardi & Gerken, 1995).

Already newborns show preferences for consonant compared to dissonant music (e.g., Trainor & Heinmiller, 1998; Zentner & Kagan, 1996) suggesting that preferences for basic acoustic properties may be innate. Conscious liking judgments, however, may be based on a more sophisticated assessment of musical characteristics that – as suggested in the model by Leder et al. (2004) – requires the integration of a variety of cognitive and affective components. Generally, school-aged children are significantly less knowledgeable about music than adults. That is, children are less likely to base their preference judgments on extra-musical information (e.g., knowledge about the composer and the historical context of a composition). It is more probable that familiarity or the emotions conveyed by music rather than music-specific knowledge is the driving force for preference judgments in children. Familiarity with music results from the repeated exposure to, for instance, children songs. In fact, Lamont (2008) showed that 3-year-old children preferably listen to children's music, especially to nursery rhymes and songs from television programs.

At last, little is known about the development of aesthetic categories such as beauty and ugliness during childhood. When do children use these concepts to describe music and other art domains? Do they experience beauty similar to adults? In line with the objectivist view, research in regard to aesthetic development during childhood has been primarily concerned with the effects of intra-musical features on aesthetic perception as quantified in liking or preference judgments (Barrett, 2006). The characterization of the quality of aesthetic responses has received far less attention.

The development of understanding emotions has been shown to be positively correlated with the ability to use affective language to describe subjective emotional states in school-aged children (Ornaghi & Grazzani, 2012). Thus, the investigation of the use of aesthetic descriptors in children needs to take into account the stages of emotional and linguistic development that may provide the necessary grounds for the experience of beauty.

1.3 Indices of aesthetic processing: Event-related potentials

One possibility to study the nature and time course of aesthetic processing in the human brain is provided by the event-related potential technique (ERP). ERP is a non-invasive method that enables the extraction of specific components from the continuous electroencephalogram (EEG). The EEG measures voltage fluctuations as a result from synchronized neural activity originating from post-synaptic potentials of neocortical pyramidal cells (Luck, 2005). ERPs are quantified by averaging brain responses time-locked to internal or external events and, hence, may be associated with specific cognitive or affective functions. The amplitude and latency of ERPs provide information about the strength and the time course of underlying brain activity. Due to its high temporal resolution in the range of milliseconds, the ERP-technique is especially suited to study rapid neural processes.

Auditory ERP components are typically classified according to their latency. Brain stem evoked potentials (BAEPs) are generated in the cochlea and brain stem and can be measured at already at 1 to 8 ms after stimulus onset. BAEPs are sensitive to intensity and location (Picton, Stapells, & Campbell, 1981) and provide first information about the encoding of acoustic stimulus properties. Middle latency evoked potentials (MAEPs) are generated in the thalamus and auditory cortex and can be measured at around 10 to 80 ms after stimulus onset (Liegeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994). MAEPs reflect subsequent stages of stimulus feature processing such as the encoding of pitch information (Alho, Grimm, Mateo-Leon, Costa-Faidella, & Escera, 2012). Of main interest for the present work are the long-latency auditory evoked potentials (LLAEPs) that may be associated with higher-order processes involved in music perception and cognition as well as in music appreciation.

The ERP-technique has been used to study neural correlates of music processing (e.g., Koelsch, Gunter, Friederici, & Schröger, 2000; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Regnault, Bigand, & Besson, 2001; Schön, Regnault, Ystad, & Besson, 2005; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005; Tervaniemi, 2003). There are, however, relatively few studies that have made use of the ERP technique to investigate aesthetic and evaluative processes during music listening (Brattico, Jacobsen, De Baene, Glerean, & Tervaniemi, 2010; Müller, Höfel, Brattico, & Jacobsen, 2010).

Several components may play a specific role during the time course of cognitive and affective music processing underlying music appreciation. For instance, LLAEPs such as the N1 and P2 are associated with stimulus feature encoding and stimulus identification needed for subsequent cognitive and affective processing stages. Later components such as the N2b and P3a reflect higher-order mechanisms of attention and memory. Finally, the late positive potential (LPP) has been repeatedly associated with the processing of valence and arousal. These components are briefly characterized in the following.

1.3.1 Stimulus feature encoding and stimulus identification: N1 and P2

The N1 component of the ERP occurs at around 100 ms after stimulus onset. It is typically maximal over fronto-central electrode sites. The N1 amplitude has been shown to be sensitive to acoustic features such as intensity, frequency, duration, and inter-stimulus interval (ISI) (Näätänen & Picton, 1987; Woods, 1995). Further, Schön et al. (2005) found that the N1 amplitude in musicians was sensitive to the rated pleasantness of harmonic intervals suggesting that even isolated musical events may trigger immediate affective responses in the listener.

The N1 component is commonly followed by a positive peak between approximately 150 and 200 ms after stimulus onset. The P2 is maximal over the vertex and reflects processes of stimulus identification and classification. Similar to the N1, the P2 has been shown to be sensitive to acoustic properties such as intensity (Adler & Adler, 1989) and frequency (Wunderlich & Cone-Wesson, 2001). Both N1 and P2 occur in response to attended and non-attended stimuli. Moreover, the P2 amplitude has been shown to be associated with rapid auditory learning (Seppänen, Hämäläinen, Pesonen, & Tervaniemi, 2012; Tremblay, Inoue, McClannahan, & Ross, 2010). The exact functional meaning of the P2 is currently not fully understood (for a review, see Crowley & Colrain, 2004).

A recent study has shown that the P2 amplitude may be diminished when the temporal occurrence of a stimulus can be predicted (Sowman, Kuusik, & Johnson, 2012). Studying the processing of phrase boundaries in music, Knösche et al. (2005), and Nan, Knösche, and Friederici (2009) found that the P2 was significantly enhanced when the onset of the phrase was preceded by a pause. It was argued that the increase in P2 amplitude possibly reflects recovery of neural populations. Notwithstanding, the P2 amplitude might also be modulated by top-down processes during phrase boundary

processing in music such as harmonic context, and musical expertise (Neuhaus, Knösche, & Friederici, 2006).

1.3.2 Expectation, attention, and target detection: N2b and P3a

The N2b and P3a component of the ERP are believed to reflect higher-order mechanisms such as attention and memory. The N2b is commonly associated with the detection of auditory targets (Folstein & Van Petten, 2008). For instance, in typical oddball paradigms, the N2b is elicited when a rarely presented target is attentively detected (Näätänen, Simpson, & Loveless, 1982; Novak, Ritter, Vaughan, & Wiznitzer, 1990). The N2b is maximal over centro-parietal sites and occurs within 200 to 300 ms after stimulus onset. Palmer, Jewett, and Steinhauer (2009) showed that the amplitude of the N2b was larger for unexpected compared to more expected target events which consisted of single-tone timbre changes embedded at earlier (unexpected) or later positions (more expected) within musical sequences. Further, the amplitude of the N2b has been shown to be reduced when sound discrimination becomes more difficult, that is, when task demands increase (Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). Together, the results suggest that the N2b reflects conscious stimulus processing which is modulated by the degree of expectedness of a target stimulus (Palmer et al., 2009) and the ease with which this target stimulus can be discriminated from the background information (Tervaniemi et al., 2005).

The N2b is commonly followed by the P3a component which has been frequently linked to stimulus-dependent attention mechanisms. The P3a is a subcomponent of the P300 and has been interpreted as reflecting top-down control of focal attention towards novel auditory events (for a review, see Polich, 2007). In contrast to the N2b, the P3a may occur to infrequent ‘novel’ auditory events also in the absence of a task. However, the P3a amplitude is sensitive to task demands. Increasing task demands result in the reduction of the P3a amplitude suggesting that inhibitory processes are engaged to facilitate stimulus processing. That is, the processing of novel events may require additional attention resources that become available through the inhibition of unrelated activity resulting in an enlarged P300/P3a amplitude. When task demands increase, attention resources might be limited because more resources are needed for inhibitory control (Polich, 2007).

Although only a few studies have addressed effects of expectation, attention and task demands during music processing as reflected in the N2b-P3a complex of the ERP

(Palmer et al., 2009; Tervaniemi et al., 2005), these components provide a promising measure of top-down mechanisms that may contribute to facilitated stimulus processing of key musical events such as phrase boundaries.

1.3.3 Evaluative impression formation: Early negativity and ERAN

Evaluative processing has been associated with two specific components in the ERP (Jacobsen & Höfel, 2003). First, an early frontal negativity may reflect the formation of a first impression of the percept in terms of its positive or negative value. Second, the LPP (see also 1.3.4) has been associated with evaluative categorization.

The early frontal negativity has been repeatedly obtained in response to non-beautiful but not to beautiful judgments of graphical black and white pattern (Höfel & Jacobsen, 2007a; Jacobsen & Höfel, 2003). Similar early effects have been found for the processing of attractive compared to non-attractive faces (Werheid, Schacht, & Sommer, 2007). Here, the processing of attractive faces was reflected in an early posterior negativity (EPN). Both components reported on were elicited within 400 ms after stimulus onset while their respective scalp distribution and polarity differed depending on the stimulus material used. The results suggest that domain-specific evaluative mechanisms are triggered early during stimulus processing.

In music, the early right anterior negativity (ERAN) as first reported by Koelsch et al. (2000) is typically elicited in response to irregularities in musical syntax (e.g., untypical chords such as the Neapolitan chord). Music experts commonly show larger ERAN amplitudes for irregular or unexpected musical events than laymen (Koelsch, Schmidt, & Kansok, 2002). The ERAN may be linked to stimulus valence through an early analysis of whether the input conforms to the rules of Western music or not. That is, irregular, unexpected or even incorrect musical events may at the same time elicit a negative impression of the stimulus material presented that, in turn, provides the basis for subsequent evaluative categorization processes.

Koelsch, Kilches, Steinbeis, and Schelinski (2008) found an ERAN to unexpected and very unexpected chords embedded in piano sonatas. The ERAN was not affected by the emotional content of the used music suggesting that music-syntactic processing is independent of affective processing. Notwithstanding, Brattico et al. (2010) found an early right negativity – reminiscent of the ERAN – in response to incorrectness and disliking

judgments of musical chord sequences ending with congruous, ambiguous, or harmonically inappropriate chords. This result suggests that evaluative processing may modulate the early negative component. Although the findings are yet inconclusive about the link between evaluative and correctness processing of music, they provide first evidence for an early effect of evaluative impression formation also in music.

1.3.4 Affective stimulus processing: Late positive potential (LPP)

The LPP has been repeatedly associated with the processing of affective stimuli such as affective pictures (Schupp, Junghöfer, Weike, & Hamm, 2003), attractive faces (Werheid et al., 2007), voices (Spreckelmeyer, Kutas, Urbach, Altenmüller, & Münte, 2006), and music (Brattico et al., 2010; Müller et al., 2010). The LPP occurs between 400 and 1000 ms after stimulus onset and typically shows a posteriorly maximal distribution. The LPP is a sustained response that might last until stimulus offset and therefore is believed to reflect a phasic increase in attention to motivating stimuli (Hajcak, MacNamara, & Olvet, 2010). The amplitude of the LPP may be modulated by the intensity of the presented emotion, hence, by stimulus arousal (Schupp et al., 2003). It was also found that valence might affect LPP amplitude (Werheid et al., 2007). Hajcak et al. (2010) advocate that the LPP is generated by the activation of the norepinephrine system whose neurons extend to brain areas that are known to be part of the reward circuitry such as the amygdala and cingulate gyrus. Although it is commonly assumed that the LPP requires conscious stimulus processing (Hajcak et al., 2010), Frühholz, Jellinghaus, and Herrmann (2011) showed that the LPP occurred also during implicit processing of attractive faces.

The LPP seems to be a domain-general response reflecting evaluative processing. Apart from visual and auditory stimuli it has been obtained in response to affective hand gestures (Flaisch, Hacker, Renner, & Schupp, 2011) and during evaluative processing of attitudes (Cacioppo, Crites, & Gardner, 1996). As suggested by Jacobsen and Höfel (2003), the LPP might reflect the evaluative categorization of stimuli which after the initial impression formation as reflected in the early negativity constitutes the second stage of evaluative processing.

In music, the LPP was found in response to liking judgment tasks of chord sequences (Brattico et al., 2010). Müller et al. (2010) found an enhanced LPP amplitude for aesthetic judgments compared to cognitive judgments of the same musical chord sequences as used

by Brattico et al. (2010). Interestingly, the difference between the types of judgments occurred only in laymen but not in music experts. The findings suggest that laymen relied more on their internal affective system when making beautiful judgments than music experts.

2 Aims of the thesis

The present thesis investigated cognitive and neural determinants of music appreciation and music aesthetics using both behavioural paradigms as well as the event-related potential technique.

Study I aimed to contribute to an understanding of the nature of aesthetic experiences by defining the content and structure of the mental concept underlying the term aesthetics. To this end, adjectives describing the aesthetics of music were collected from over 250 participants (music experts and laymen) by means of a free association task.

Study II aimed to describe the developmental stage of school-aged children regarding music emotions, preference judgments, and the concept of beauty. To this aim, 6–9-year old children listened to two piano tunes composed in major and minor mode respectively, and a free tonal version of the same melody. Using a child-adapted rating scale, they judged each tune regarding its emotional content, its beauty, and according to their preference.

The goal of **Study III** was to unravel whether neural processes underlying evaluative and cognitive judgments of music would be modulated by participants' long-term music preference. Latin-American music fans and heavy metal music fans listened to excerpts of both genres while their ERPs were recorded for a liking judgment and a genre classification task.

Study IV investigated whether expressive timing facilitates the neural processing of phrase boundaries in music. ERP responses were recorded while participants listened to short atonal two-phrased melodies in a passive and an active condition in which the phrase boundary served as detection target. The melodies were presented as deadpan versions with isochronous tone onsets and as expressive versions containing timing fluctuations emphasizing the two-phrase structure of the melodies.

3 Methods

3.1 Behavioural experiments

3.1.1 Participants

In **Study I**, the participants were young adults aged from 19 to 29 years. The sample consisted of 126 music experts and 162 laymen. Professional and semi-professional musicians were categorized as music experts, whereas beginners, amateur musicians and participants without any musical training were assigned to the group of laymen. At the time of the experiment, music experts had accomplished on average 14.0 ($SD = 5.47$) years of instrument lessons and/or 6.0 years of singing lessons ($SD = 6.09$). Laymen had had musical training in instrument playing or singing on average for 4.2 years ($SD = 4.49$) and 2.4 years, ($SD = 3.72$) respectively. Participants were Finnish and Swedish native speakers recruited from the University of Helsinki, University of Jyväskylä, University of Turku, and Sibelius Academy in Helsinki.

In **Study II**, 127 school-aged children participated. They were aged between 6 and 9 years at the time of the experiment. The children were recruited from elementary schools in the Helsinki area. The data analysis was performed on 78 children because of an otherwise unbalanced distribution of the background variable *music education*¹. For the data analysis, the children were assigned to two different age groups: 6–7-year-old and 8–9-year-old children.

3.1.2 Stimuli

In **Study I**, no stimuli were presented to avoid any bias introduced by experimenter-selected music. For **Study II**, three versions of one short piano tune were composed for the experiment. The first version was composed in major mode, the second version in minor mode, and the third tune lacked any tonal centre and was therefore classified as free tonal (Figure 2).

¹ In Study II, 'music education' refers to musical training outside the school curriculum.

A.



B.



C.



Figure 2. Scores of the piano tunes used in Study II. A. The melody in major mode. B. The melody in minor mode. C. The free tonal version of the melody. The melodies can be listened to at www.eva-istok.de.

The three melodies consisted of simple chords and simple melodic lines and, hence, were similar to prototypical children's songs. Each tune lasted approximately 20 seconds. The tempo was held constant across the different versions and the melodies were lacking any emotional expression.

3.1.3 Experimental setup and data analysis

3.1.3.1 Free verbal association task

A free verbal association task was used to collect adjectives describing the aesthetics of music in **Study I**. The participants were instructed as follows: "Please write down as fast as possible adjectives that, according to your knowledge, might best describe the aesthetic value of a musical piece. You can choose as many adjectives as you like". The task had to be accomplished within 5 minutes. Written instructions were given in Finnish and participants were asked to give their answers in their native tongue (Finnish, and in rare cases, in Swedish).

The data were collected during several lectures at the different Universities (see participants). All participants in a given group performed the task at the same time and afterwards filled in a questionnaire regarding demographic information in general and their musical background in particular.

To handle the huge amount of adjectives that were collected, several data reduction procedures were applied. The final analysis was based on 43 adjectives that were named by at least 5 % of the participants. For these remaining adjectives, the frequency of occurrence and the cognitive salience index (Sutrop, 2001) were calculated. In addition, valence ratings for each of the adjectives were obtained from several independent judges.

Logistic regression was used to test the effects of age, gender, and musical expertise on the choice of adjectives. Finally, multidimensional scaling (nonmetric MDS, PROXSCAL) was applied to visualize the content and structure of the cognitive concept underlying aesthetic responses to music. To this aim, the co-occurrence of adjectives within one list was calculated. The co-occurrence matrix served as input for MDS which is generally used to extract underlying dimensions based on similarities or dissimilarities of items.

3.1.3.2 Child-adapted rating scales

In **Study II**, children were asked to listen to the three music pieces and to rate them according to whether they liked them or not, according to the emotion (happiness, sadness) expressed by the music, and according to whether the piece sounded beautiful or ugly. To quantify the aesthetic responses of the participants, special child-adapted rating scales were developed. A picture of seven adjacent circles of increasing size was used for the quantification of preference ratings. The children were told that the more they like the music the larger should be the circle they choose. For the emotion and aesthetic ratings, drawings of a laughing sun (happiness), a crying cloud (sadness), a princess (beauty), and a witch (ugliness) had to be attached to a separate sheet showing a note symbol and four lines with seven squares extending from the note symbol (Figure 3). It was explained that the closer the drawing would be attached to the note symbol, the more adequate was the attribute depicted in the drawing (happiness, sadness, beauty, ugliness) for a characterization of the music piece.

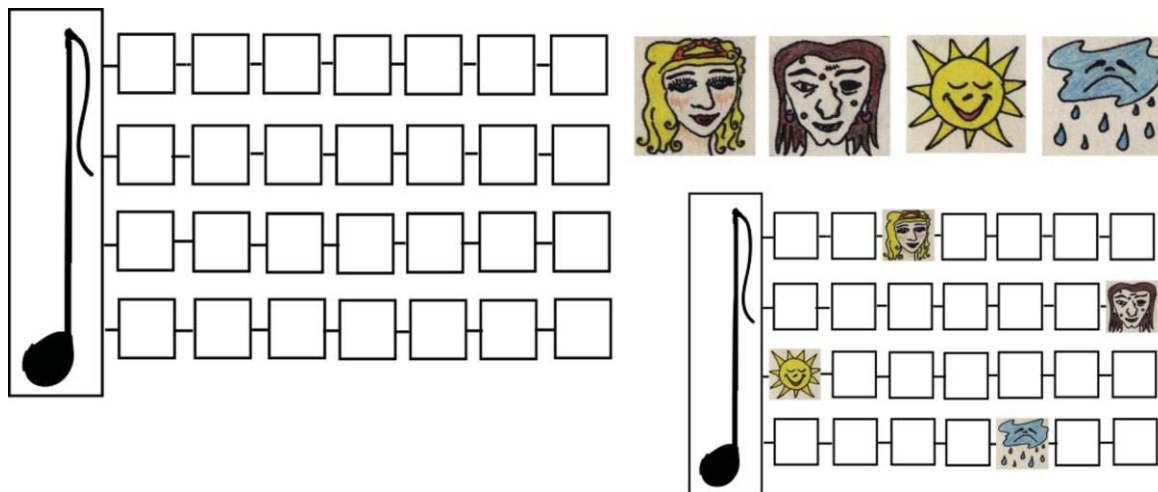


Figure 3. Child-adapted rating scale used for preference, emotion, and aesthetic ratings of the three piano tunes.

The data were collected during school lessons. Instructions were given in Finnish. All pupils of a class performed the experiment at the same time. Before the actual experiment, the children completed a simplified questionnaire asking for their age, their gender, and whether they were playing any instrument. Thereafter each of the pieces was played twice. After the first presentation, children had to invent a little title for the tune and to indicate whether and how much they liked it. Then the melody was played again and the children

had to perform the emotion and aesthetic ratings. The procedure was repeated for each of the tunes, whereby the order of presentations varied between different school classes.

The data collected by means of the rating scales were treated as interval data and analysed with parametric statistical test. Specifically, a mixed analysis of variance (ANOVA; repeated measures) was conducted for the preference, emotion, and aesthetic ratings separately. The following factors were entered: Mode (major, minor, free tonal), Age group (6–7-year-old children, 8–9-year-old children), Gender (girls, boys), and Music education (children with music education, children without music education).

3.2 ERP recordings

3.2.1 Participants

The participants of **Study III** were selected based on their preference for Latin-American music and heavy metal music respectively. In total, 34 volunteers participated in the experiment of which seven were excluded from data analysis because they either did not show a clear preference profile for each of the genres or because the quality of ERP-recordings was poor. The data analysis was performed for 14 LAm-fans (mean age 29.9 years, $SD = 5.8$; 8 females) and 13 HMm-fans (mean age 26.3 years, $SD = 5.4$; 7 females).

In **Study IV**, 10 musicians (professional and semi-professional, mean age = 25.7 years; $SD = 3.40$) and 10 laymen (mean age = 24.0 years; $SD = 3.24$) participated. In each group, half of the participants were female. The groups were recruited in the Helsinki metropolitan area.

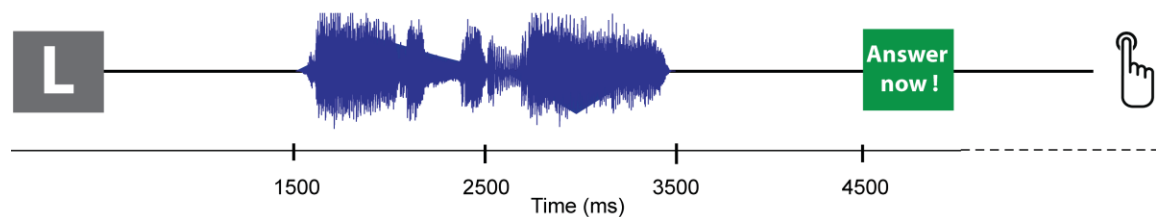
All volunteers participating in **Study III** and **IV** were Finnish native speakers. All of them indicated to have normal hearing. None of them reported on neurological problems and none took medication affecting the central nervous system. The studies were approved by the ethical committee of the former Department of Psychology, University of Helsinki.

3.2.2 Stimuli and experimental procedures

The stimulus set² in **Study III** consisted of short excerpts (2 seconds) of audio recordings of Latin-American music (LA-music) and heavy metal music (HM-music). An additional set of electronic music was presented but not analysed. The excerpts were chosen based on mean liking and mean familiarity ratings in a behavioural experiment. Only unfamiliar (mean familiarity rating < 2.5) excerpts were included in the ERP measurement when the respective mean liking ratings significantly differed between LAm-fans and HMm-fans.

Each music excerpt was presented twice and participants performed a liking judgment task and a genre classification task on each of the excerpts. Genre and task were fully crossed and stimulus presentation was randomized. To avoid contamination of the EEG signal with motor preparation responses, a delayed response design was used (Figure 4).

Liking task: Do you like the following piece ?



Genre classification task: Is the following piece electronic music ?

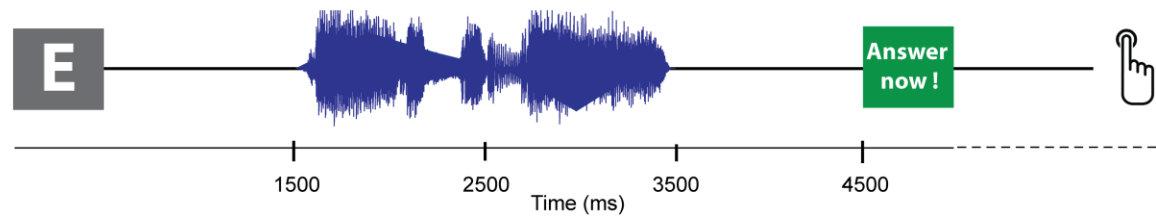


Figure 4. Experimental design of Study III. The task (L = liking task; E = genre classification task) was prompted before the music excerpt was presented. When the music stopped another cue signaling the allowance to press the answer button appeared on the screen.

In **Study IV**, simple scale-like melodies³ served as stimulus material. They were based on minor seconds, major seconds, or major thirds and presented in either ascending or descending order. Half of the melodies contained a pitch leap larger than the basic interval of the melodies indicating the beginning of a new phrase, hence, the phrase boundary. The

² Audio examples can be found at www.eva-istok.de

³ Examples of the melodies can be found at www.eva-istok.de

size and position of the pitch leap varied between melodies to vary the difficulty level and to avoid that participants could predict the grouping structure.

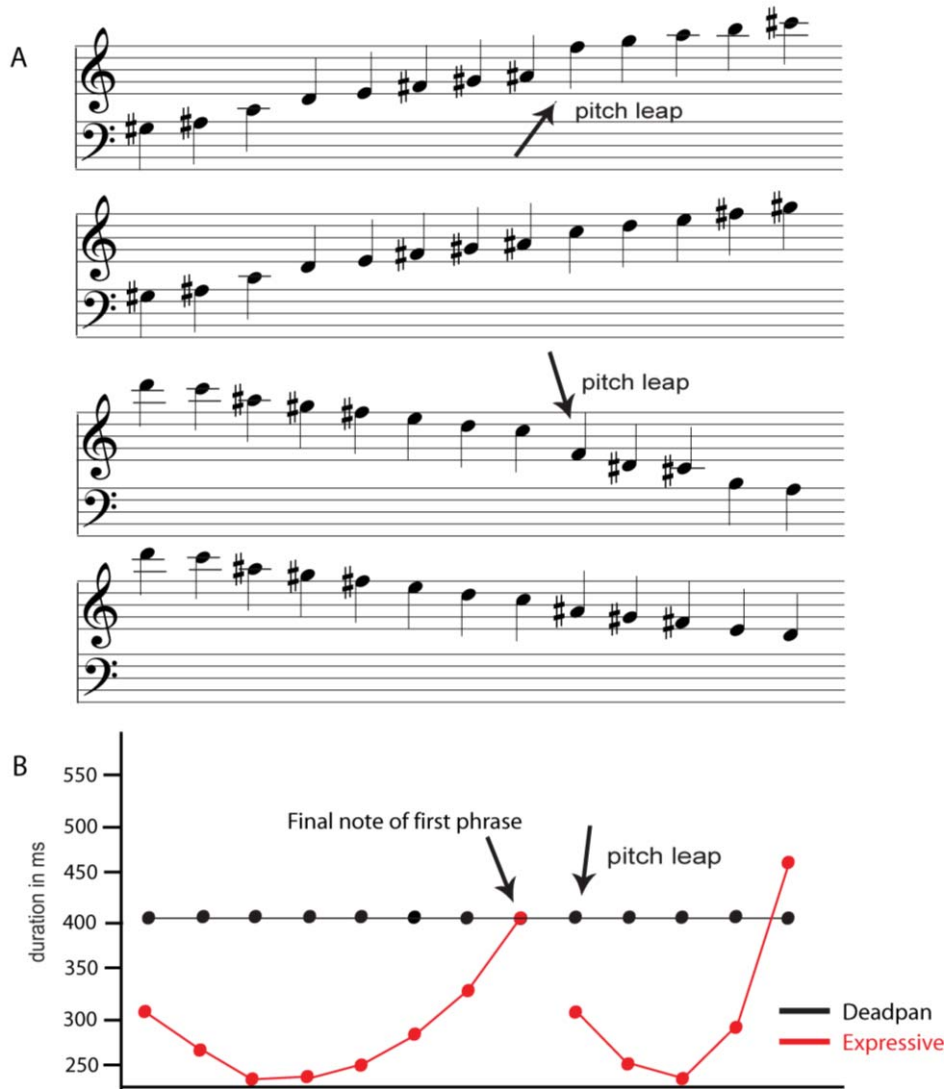


Figure 5. Melody examples of Study IV. A. Ascending and descending melodies consisting of major seconds with and without pitch leap. B. Duration (in ms) of each individual note for deadpan and expressive versions of the same melody.

The other half of the melodies was presented without an additional pitch leap. Both melodies with and without pitch leap were presented in two versions: a) as deadpan melodies with isochronous note onsets and b) as expressive melodies containing subtle timing variations that highlighted the melodies' two-phrase structure.

Importantly, the expressive versions were adjusted in tempo so that the length of the final note in the first phrase was the same as in the corresponding deadpan version (see Figure 5B). The melodies were generated by the software programme Director Musices developed at KTH, Sweden (Friberg, Colombo, Frydén, & Sundberg, 2000).

The experiment consisted of two different sessions. During the first session, participants passively listened to the melodies while they were watching a silent movie of their choice. During the second session, participants were asked to actively listen to the melodies and to detect the pitch leap. Between the two sessions was a break of at least one week.

In both **Study III** and **IV**, the stimulus material was presented via headphones (Sony MDR-7506 Professional) whereby the sound level was adjusted to 40 dB above the individual subject's hearing threshold. The software programme Presentation (version 12.1 04. 10.08; Neurobehavioural Systems, Inc., Albany, CA, USA) was used for stimulus presentation.

3.2.3 Data acquisition and analysis

In **Study III** and **Study IV**, the EEG was recorded with 128 and 64 Ag/AgCl electrodes (Active Two system, Biosemi) respectively. Electrodes were mounted in nylon caps. The reference electrode was placed on the nose. The horizontal electro-oculogram (EOG) was recorded with two electrodes placed at the outer canthi of the left and right eye, whereas the vertical EOG was recorded with two electrodes attached above and below the left eye.

The continuous EEG recordings were filtered offline between 0.5 and 40 Hz using finite impulse response filtering and afterwards down-sampled to 256 Hz. Missing channels were interpolated. Independent component analysis (ICA) was used for artefact removal. In **Study III**, the EEG signal was divided into epochs of 2000 ms length starting with the onset of each genre excerpt and an additional pre-stimulus baseline of 200 ms. In **Study IV**, ERP responses were recorded to the beginning of the second phrase, hence, to the phrase boundary. The continuous EEG recordings were therefore divided into epochs starting 100 ms before and 600 ms after phrase boundary onset. In **Study III** and **IV**, epochs that contained artefacts exceeding a threshold of $\pm 100 \mu\text{V}$ were automatically rejected.

In Study **III**, grand-averages were computed for four different conditions: a) LA-music; liking judgment task, b) LA-music; genre classification task, c) HM-music; liking judgment task, d) HM-music; genre classification task. Mean amplitudes were calculated for three different time windows: 100-200 ms (N1), 230-370 ms (early negativity), and 600-900 ms (LPP). A mixed analysis of variance (repeated measures ANOVA) was used for statistical analysis including the factors ‘Genre’ (LA-music, HM-music), ‘Task’ (liking judgment task, genre classification task), ‘Group’ (LAm-fans, HMm-fans), ‘Anterior-posterior distribution’ (F-line [D5, C26, C21, C13, C5], C-line [D20, D14, A1, B20, B23], P-line [D29, A18, A19, A31], O-line [A11, A15, A23, A26, B8]), and ‘Laterality’ (left [D5, D20, D29, A11], middle-left [C26, D14, A18, A15], midline [C21, A1, A19, A23], middle-right [C13, B20, A31, A28], and right [C5, B23, B13, B8]).

In **Study IV**, grand-averages were calculated for the following conditions: a) Deadpan melody with pitch leap, b) Expressive melody with pitch leap, c) Deadpan melody without pitch leap, d) Expressive melody without pitch leap. The mean amplitudes of ERP responses were calculated for two time windows for both the passive and active recordings: 140-220 ms (P2/N2b) and 300-400 ms (CPS/P3a). Four different regions of interest (ROI) were defined for statistical analysis: left fronto-central [FC1, FC3, C1, C3, C5], right fronto-central [FC2, FC4, C2, C4, C6], left centro-parietal [CP1, CP3, CP5, P1, P3], and right centro-parietal [CP2, CP4, CP6, P2, P4]. A mixed ANOVA (repeated measures) was calculated with the factors ‘Expression’ (deadpan, expressive), ‘Leap’ (pitch leap, no pitch leap), ‘ROI’ (left fronto-central, right fronto-central, left centro-parietal, right centro-parietal), and ‘Group’ (musicians, laymen) for the passive and active session separately.

Whenever Mauchly’s test of sphericity was significant, Epsilon corrected degrees of freedom were used. Effect size was estimated using Eta-squared. Pairwise comparisons were based on simple main effects analysis and Bonferroni corrected. Effects sizes of pairwise comparisons were determined using Cohen’s *d*.

The analysis of the co-occurrence of adjectives revealed that *beautiful* and *touching* were most often listed together (Figure 6) supporting the notion that music emotions can be best described as being moved or touched by music (Konečni, 2008).

Differences regarding the conceptual space underlying aesthetic responses to music were found between musicians and laymen and male and female participants. Surprisingly, laymen rather than musicians used music-specific vocabulary referring to harmony and rhythm. It is possible that music experts did not produce such vocabularies because melody, harmony, and rhythm do represent basic analytical concepts of music theory and composition that are not seen to be adequate to describe the aesthetics of music. Laymen also produced more adjectives related to mood and mood regulation than music experts. This finding suggests that there may be differences in the functional use of music between the two groups. It is possible that laymen more frequently use music to regulate their emotional states than music experts. Musicians frequently named adjectives such as *varying* and *original*. This suggests that the stimulating character of music, its novelty and originality, seem to be of greater importance for music experts.

Gender differences were obvious for the use of the adjective *beautiful*. Male participants listed *beautiful* significantly less often than female participants, though *beautiful* still was the most frequently named term named by males. There was a tendency for male participants to use more negatively loaded adjectives than females.

Generally, the results of **Study I** suggest that there is a need to consider aesthetic judgments, as often also required during experimental tasks, as an affective and not purely cognitive response to music. Future studies could benefit from an inclusion of psychophysiological measures to complement subjective ratings of beauty with more objective data.

4.2 Study II: Aesthetic judgments in school-aged children

In **Study II**, the analysis of the preference, emotion, and aesthetic ratings of three simple piano tunes revealed that the melody in major mode received highest preference scores overall. In relation to the other tunes, this tune was preferred to the one composed in minor mode by 8–9-year-old children and by 6–7-year-old children without specific music training.

The preference ratings obtained in response to the three different pieces (Figure 7) suggest that the children based their judgments on the affective content of the music rather than in terms of the pieces' conformity to the rules of Western tonal music. If the children had based their preference judgments on the correctness of the tunes, the free tonal piece should have been least liked. In neither group, however, did the difference between preference ratings for the major piece and the free tonal version reach significance. It is therefore likely that the minor piece received the lowest preference scores because it conveyed sadness, a music emotion that is preferably avoided in experimental situations (Hunter & Schellenberg, 2010).

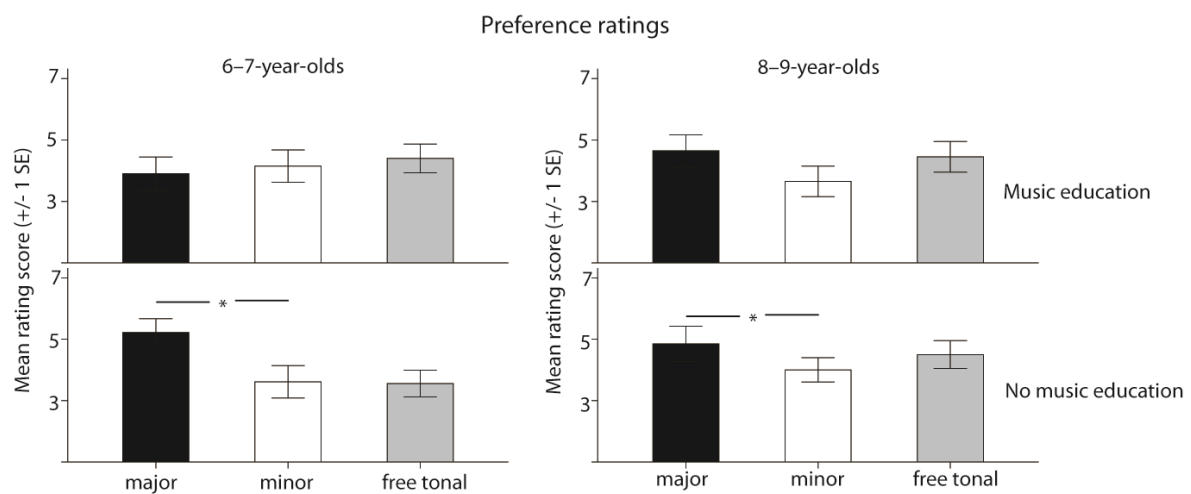


Figure 7. The bars show the mean preference rating scores (± 1 SE) of the three piano tunes used in Study II separately for 6–7-year-olds and 8–9-year-olds with and without music education.

The major piece obtained highest happiness scores overall. Except of 6–7-year-olds without music training, all children rated this tune to be happier than the minor and, except of the 8–9-year-old children without musical training, also happier than the free tonal versions (Figure 8). It is possible that music education accelerates the development of understanding music emotions based on mode differences in early school-age years. Nevertheless, as the data show, the connotation between happiness and the major mode may also develop by passive exposure to Western tonal music. That is, because there is a prevalence of children's songs composed in major mode (Cohen, Thorpe, & Trehub, 1987; Kastner & Crowder, 1990), familiarity might account for the earlier association between major mode and happiness.

Only the 8–9-year-old children found the minor tune to be sadder than the melody in major. The sadness ratings of the younger group of children did differ between the music pieces. This finding is in line with previous studies (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Gerardi & Gerken, 1995) suggesting that the association between major mode and happiness emerges earlier in development than the association between minor mode and sadness.

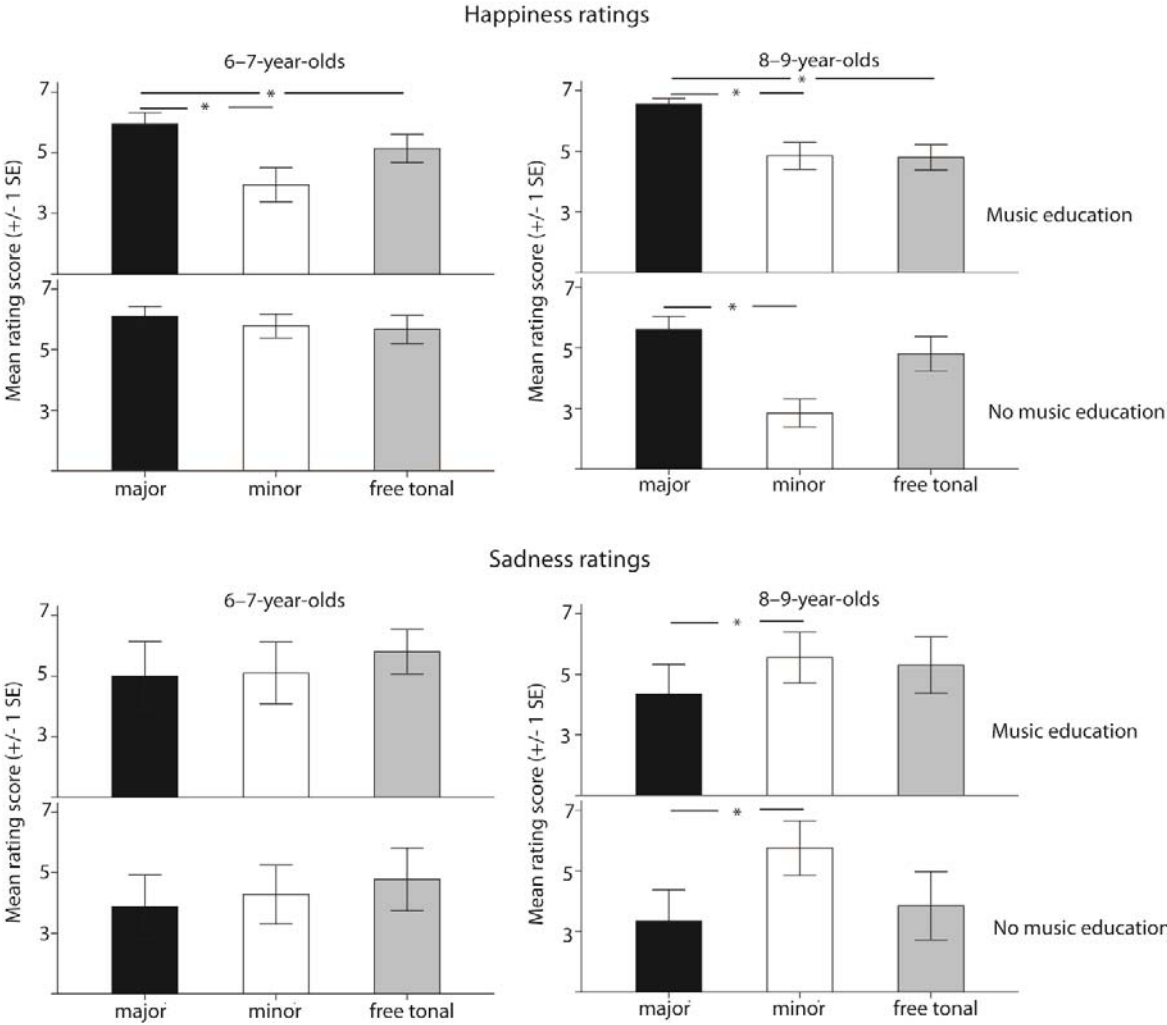


Figure 8. The bars show the mean happiness and sadness rating scores (+/- 1 SE) of the three piano tunes used in Study II separately for 6–7-year-olds and 8–9-year-olds with and without music education.

The group of 8–9-year-olds found the melody in major mode more beautiful than the one in minor mode and the free tonal one, whereas the younger group of children did not differentiate between the melodies in terms of their beauty (Figure 9). This finding implies

that beauty as aesthetic category for music may not be available before the age of eight years.

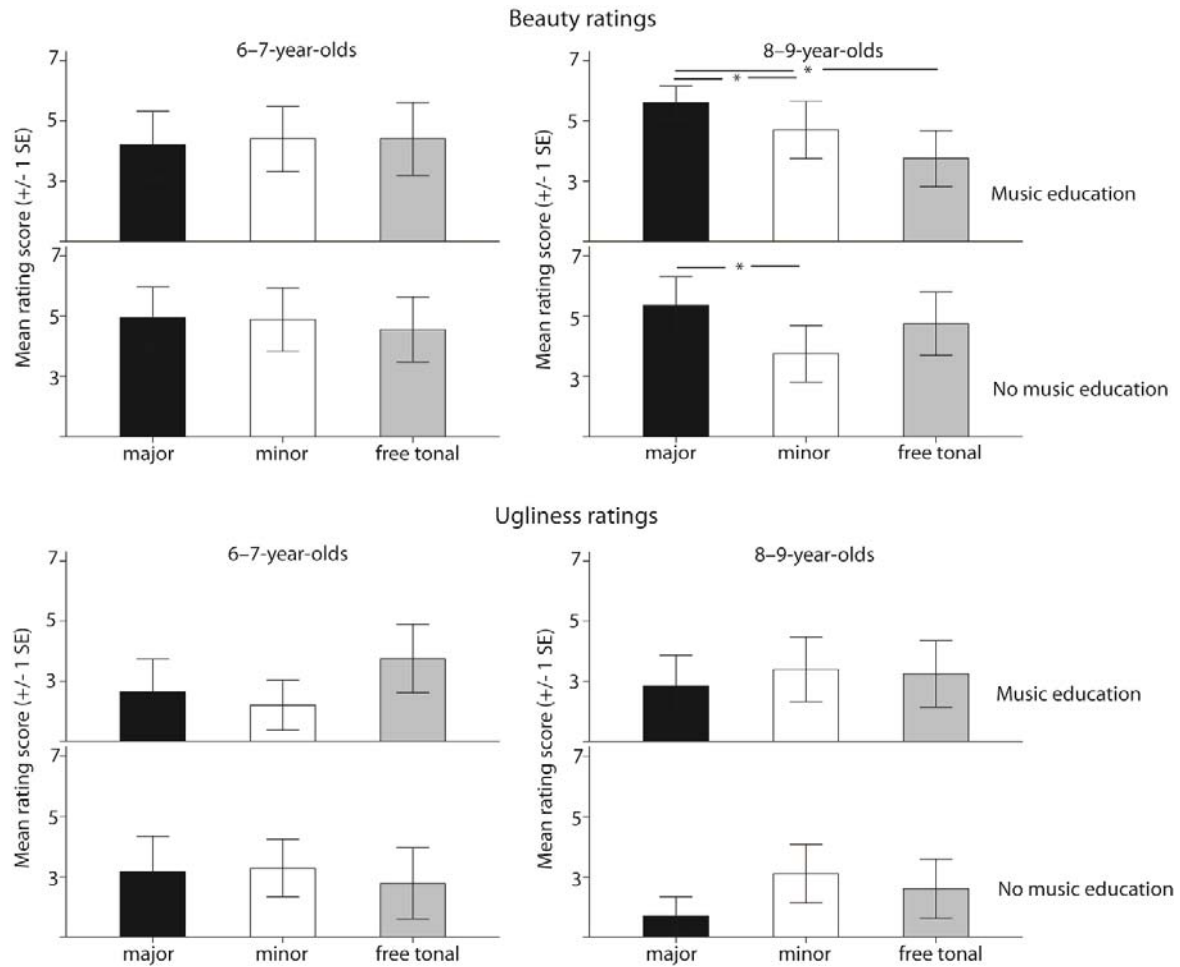


Figure 9. The bars show the mean beauty and ugliness rating scores (+/- 1 SE) of the three piano tunes used in Study II separately for 6-7-year olds and 8-9-year-olds with and without music education.

The children did not distinguish between the three different melodies in terms of ugliness, although the major tune was rated to be least ugly by the 8-9-year old children. Generally, the ugliness ratings were comparably low. This finding cannot be attributed to a possible avoidance of negative words and associated emotions because children seemed to naturally judge music in regard to sadness. It rather suggests that ugliness may not be an adequate term to describe music in childhood.

4.3 Study III: Music genre preference and evaluative judgments

Study III shows that attentive listening to both preferred and non-preferred music may automatically initiate affective responses in listeners as indexed by the occurrence of an LPP between 600 and 900 ms after stimulus onset in all conditions. In addition, the valence of the presented music may affect early stimulus processing.

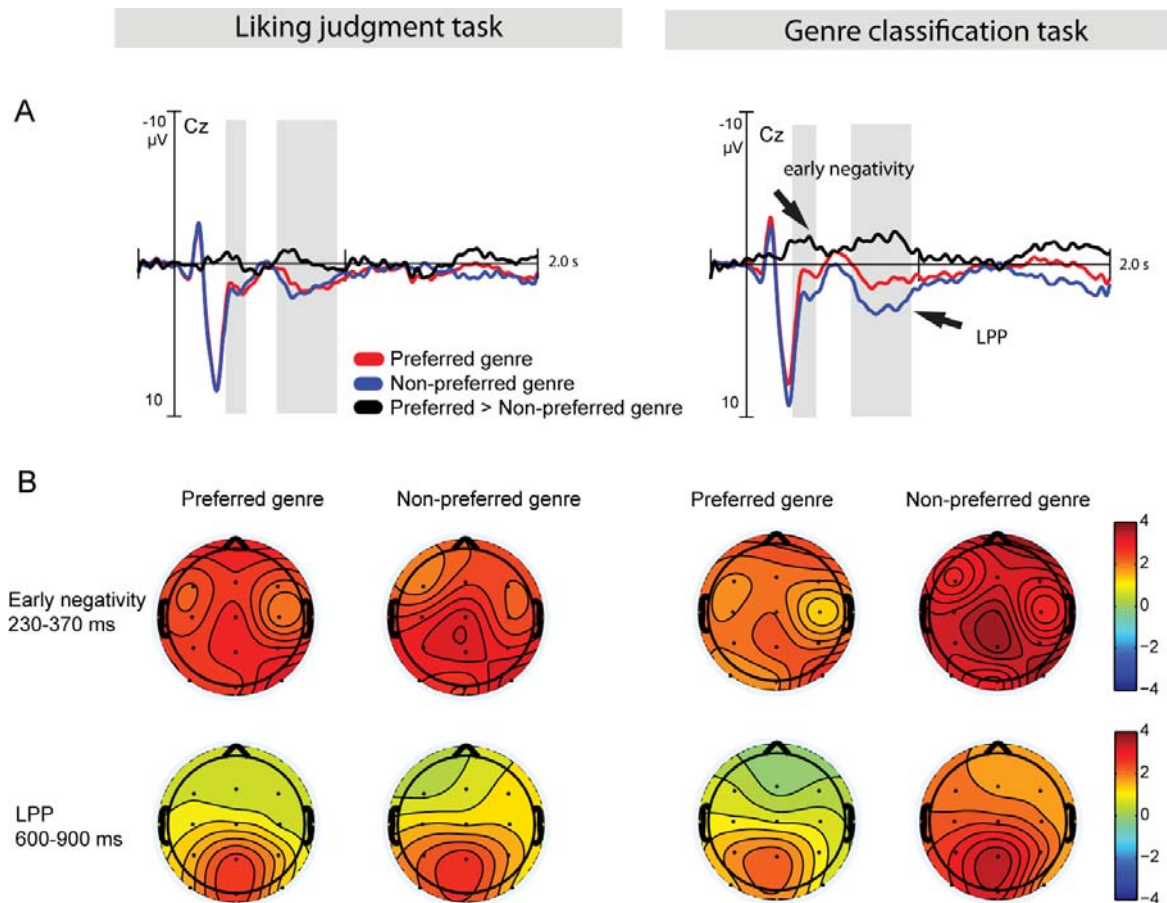


Figure 10. Grand-average ERPs at Cz (A) and voltage maps (B) in response to preferred and non-preferred music for the liking judgment task (left panel) and the genre classification task (right panel) separately. An LPP was elicited in all conditions. Valence modulated the early negativity and the LPP during the genre classification task.

The preferred compared to the non-preferred genre elicited a larger early negativity between 230 and 370 ms after stimulus onset. The non-preferred genre was associated with an increased LPP amplitude suggesting that the processing of non-preferred music might have resulted in the recruitment of additional neural resources needed for the evaluation of less familiar music. Surprisingly, valence affected music processing only during the genre classification task.

Several aspects might account for this finding. First, it may be that each music excerpt might have triggered affective responses at different time points during music presentation. Therefore, it is possible that by averaging ERP responses time-locked to the beginning of each excerpt, time-varying affective responses could not been captured. In this case, differences between the preferred and non-preferred genre, however, should not have occurred during the genre classification task either.

Second, familiarity instead of preference might account for the result. Since those two aspects are so intimately connected, the experimental design of **Study III** does not allow disentangling preference from familiarity. Notwithstanding, according to the theory of processing fluency developed by Reber et al. (2004), familiarity to a great deal determines how fluent stimuli may be processed. That is, the more familiar a stimulus is, the more fluent is its processing. As a result, fluent stimulus processing leads to aesthetic pleasure. Thus, even if familiarity rather than music preference was the main factor contributing to the results obtained in **Study III**, it is likely that listening to very well known genres elicited affective responses in the participants because of the ease or difficulty with which LAm-fans and HMm-fans could process the respective genre.

Most likely, effortful liking judgments may have obscured automatic evaluative judgments. By drawing attention to internal affective responses, they become conscious, are no longer involuntary and even might be controlled. Hence, automatic evaluative responses as suggested by Zajonc (1980) may only be captured during implicit tasks.

LA-music elicited a larger N1 amplitude than HM-music (Figure 11). An analysis of timbre qualities of the stimulus material using the MIRToolbox developed by Lartillot, Toivainen, and Eerola (2008) revealed differences in brightness and roughness between the two genres that most probably account for the difference in N1 amplitude. In contrast to previous studies where the N1 was elicited in response to a single musical chord (Regnault et al., 2001) or interval (Schön et al., 2005), the N1 elicited in our experiment reflects the average response within the time window from 100 to 200 ms after stimulus onset of a variety of different audio recordings.

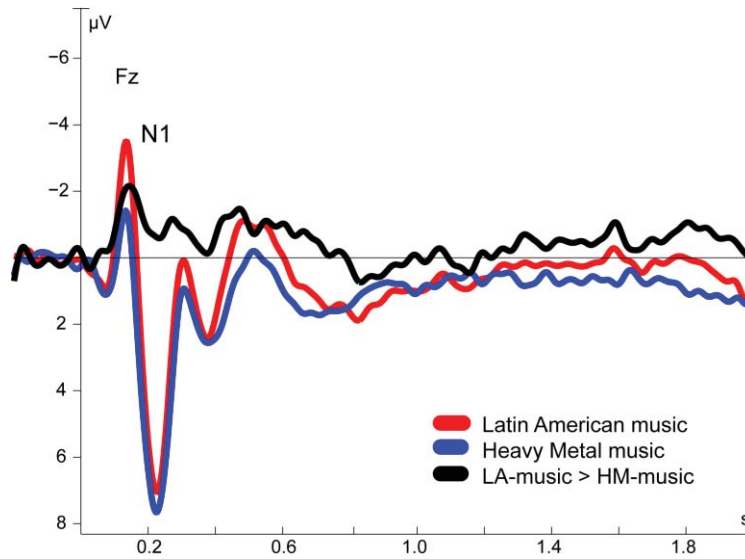


Figure 11. Grand-average ERPs at Fz elicited by LA-music and HM-music. The N1 amplitude was enhanced for LA-music.

We therefore speculate that the N1 effect is due to a combination of timbre properties (roughness, brightness) and the concomitant pleasant or unpleasant sensation caused by the respective music excerpt.

4.4 Study IV: Expressive timing and phrase boundary processing

In **Study IV**, a clear fronto-central P2 component was elicited in response to the phrase boundary for expressive and deadpan melodies during the passive listening session (Figure 12). The response was elicited independently of whether a pitch leap additionally marked the phrase boundary or not. This finding suggests that, during passive exposure to phrases, the P2 reflects local stimulus feature encoding that is independent of the timing structure of the preceding context. Interestingly, a P3a-like response occurred when both expressive timing and the pitch leap together marked the phrase boundary. This response was completely absent for expressively played melodies without a pitch leap indicating that the processing of the phrase boundaries was affected by attention allocation mechanisms that may especially serve the processing of ‘novel’ stimuli.

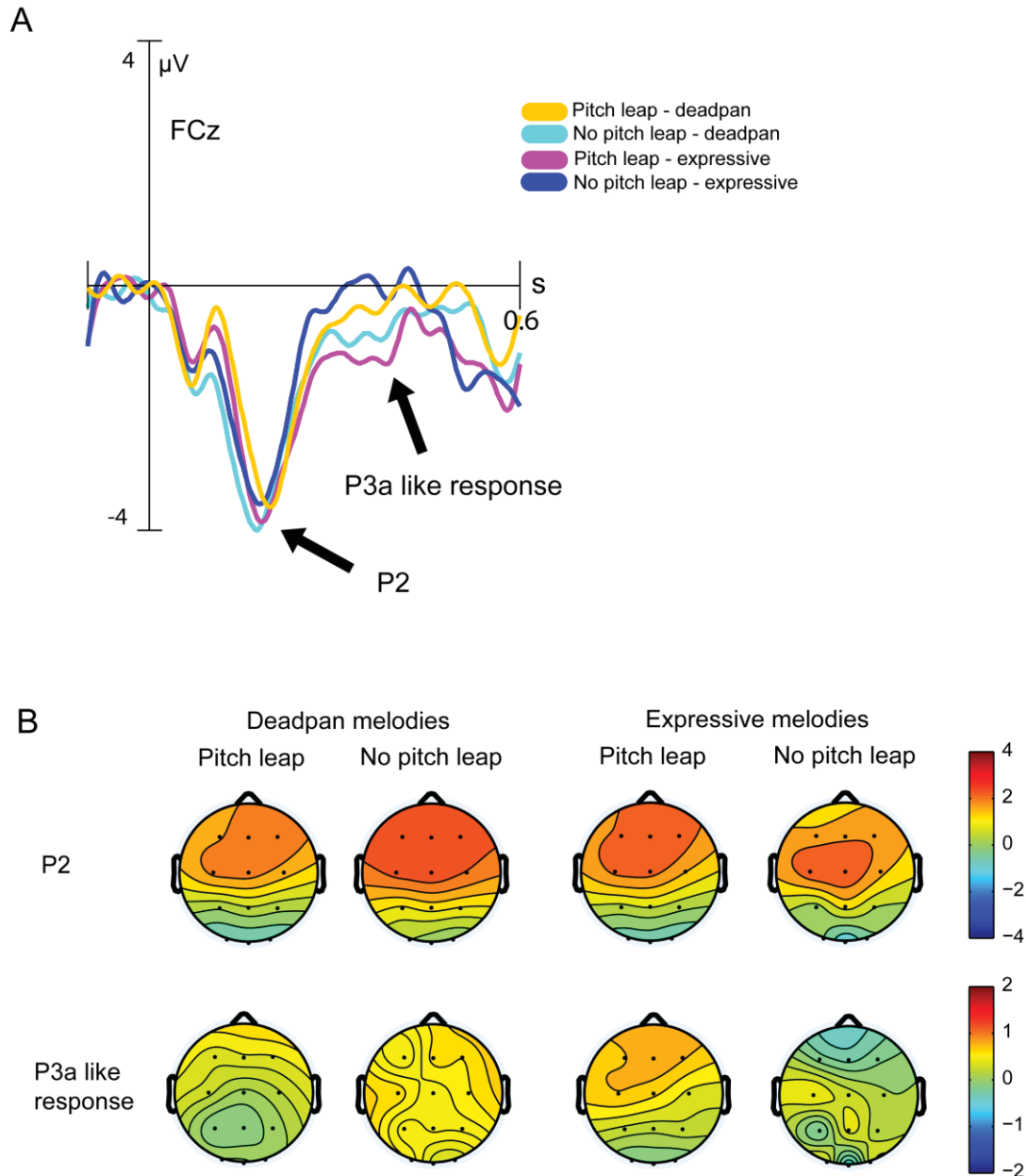


Figure 12. A. Grand-average ERPs at FCz in response to the phrase boundary during the passive listening session. B. Voltage maps of the P2 and P3a-like response.

During active phrase boundary processing (Figure 13), a clear N2b/P3a complex was elicited for target phrase boundaries, whereas the processing of non-target boundaries was marked by the occurrence of the P2 component. The P2 amplitude was larger for expressive melodies compared to deadpan melodies suggesting that the P2 amplitude is sensitive to top-down mechanism during attentive phrase boundary processing. Specifically, the extraction of global timing information seems to affect stimulus identification. In previous studies, the P2 amplitude was reduced for phrase boundaries

preceded by a pause (Knösche et al., 2005; Neuhaus et al., 2006) suggesting that the effect was due to neural recovery. In **Study IV**, such an effect can be excluded because the inter-onset interval (IOI) preceding the phrase boundary was identical for expressive and deadpan melodies.

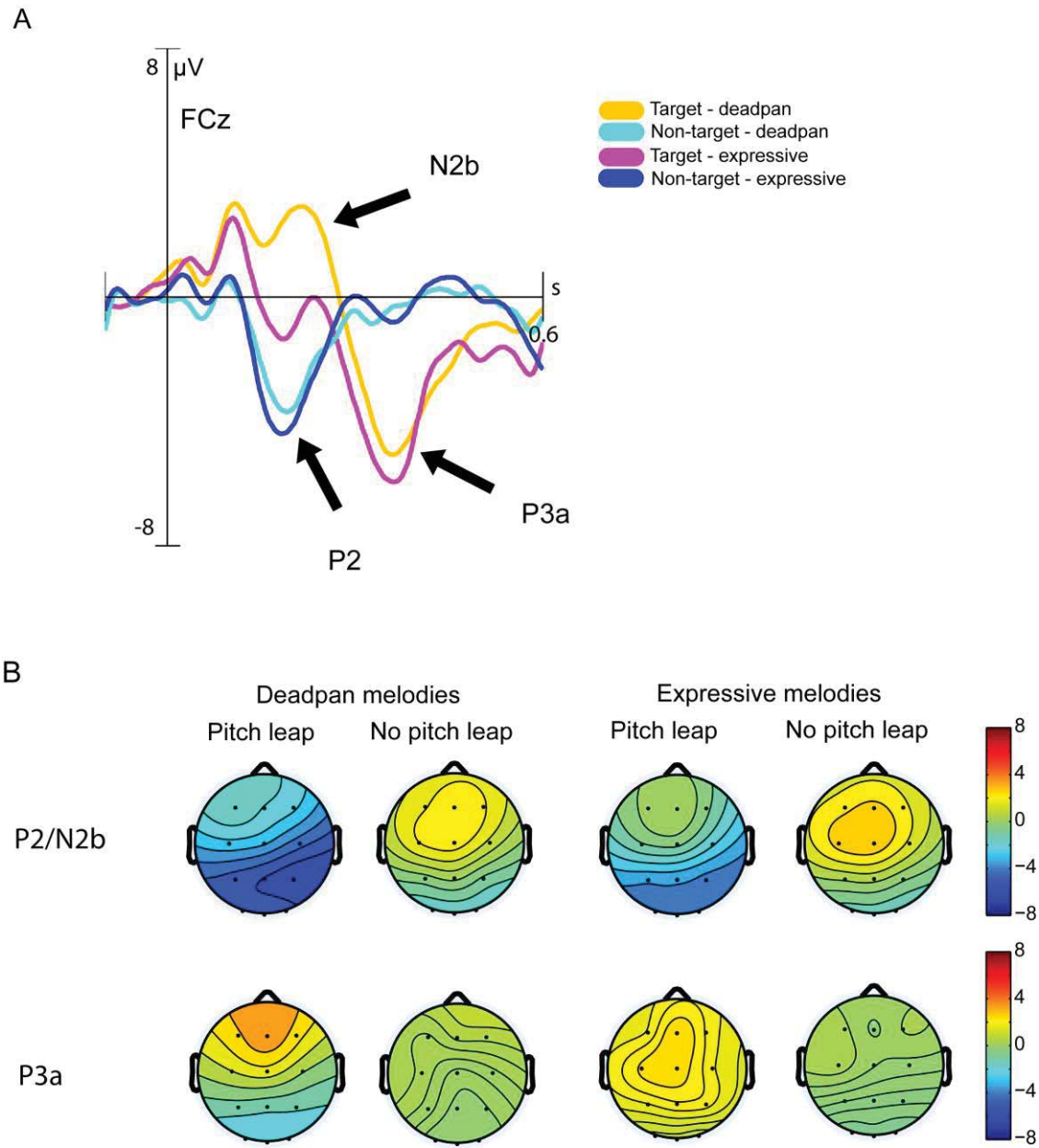


Figure 13. A. Grand-average ERPs at FCz in response to the phrase boundary during the active detection task. B. Voltage maps of the P2/N2b and P3a.

The N2b component was smaller for expressive compared to deadpan melodies. Palmer and colleagues (2009) showed that the N2b amplitude was modulated by the position of

target events embedded in melodies. The later the target occurred, hence, the more expected the target was, the smaller was the N2b amplitude. This finding implies that the occurrence of the phrase boundary in **Study IV** was less unexpected when expressive timing indicated its occurrence. During deadpan melodies no such timing cues were available and, hence, the phrase boundary occurred unexpectedly as reflected in an increased N2b amplitude. In line, the P3a amplitude was significantly reduced in deadpan melodies suggesting that the recognition of phrase boundaries was more difficult than for expressive melodies.

Study IV further revealed differences between musicians and laymen for the active but not the passive processing of target phrase boundaries (Figure 14). In musicians, the N2b was characterized by shorter latency and reduced amplitude compared to laymen suggesting that musicians detected the pitch leap faster than laymen.

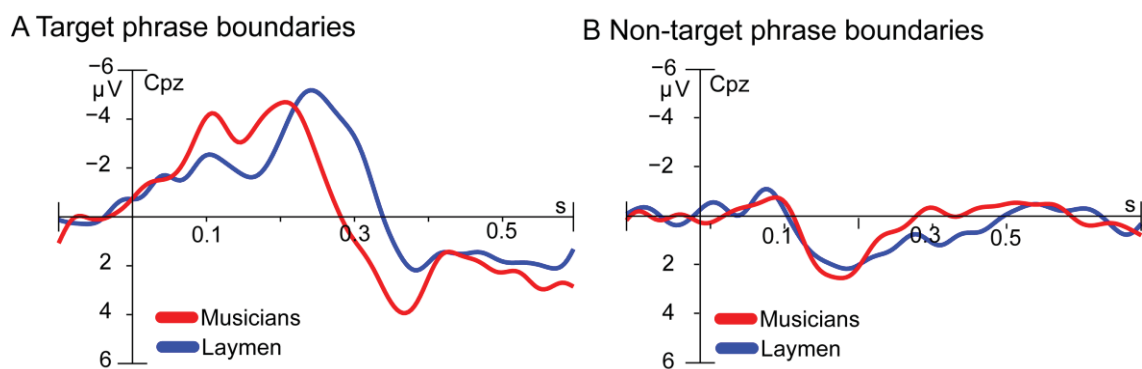


Figure 14. Grand-average ERPs at Cpz in response to the phrase boundary for target (A) and non-target (B) phrase boundaries for musicians and laymen separately.

The P3a amplitude was larger for musicians compared to laymen for deadpan melodies. This result shows that the detection of the target phrase boundary was more difficult for laymen than for music experts because reduced P3a amplitudes have been shown to reflect increasing task demands (Polich, 2007). The P3a amplitude did not differ between musicians and experts for expressive melodies suggesting that expressive timing facilitates the processing of subsequent auditory events by increasing attention resources towards the anticipated sound object especially when pitch discrimination abilities are not as precise as in trained musicians.

5 General discussion

5.1 The nature of aesthetic experiences in music

Since its beginnings in the 19th century, empirical aesthetics has been mainly concerned with the investigation of various stimulus features and their effects on aesthetic judgments of perceivers. The systematic description of the nature of aesthetic experiences from a psychological perspective has attracted considerably less attention. In his book *Aesthetics and Psychobiology* (1971), Daniel E. Berlyne provided an elaborate attempt to characterize human aesthetic experiences as dependent on the arousal a stimulus is able to elicit in the perceiver. In music research, the interest in emotions and their close connection with music has significantly increased during the last two decades and offered new important insights regarding affective reactions to music. The prevalent application of basic emotion models and dimensional models has been critically examined by several scholars in the field (e.g., Sloboda & Juslin, 2010). As a result, recent work (Konečni, 2008; Zentner et al., 2008) promotes a more nuanced description of emotions elicited by music.

In **Study I**, it was found that the everyday concept underlying aesthetic responses to music is centred on the experience of beauty which seems to be closely tied to the feeling of being touched by music. As such, beauty presumably does not refer to a purely cognitive and analytic assessment of the quality of an aesthetic object that is limited to a handful of experts. On the contrary, beauty is conceived an emotional experience in response to music. This assumption is further supported by the finding that many of the adjectives produced in **Study I** referred to emotions and affective states.

In contrast to the findings by Jacobsen et al. (2004) who found that *ugly* was the second most listed adjective describing the aesthetics of objects, participants listed ugly only seldom in **Study I**. Thus, the application of ugliness to describe music does not seem to be appropriate. This assumption is supported also by the results of **Study II** showing that the mean scores of ugliness ratings of music by school-aged children were generally very low for all music pieces. That is, the children participating in **Study II** were seemingly not able to judge music according to the concept of ugliness. Since the older children, however, showed a refined judgment pattern regarding the beauty of the music, it is likely that the concept of ugliness does not apply to music even in school-aged children.

Together the results of **Study I** and **Study II** suggest that the concept underlying aesthetic responses to music might differ from the concept domain of visual art. The frequent listing of music-specific vocabulary (e.g., *rhythmic*, *harmonic*, *melodious*) further emphasizes possible differences between content domains. This is an important finding also in regard to the applicability of general models of aesthetic processing to music. That is, although a general model of aesthetic processing is helpful for the formulation of hypotheses to be empirically tested, it might be warranted to further develop models that account for the domain-specific characteristics of paintings, music, architecture and the like.

5.2 Music preference and aesthetic judgments in young children

Study II demonstrates that school-aged children are able to express their subjective music preferences by the age of 8 to 9 years. Further, the data show that the connotation between basic emotions such as sadness and happiness and musical mode is formed at around the same age. Happiness seems to be associated with the major mode earlier in development than sadness with the minor mode. Finally, **Study II** reveals that school-aged children are able to judge music according to aesthetic criteria.

The results of **Study II** show that early music education may promote the ability to recognize emotions based on mode information. Although the piece in major overall received the highest happiness ratings, 6–7-year-olds without musical background did not distinguish the three pieces in regard to happiness. 8–9-year-olds without specific musical training clearly differentiated the emotional tone of the major and the minor tune suggesting that also passive exposure alone aids the recognition of basic emotions in music. Without musical training, the ability to recognize music emotions based on mode information seems to develop somewhat later on.

Study II also showed that 8–9-year-olds differentiated the music material in regard to beauty suggesting that, at a certain age, children start to use the concept of beauty similarly to adults.

In sum, **Study II** confirms previous findings and provides new insights into the development of aesthetic experiences in young children. Taking into account several aspects that may contribute to music appreciation, namely preference, music emotion recognition, and the application of aesthetic concepts such as beauty and ugliness, the results show that the different components may follow different developmental timelines.

Of these aspects, the association between major mode and happiness seems to occur earliest. Interestingly, at this age preferences seem to be related to the positive emotional tone conveyed by music. This again highlights the importance to consider emotions as fundamental to aesthetic experiences of music.

Study II, however, is limited in regard to the musical material used. It would be of great importance to study aesthetic responses in children using music that is ecologically valid and appropriate regarding the general cognitive and affective development of young participants.

5.3 Neural correlates of aesthetic processing

Study III and **Study IV** aimed at determining neural processes underlying evaluative judgments of music and music appreciation. **Study III** took a bottom-up approach in examining the nature and time course of evaluative compared to cognitive judgments of music. **Study IV** aimed at describing basic neural processes of auditory sound organization that eventually may lead to pleasurable experiences during music listening.

The results of **Study III** in which HMm-fans and LAm-fans listened to excerpts of both music genres and performed a liking and a genre classification task revealed that listening to the preferred genre elicited a more negative response between 230 and 370 ms after stimulus onset compared to the non-preferred genre. The non-preferred genre, in turn, was characterized by an increased LPP amplitude between 600 and 900 ms after stimulus onset.

First, the results of **Study III** strongly support the idea of at least two successive neural processing stages during evaluative processing (Jacobsen & Höfel, 2003). The findings suggest that the initial encoding of genre-specific sonic properties immediately triggers affective responses in listeners via processes that may be related to familiarity and processing fluency. It seems likely that the voluntary attending to the outcome of the evaluative system during the liking task, however, obscures these spontaneous responses.

Second, the results suggest that valence affects neural stimulus processing at early stages. A similar early effect of valence has been obtained for non-beautiful judgments of graphical pattern (Höfel & Jacobsen, 2007b; Jacobsen & Höfel, 2003) and disliking judgments of chord sequences (Brattico et al., 2010). Also Werheid et al. (2007) found that attractive faces triggered a stronger early posterior negativity (EPN) compared to non-

attractive faces. Differences in scalp distribution and polarity between the ERPs obtained in the different studies suggest that this early processing stage is stimulus-driven and domain-specific rather than domain-general.

Importantly, the LPP obtained during both tasks and for both the preferred and non-preferred genre suggests that the presentation of music automatically elicited affective response during music listening. That is, **Study III** promotes the assumption, as also proposed by the model of Leder et al. (2004), that aesthetic processing is constantly accompanied by affective responses. Generally, the findings are in line with the previously suggested model. Music preference seems to modulate early stages of aesthetic processing. As the results show, perceiver-related aspects (e.g., subjective musical preferences) may not only determine the explicit classification of an aesthetic object but already interact with stimulus processing at the level of memory integration or even earlier.

Study IV shows that expressive timing as used by performers to highlight the musical structure may facilitate the organization of sound in sequential units even during passive exposure. These variations in timing may account for why expressively played music is typically conceived of being more pleasant and of greater aesthetic value than music that is presented without any variations in its micro-structure (deadpan music). Assuming that the P3a-like component reflects the allocation of attention resources for the processing of novel or unexpected events (Polich, 2007), it seems that the extraction of timing information provided sufficient information about the possible occurrence of such a novel event. This, in turn, may have resulted in the allocation of attention resources to the phrase boundary. For expressive melodies without a pitch leap, the P3a amplitude was completely absent during passive exposure. That is, the continuation of the melody in equal interval steps at the phrase boundary was not processed as novel or unexpected event despite the preceding timing information.

The active detection of the target phrase boundary resulted in the elicitation of a clear N2b-P3a complex. The N2b was smaller for expressive melodies. Because the N2b has previously been shown to be smaller for stimuli that were expected compared to those that were less expected (Palmer et al., 2009), the results suggest that phrase boundaries preceded by expressive timing variations were more predictable than those occurring within deadpan melodies. Similarly, also the P3a amplitude was reduced when the phrase boundary was embedded in deadpan melodies. This finding implies that the recognition of

the phrase boundary was more difficult when the beginning of the new phrase could not be predicted on grounds of the preceding temporal information as in deadpan melodies.

In sum, **Study IV** demonstrates that expressive timing facilitates the recognition of phrase boundaries in music. The results provide evidence for the assumption that temporal information is extracted on a global level to generate expectations and to activate the allocation of attention towards a specific musical event. Importantly, the mechanisms unravelled in **Study IV** might account for fluent stimulus processing that eventually may lead to aesthetic pleasure (Reber et al., 2004).

Taken together, **Study III** and **Study IV** approach music appreciation and aesthetics from a neurocognitive and affective perspective. The results of **Study III** demonstrate that evaluative responses to music may be determined by background variables such as music preferences and likes. Moreover, **Study III** shows that valence might affect stimulus processing even implicitly. **Study IV** provides first evidence of facilitated processing of musical structure due to expressive timing. Although **Study III** and **Study IV** cover only a small portion of what may contribute to pleasurable experiences during music listening, they might help to gather further knowledge about those neural mechanisms underlying music appreciation and aesthetics.

5.4 A model of music appreciation and aesthetics

Figure 15 shows a tentative model of music appreciation and aesthetics. The basic outline of the model is adapted from the general model of aesthetic experience developed by Leder et al. (2004). It, however, illustrates the specific characteristics of aesthetic responses during music listening. The present model is simplified in that it does not include all previously proposed processing stages. It should be considered as a starting point that needs to be elaborated on grounds of more empirical data.

According to the model, aesthetic experiences of music are to a large extent determined by individual listener features (e.g., personality, music expertise, music preference, current mood), situational factors (e.g., music listening in public or private settings), the listening mode (attentive listening versus background music), and the intended purpose of music listening (e.g., mood or emotion regulation, relaxation, activation).

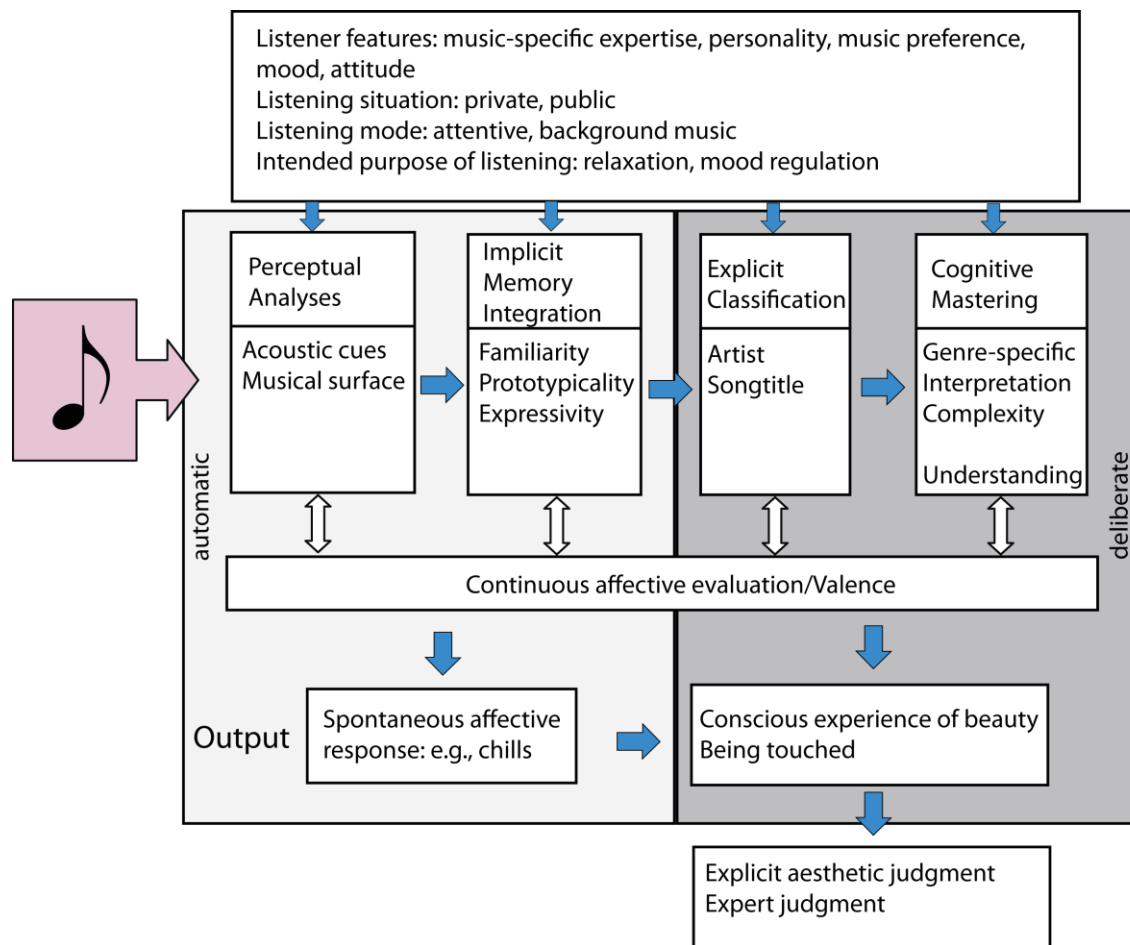


Figure 15. A model of music appreciation and aesthetics.

All of these aspects are believed to affect already the initial encoding of the acoustic features of the music. This assumption is based on previous findings showing that, for instance, attention modulates early stimulus processing as indexed by the N1 component of the ERP (Woldorff et al., 1993). There is also evidence that the N1 is affected by the pleasantness of presented musical material (Regnault et al., 2001; Schön et al., 2005). Thus, it is conceivable that the continuous affective evaluation retroacts on the perceptual analysis. In line, the findings of **Study III** suggest that early encoding of genre-specific acoustic cues is modulated by individual music preferences and the valence of music. In future studies, it needs to be tested whether the listening context, personality, mood, and other related features similarly interact with the early neural processing of music.

The mentioned features do not only affect early stimulus processing, but are assumed to equally influence subsequent processing stages, such as implicit memory integration, explicit classification, and cognitive mastering. For example, expertise with a certain

genre may specifically facilitate the cognitive mastering and understanding of corresponding music selections due to, for instance, familiarity and processing fluency (Reber et al., 2004).

Similar to Leder et al. (2004), the present model makes a distinction between automatic (light-grey shaded area) and deliberate processing (dark-grey shade area). In the model of Leder et al. (2004) aesthetic emotion and aesthetic judgments are assumed to represent the output of aesthetic processing stages. Somewhat different, the present model suggests that aesthetic processing may result in both spontaneous affective responses (e.g., psychophysiological reactions, such as chills or electrodermal activity) that need no deliberate processing and conscious experiences of beauty or being touched by music that can be verbalized if required ('The music is beautiful', 'I like the music', 'The music sounds pleasant'). Expert judgments (e.g., 'The composition is sophisticated', 'The music belongs to the most important works of the 19th century') do follow the same processing stages. That is, although the affective involvement during music listening is not verbally expressed, it is believed that also expert judgments are accompanied by evaluative processes.

6 Conclusions

The present thesis investigated the structure and content of the concept underlying aesthetic responses to music (**Study I**), the ability of school-aged children to perform aesthetic judgments (**Study II**), the neural processing of music in regard to evaluative judgments of music (**Study III**) and facilitated processing of musical groups marked by expressive timing cues (**Study IV**).

Each of the studies provides new insights regarding the processes underlying music appreciation. Specifically, **Study I** indicates that beauty is central to the conceptual space underlying aesthetic responses to music. Importantly, the result show that there is a need for individual aesthetic models that incorporate the demands of specific aesthetic domains, such as visual art, sculpture, dance, and music. Future studies aiming at investigating aesthetic responses to music may make use of the finding that beauty is regarded as emotional phenomenon and not as purely cognitive concept used for expert descriptions of music.

Study II corroborates and extends current knowledge about the developmental stages of music preference, music emotions, and the use of aesthetic categories. It shows that school-aged children are able to express their preferences, to categorize music in terms of basic emotions, and to use the concept of beauty in an adult-like manner. Because the study was limited to three music pieces differing in mode only, it would be important to investigate children's aesthetic responses using a greater variety of music excerpts. Moreover, to better delineate the development of each of the suggested components (music preference, music emotions, beauty), future studies could benefit from an inclusion of additional age groups.

Study III shows that music preference and the valence of music affect early neural music processing stages. This is an important finding that allows concluding that even the encoding of genre-specific acoustic cues is determined by individual listener features. It is an open question, however, why music preference and valence affected early stimulus processing only when no overt liking judgment was required. Although it was suggested that evaluative responses might have been obscured by focusing on the outcome of the evaluative system, this assumption needs further testing.

At last, **Study IV**, shows that expressivity in music may guide our perception and facilitate the organization of sounds into meaningful units. On a general level, the results

suggest that auditory information is implicitly and constantly being evaluated in terms of its significance. Expressive timing, however, is only one aspect that contributes to a performance's expression. Variation in loudness, pitch, timbre and other acoustic cues equally belong to the micro-structure of music performances. Thus, future studies should investigate the effects of both individual parameters and their combinations on the neural processing of phrase boundaries.

In sum, the application of both behavioural methods (**Study I and II**) and the event-related potential technique (**Study III and IV**) has proven to be beneficial for an understanding of aesthetic processes from a psychological perspective. Approaching the topic with different methodological procedures has enabled the extension of current models of aesthetic processing to a specific model of music appreciation and aesthetics. Although the model provides only a first step towards a more detailed description of the processes underlying music appreciation, it may serve the formulation of testable hypotheses and stipulate further research.

References

- Adler, G., & Adler, J. (1989). Influence of stimulus intensity on AEP components in the 80– to 200–millisecond latency range. *Audiology*, 28(6), 316–324.
- Alho, K., Grimm, S., Mateo-Leon, S., Costa-Faidella, J., & Escera, C. (2012). Early processing of pitch in the human auditory system. *European Journal of Neuroscience*, 36(7), 2972–8.
- Allesch, C. G. (2006). *Einführung in die psychologische Ästhetik*. Wien: facultas.wuv.
- Barrett, M. S. (2006). Aesthetic response. In G. E. McPherson (Ed.), *The child as musician: A handbook of musical development* (pp. 173–191). New York, NY: Oxford University Press.
- Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, 80(3), 1–10.
- Brattico, E., Jacobsen, T., De Baene, W., Glerean, E., & Tervaniemi, M. (2010). Cognitive vs. affective listening modes and judgments of music – an ERP study. *Biological Psychology*, 85(3), 393–409.
- Cacioppo, J. T., Crites, S. L. J., & Gardner, W. L. (1996). Attitudes to the right: Evaluative processing is associated with lateralized late positive event-related brain potentials. *Personality and Social Psychology Bulletin*, 22(12), 1205–1219.
- Cela-Conde, C. J., Agnati, L., Huston, J. P., Mora, F., & Nadal, M. (2011). The neural foundations of aesthetic appreciation. *Progress in Neurobiology*, 94(1), 39–48.
- Cohen, A. J., Thorpe, L. A., & Trehub, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology*, 41(1), 33–47.
- Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, 115(4), 732–744.
- Cunningham, J., G., & Sterling, R., S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion*, 12(4), 399–413.
- Dolgin, K. G., & Adelson, E. H. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music*, 18(1), 87–98.
- Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music perception*, 30(3), 307–340.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3/4), 169–200.
- Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review*, 17(3), 335–341.
- Flaisch, T., Hacker, F., Renner, B., & Schupp, H. T. (2011). Emotion and the processing of symbolic gestures: An event-related brain potential study. *Social Cognitive and Affective Neuroscience*, 6(1), 109–118.

- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1), 152–170.
- Friberg, A., Colombo, V., Frydén, L., & Sundberg, J. (2000). Generating musical performances with director musices. *Computer Music Journal*, 24(3), 23–29.
- Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. *Journal of the Acoustical Society of America*, 105(3), 1469–1484.
- Frühholz, S., Jellinghaus, A., & Herrmann, M. (2011). Time course of implicit processing and explicit processing of emotional faces and emotional words. *Biological Psychology*, 87(2), 265–274.
- Gabrielsson, A. (1987). Once again: The theme from Mozart's Piano Sonata in A Major (K. 331). A comparison of five performances. In: A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music*. Stockholm: Publication of the Royal Swedish Academy of Music No.55.
- Gabrielsson, A. (2001–2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special issue*, 122–147.
- Gerardi, G. M., & Gerken, L. (1995). The development of affective responses to modality and melodic contour. *Music Perception*, 12(3), 279–290.
- Gjerdingen, R., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research*, 37(2), 93–100.
- Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, 35(2), 129–155.
- Hideaki, K., & Semir, Z. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91(4), 1699–705.
- Höfel, L., & Jacobsen, T. (2007a). Electrophysiological indices of processing symmetry and aesthetics: A result of judgement categorization or judgment report? *Journal of Psychophysiology*, 21(1), 9–21.
- Höfel, L., & Jacobsen, T. (2007b). Electrophysiological indices of processing aesthetics: Spontaneous or intentional processes? *International Journal of Psychophysiology*, 65(1), 20–31.
- Hunter, P. G., & Schellenberg, E. G. (2010). Introduction to the study of music and emotion. In M. R. Jones, R. R. Fay & A. N. Popper (Eds.), *Music perception* (pp. 129–164). New York, Dordrecht, Heidelberg London: Springer Science + Business Media.
- Ishizu, T., & Zeki, S. (2011). Toward a brain-based theory of beauty. *PLoS One*, 6(7). e21852. doi: 10.1371/journal.pone.0021852.

- Jacobsen, T., & Höfel, L. (2003). Descriptive and evaluative judgment processes: Behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cognitive, Affective & Behavioral Neuroscience*, 3(4), 289–299.
- Jacobsen, T., Buchta, K., Kohler, M., & Schröger, E. (2004). The primacy of beauty in judging the aesthetics of objects. *Psychological Reports*, 94(3), 1253–1260.
- Jacobsen, T., Schubotz, R. I., Höfel, L., & Cramon, D. Y. (2006). Brain correlates of aesthetic judgment of beauty. *NeuroImage*, 29(1), 276–285.
- Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music*, 31(3), 273–302.
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129(5), 770–814.
- Kastner, M. P., & Crowder, R. G. (1990). Perception of the Major/Minor distinction: IV. emotional connotation in young children. *Music Perception*, 8(2), 189–201.
- Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., et al. (2005). Perception of phrase structure in music. *Human Brain Mapping*, 24(4), 259–273.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "nonmusicians" are musical. *Journal of Cognitive Neuroscience*, 12(3), 520–541.
- Koelsch, S., Kilches, S., Steinbeis, N., & Schelinski, S. (2008). Effects of unexpected chords and of performer's expression on brain responses and electrodermal activity. *PloS One*, 3(7).
- Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, 39(5), 657–663.
- Konečni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2), 115–129.
- Lamont, A. (2008). Young children's musical worlds: Musical engagement in 3.5-year-olds. *Journal of Early Childhood Research*, 6(3), 247–261.
- Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker, (Ed.), *Data analysis, machine learning and applications, studies in classification, data analysis, and knowledge organization* (pp. 261–268). Berlin, Heidelberg: Springer.
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489–508.
- Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology*, 92(3), 204–214.

- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Meyer, L. B. (1956). *Emotion and meaning in music*. London: University of Chicago Press.
- Mulder, J., Ter Bogt, T. F. M., Raaijmakers, Q. A. W., Gabhainn, S. N., & Sikkema, P. (2010). From death metal to R&B? Consistency of music preferences among dutch adolescents and young adults. *The Psychology of Music*, 38(1), 67–83.
- Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons – an ERP study. *International Journal of Psychophysiology*, 76(1), 40–51.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425.
- Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology*, 14(1–2), 53–98.
- Nan, Y., Knösche, T. R., & Friederici, A. D. (2009). Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biological Psychology*, 82(1), 70–81.
- Neuhaus, C., Knösche, T. R., & Friederici, A. D. (2006). Effects of musical expertise and boundary markers on phrase perception in music. *Journal of Cognitive Neuroscience*, 18(3), 472–493.
- North, A., & Hargreaves, D. (2008). *The social and applied psychology of music*. London, UK: Oxford University Press.
- North, A. C., & Hargreaves, D. J. (1997). Liking, arousal potential, and the emotions expressed by music. *Scandinavian Journal of Psychology*, 38(1), 45–53.
- North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology*, 70(2), 255–272.
- Novak, G. P., Ritter, W., Vaughan, H. G., & Wiznitzer, M. L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*, 75(4), 255–275.
- Ornaghi, V., & Grazzani, I. (2012). The relationship between emotional-state language and emotion understanding: A study with school-age children. *Cognition & Emotion*, 27(2), 356–66.
- Palmer, C., Jewett, L. R., & Steinhauer, K. (2009). Effects of context on electrophysiological response to musical accents. *Annals of the New York Academy of Sciences*, 1169, 470–480.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6), 717–733.

- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(2), 111–141.
- Picton, T. W., Stapells, D. R., & Campbell, K. B. (1981). Auditory evoked potentials from the human cochlea and brainstem. *The Journal of Otolaryngology. Supplement*, 9, 1–41.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148.
- Ramachandran, V. S., & Hirstein, W. (1999). The science of art. A neurological theory of aesthetic experience. *Journal of Consciousness Studies*, 6(6–7), 15–51.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality & Social Psychology Review*, 8(4), 364–382.
- Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, 13(2), 241–255.
- Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6), 1236–1256.
- Russel, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescent's mood regulation. *Psychology of Music*, 35(1), 88–109.
- Schäfer, T., & Sedlmeier, P. (2009). What makes us like music? *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)*, Jyväskylä, Finland, 487–490.
- Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research*. (pp. 361–392). New York, NY, US: Oxford University Press.
- Schön, D., Regnault, P., Ystad, S., & Besson, M. (2005). Sensory consonance: An ERP study. *Music Perception*, 23(2), 105–117.
- Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Attention and emotion: An ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, 14(8), 1107–1110.
- Schwartz, K. D., & Fouts, G. T. (2003). Music preferences, personality style, and developmental issues of adolescents. *Journal of Youth and Adolescence*, 32(3), 205–213.
- Seppänen, M., Hämäläinen, J., Pesonen, A. K., & Tervaniemi, M. (2012). Music training enhances rapid neural plasticity of N1 and P2 source activation for unattended sounds. *Frontiers in Human Neuroscience*, 6:43.

- Silvia, P. J. (2005). Emotional responses to art: From collation and arousal to cognition and emotion. *Review of General Psychology*, 9(4), 342–357.
- Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, 35, 377–396.
- Sloboda, J. A. (2010). Music in everyday life. The role of emotions. In P. N. Juslin, & J. A. Sloboda (Eds.), *Handbook of music and emotions. Theory, research, applications*. (pp. 493–514). Oxford NY: Oxford University Press.
- Sloboda, J. A., & Juslin, P. N. (2010). At the interface between the inner and outer world. Psychological perspectives. In P. N. Juslin, & J. A. Sloboda (Eds.), *Handbook of music and emotion. Theory, research, applications* (pp. 73–97). Oxford NY: Oxford University Press.
- Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Experimental Brain Research*, 222, 149–57.
- Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmüller, E., & Münte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Research*, 1070(1), 160–170.
- Sutrop, U. (2001). List task and a cognitive salience index. *Field Methods*, 13(3), 263.
- Tatarkiewicz, W. (1963). Objectivity and subjectivity in the history of aesthetics. *Philosophy and Phenomenological Research*, 24(2), 157–173.
- Tervaniemi, M. (2003). Musical sound processing: EEG and MEG evidence. In I. Peretz, & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 294–309). New York, NY, US: Oxford University Press.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, 161(1), 1–10.
- Thoma, M. V., Ryf, S., Mohiyeddini, C., Ehlert, U., & Nater, U. M. (2012). Emotion regulation through listening to music in everyday situations. *Cognition & Emotion*, 26(3), 550–560.
- Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, 21(1), 77–88.
- Trehub, S. E., Hannon, E. E., & Schachner, A. (2010). Perspective on music and affect in the early years. In P. N. Juslin, & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 645–668). Oxford New York: Oxford University Press.
- Tremblay, K. L., Inoue, K., McClannahan, K., & Ross, B. (2010). Repeated stimulus exposure alters the way sound is encoded in the human brain. *PloS One*, 5(4). e10283. doi: 10.1371/journal.pone.0010283.

- Werheid, K., Schacht, A., & Sommer, W. (2007). Facial attractiveness modulates early and late event-related brain potentials. *Biological Psychology*, 76(1–2), 100–108.
- Witvliet, C. V. O., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition & Emotion*, 21(1), 3–25.
- Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, 90(18), 8722–8726.
- Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology*, 44, 102–109.
- Wunderlich, J. L., & Cone-Wesson, B. K. (2001). Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. *The Journal of the Acoustical Society of America*, 109(4), 1526–1537.
- Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151–175.
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4), 494–521.
- Zentner, M. R., & Kagan, J. (1996). Perception of music by infants. *Nature*, 383, 29.
- Zweigenhaft, R.L. (2008). A do re mi encore. A closer look at the personality correlates of music preferences. *Journal of Individual Differences*, 29(1), 45–55.