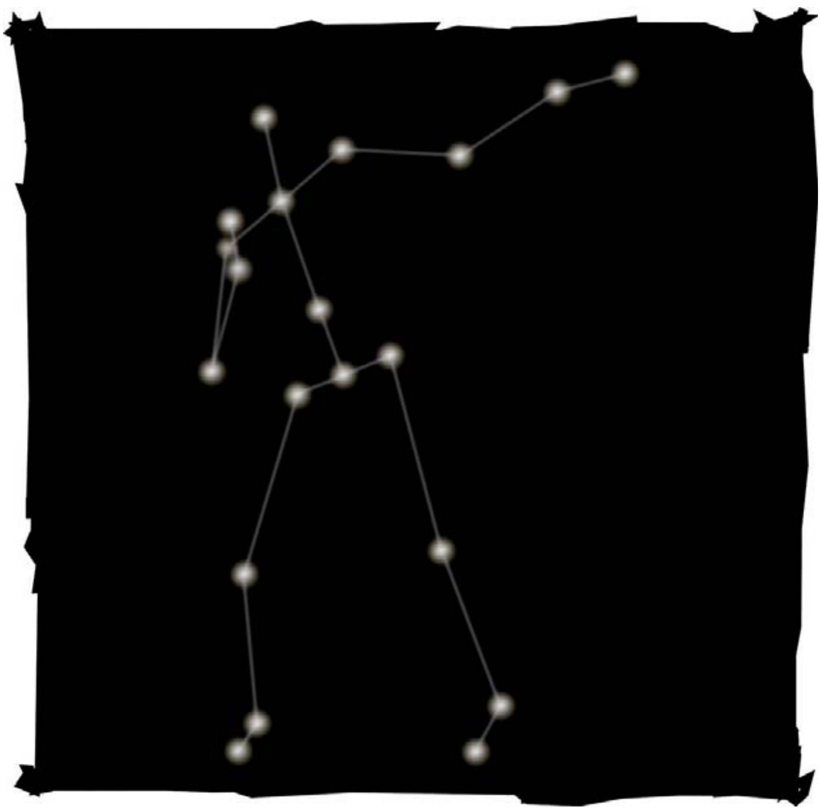
JYVÄSKYLÄ STUDIES IN HUMANITIES
215

Birgitta Burger

Move the Way You Feel

Effects of Musical Features, Perceived Emotions, and Personality on Music-Induced Movement



JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN HUMANITIES 215

Birgitta Burger

Move the Way You Feel

Effects of Musical Features, Perceived Emotions, and Personality on Music-Induced Movement

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa S212
joulukuun 7. päivänä 2013 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities of the University of Jyväskylä,
in Auditorium S212, on December 7, 2013 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2013

Move the Way You Feel

Effects of Musical Features, Perceived Emotions, and Personality on Music-Induced Movement

JYVÄSKYLÄ STUDIES IN HUMANITIES 215

Birgitta Burger

Move the Way You Feel

Effects of Musical Features, Perceived Emotions, and Personality on Music-Induced Movement

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2013

Editors
Petri Toiviainen
Department of Music, University of Jyväskylä
Pekka Olsbo, Ville Korkiakangas
Publishing Unit, University Library of Jyväskylä

Cover picture by Birgitta Burger.

URN:ISBN:978-951-39-5466-6
ISBN 978-951-39-5466-6 (PDF)

ISBN 978-951-39-5465-9 (nid.)
ISSN 1459-4323 (nid.), 1459-4331 (PDF)

Jyväskylä University Printing House, Jyväskylä 2013

*"Hab keine Angst vor deinen Schwächen,
Fürchte nie deine Fehler aufzudecken.
Sei bedacht, beruhigt und befreit,
Sei auch verrückt von Zeit zu Zeit."*

(Don't be afraid of your weaknesses,
Never fear to discover your mistakes.
Be deliberate, calm, and free,
Be crazy, too, from time to time.)

– Silbermond: *Krieger des Lichts* (2009) –

ABSTRACT

Burger, Birgitta
Move the way you feel: Effects of musical features, perceived emotions, and personality on music-induced movement
Jyväskylä: University of Jyväskylä, 2013, 79 p.
(Jyväskylä Studies in Humanities
ISSN 1459-4323 (nid.), 1459-4331 (PDF); 215)
ISBN 978-951-39-5465-9 (nid.)
ISBN 978-951-39-5466-6 (PDF)

Listening to music makes us move, ranging from barely visible movement like faint foot tapping to full-body dance movement. Such movements appear to be synchronized to the beat of the music and associated with other music-intrinsic features, while at the same time showing highly individualistic properties. The aim of this thesis is to investigate different factors that might affect movements induced by music. These factors include, for instance, musical features, perceived emotions in music, genre, and personality traits of the dancers. A motion capture study, in which 60 participants were recorded while moving to 30 music excerpts, forms the core of this thesis research. From these data, several features related to movement posture, kinematics, and kinetics were extracted using MATLAB MoCap Toolbox. Subsequent correlational analyses with musical features extracted from the stimulus data indicated strong relationships between music characteristics, in particular pulse clarity and spectral flux of certain frequency bands, and movement characteristics. A perceptual experiment consolidated the connections between sub-band spectral flux and both perceived rhythmic strength and movement propensity. Moreover, emotions perceived in the music were closely associated with various movement features. A subsequent perceptual experiment in which animated stick figures were rated regarding perceived emotions revealed that music-induced movement conveyed emotional qualities to observers, though auditory and visual information differed in their perception. Furthermore, participants' personality traits were shown to significantly affect movement characteristics. Multivariate analysis approaches disclosed that emotions perceived in the music accounted notably for relationships between musical and movement features, whereas dancers' personality traits failed to affect existing relationships between music and movement. In conclusion, this thesis presents new and innovative methods for studying music-related movement. By employing current motion capture technology and advanced analysis methods, several relationships between music-induced movement and factors influencing such movements have been specified, promoting approaches and applications related to embodied music cognition and nonverbal (everyday) human behavior.

Keywords: Music-induced movement, motion capture, musical features, emotion, personality, perception

Author Birgitta Burger
Department of Music
University of Jyväskylä
Jyväskylä, Finland
birgitta.burger@jyu.fi

Supervisor Professor Petri Toiviainen
Department of Music
University of Jyväskylä
Jyväskylä, Finland

Reviewers Associate Professor Sofia Dahl
Department of Architecture, Design
and Media Technology
Aalborg University Copenhagen
Copenhagen, Denmark

Professor Edward W. Large
Department of Psychology
University of Connecticut
Storrs, CT, USA

Opponent Associate Professor Peter Keller
The MARCS Institute
University of Western Sydney
Sydney, Australia

ACKNOWLEDGEMENTS

A bit more than four years ago, my life took a path I wouldn't have ever imagined before: I got the opportunity to join the Finnish Centre of Excellence in Interdisciplinary Music Research at the University of Jyväskylä. Barely being able to pronounce this place I set out to the north – to the cold, solitude, and wilderness. What I found, was – besides masses of snow, trees, and lakes – a place full of warmth, friendship, and spirit. Finland, Jyväskylä, and especially the CoE teams both in Jyväskylä and Helsinki have provided an extraordinary environment of support and innovation. I have been proud and grateful to be part of this group that made working on this dissertation a special experience.

First and foremost, I would like to thank my supervisor Professor Petri Toiviainen for his guidance and support throughout this PhD. I am thankful for his trust, his patience, even when asking the stupidest questions, and for him being approachable at any time, either personally or online. I consider myself lucky to work so closely with him and benefit from his invaluable knowledge, experience, and advices – be it related to research or to other significant topics such as berries, mushrooms, wines, or physical phenomena explained in layman terms.

I feel honored that this thesis has been reviewed by Associate Professor Sofia Dahl and Professor Edward Large. I wish to thank them for reading the manuscript and their beneficial comments, and for their appreciation and approval of my work. I further regard it as a special privilege having Associate Professor Peter Keller as my opponent and wish to thank him for the support and interest he has shown in my work so far.

To all my colleagues a big thanks for providing such a great working environment. My first gratitude to the person I have probably spent most time with during my PhD: my office-mate Anemone van Zijl. We couldn't be more different in so many respects and yet complement each other so well. Thanks for all the fruitful discussions, ideas, opinions, support, knowledge, noises, and friendship over the past years – [Ping!]. Further thanks to Suvi Saarikallio, Geoff Luck, and Marc Thompson for the collaboration during my PhD. I so much benefited from your expertise in experiment design and analysis, numerous proofreading, and comments on my articles. Special thanks to Marc for teaching me how to do motion capture. Thanks to Markku Pöyhönen for providing support in all administrative and practical issues that made working life so much easier. Henna Peltola, thanks for being such a soulmate. Thanks to Pasi Saari, Martín Hartmann, and Iballa Burunat for countless discussions, drinks, food, and fun. Thanks to Olivier Brabant and Margarida Baltazar in anticipation of having fun in future times. Mikko Leimu, thanks for all the tech and mental support (especially during the stroke study). Thanks to Jaakko Erkkilä and Tuomas Eerola for support and being great bosses in difficult times. Thanks to Esa Ala-Ruona for introducing me to the Stroke project and for the massages. Thanks to Olivier Lartillot for development and help with MIRToolbox. Petteri Poutiainen, thanks for solving numerous last-minute computer issues. Thanks to Marko Punkanen, Mónika Benedek, and Tommi Himberg. Thanks to Jörg Fachner for talking German and – with Sonja and Alvar – providing a little "family" for me. Thanks to Jonna Vuoskoski, Vinoo Alluri, and Rafael Ferrer for giving valuable insights into being and surviving as a PhD student. Thanks to Eva Istok and Clemens Maidhof for the Köln connection. Thanks to Jakob Abesser for an awesome summer 2010. Thanks to Kristian Nymoen for profound insights into motion capture and a great mocap workshop at NIME 2011.

I'd like to say thank you to my German and Finnish friends that provided distraction and distance from work when needed. Besonderer Dank geht an meine Freunde aus Köln, die trotz der Distanz unsere Freundschaft aufrecht erhalten haben: Birgit, Kerstin, Christine und Martina. Danke an Lüder für alles. Tine, danke für Weihnachtsplätzchen backen, Ausflüge und alles andere, was wir zusammen in Jyväskylä angestellt haben.

Haluaisin vielä kiittää poikaystäväni Juhon perhettä. Annoitte minulle lämpöä ja huolenpitoa, jota vain perhe voi antaa. Kiitos Helena, Hannu, Anna-Maria ja Antti.

Heartfelt thanks go to my boyfriend Juho, who has bravely and patiently gotten along with the researcher in me during our relationship, especially in those times when significance levels, latent variables, and unexplained variances were among our favorite conversation topics. Thanks for all the support, encouragement, and distraction you provided.

Zu guter Letzt möchte ich meiner Familie für die Unterstützung, Bildung und Liebe danken, die sie mir Zeit meines Lebens geschenkt hat. Skype kann zwar das Zuhause nicht ersetzen, dennoch hat es die Distanz merklich verkürzt und häufig die Illusion erzeugt, während unserer endlosen Internetgespräche quasi neben Euch auf dem Sofa zu sitzen. Ich danke Euch dafür, dass Ihr trotz der 2000 km Distanz immer da und nah seid.

Jyväskylä, November 2013
Birgitta Burger

LIST OF PUBLICATIONS

List of publications included in this thesis (reprinted after the summary part):

I Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology 4:183*, 1–10.

II Burger, B., Ahokas, R., Keipi, A., & Toiviainen, P. (2013). Relationships between spectral flux, perceived rhythmic strength, and the propensity to move. In R. Bresin (Ed.), *Proceedings of the 10th Sound and Music Computing Conference* (pp. 179–184). Stockholm, Sweden: KTH Royal Institute of Technology.

III Burger, B., Saarikallio, S., Luck, G., Thompson, M. R. & Toiviainen, P. (2013). Relationships between perceived emotions in music and music-induced movement. *Music Perception 30*(5), 519–535.

IV Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2013). Oh happy dance: Emotion recognition in dance movement. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music and Emotion*. Jyväskylä, Finland: University of Jyväskylä.

V Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality 44*(6), 714–720.

VI Burger, B., Polet, J., Luck, G., Thompson, M. R., Saarikallio, S., & Toiviainen, P. (2013). Investigating relationships between music, emotions, personality, and music-induced movement. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music and Emotion*. Jyväskylä, Finland: University of Jyväskylä.

VII Burger, B. & Toiviainen, P. (2013). MoCap Toolbox – A Matlab toolbox for computational analysis of movement data. In R. Bresin (Ed.), *Proceedings of the 10th Sound and Music Computing Conference* (pp. 172–178). Stockholm, Sweden: KTH Royal Institute of Technology.

AUTHOR'S CONTRIBUTIONS TO PUBLICATIONS

I This article is part of an ongoing project entitled Music, Movement, & Personality (MMP). The author was responsible for data collection, movement data preprocessing, movement and musical features extraction, data analysis, and writing.

II This conference proceedings article is the outcome of a student project carried out during a practice course on Music Perception within the local Master's program Music, Mind and Technology. The author acted mainly as supervisor and initiated the experiment, provided the patch used in the data collection, and gave support in data analysis and writing.

III This article is also part of the MMP project. The author was responsible for data collection of both movement and emotion data, movement data preprocessing, movement features extraction, data analysis, and writing.

IV This conference proceedings article is another outcome of the MMP project. The author was responsible for experiment design, stimuli preparation, data collection, data analysis, and writing.

V This article is part of the MMP project as well. The author was responsible for data collection, movement data preprocessing, movement feature extraction, and contributed to the writing.

VI This conference proceedings article is related to the MMP project. The author was responsible for data collection and processing, data analysis, and writing.

VII This conference proceedings article describes the MoCap Toolbox, a Matlab toolbox for analyzing and visualizing motion capture data to which the author contributed several functions and improvements. The author was mainly responsible for designing and writing the article.

CONTENTS

ABSTRACT
ACKNOWLEDGEMENTS
LIST OF PUBLICATIONS
AUTHOR'S CONTRIBUTIONS TO PUBLICATIONS
CONTENTS

1 INTRODUCTION 11

2 EMBODIED MUSIC COGNITION 14

3 FACTORS INFLUENCING MUSIC-INDUCED MOVEMENT 17
3.1 Music 17
3.1.1 Musical features relevant for music-induced movement 19
3.1.2 Genre 20
3.2 Emotions 21
3.2.1 Emotion-specific nonverbal behavior 21
3.2.2 Emotions in music 22
3.3 Personality 25

4 AIMS AND OVERVIEW OF THE THESIS 28

5 METHODOLOGIES 30
5.1 Motion capture 30
5.1.1 Motion capture data collection 32
5.1.2 Mocap Toolbox 36
5.1.3 Movement feature extraction 38
5.2 Musical feature extraction 41
5.3 Experiment designs 44
5.4 Analysis methods 44

6 STUDIES 46
6.1 Aims of each study 46
6.2 Methods 48
6.2.1 Participants 48
6.2.2 Stimuli 48
6.2.3 Apparatus 49
6.2.4 Procedures 49
6.2.5 Movement features 50
6.2.6 Musical features 50
6.3 Results and discussion 51
6.3.1 Study I – Rhythm- and timbre-related musical features and music-induced movement 51

6.3.2 Study II – Spectral flux, perceived rhythm, and the propensity to move 52
6.3.3 Study III – Perceived emotions and music-induced movement 53
6.3.4 Study IV – Emotion recognition in music-induced movement 55
6.3.5 Study V – Personality and music-induced movement 57
6.3.6 Study VI – Music, emotion, personality, and music-induced movement 58

7 GENERAL DISCUSSION AND CONCLUSION 59
7.1 Limitations and future work 62
7.2 Concluding remarks 65

YHTEENVETO (FINNISH SUMMARY) 66

REFERENCES 67

APPENDIX 1 SONG LIST 79

1 INTRODUCTION

Music makes us move, be it consciously and intentionally in the night club, or subconsciously and spontaneously when listening to the radio while doing, for instance, the dishes. Keller and Rieger (2009) stated that simply listening to music can induce movement, and in a self-report study conducted by Lesaffre et al. (2008), most participants reported moving when listening to music. Furthermore, Janata, Tomic, and Haberman (2012) reported a study in which they asked participants to tap to music and found that they not only moved the finger/hand, but also other body parts, such as feet and head. Additionally, the tapping condition (isochronous versus free tapping) influenced the amount of movement: the more "natural" the tapping condition, the more movement was exhibited. Thus, it seems that there is a close relationship between music and movement exhibited during listening, indicating that music-induced movement is a central element of every-day behavior / music listening.

Music-induced movement covers any movement that occurs when humans listen to music, ranging from professional dance and (layperson/every-weekend) dancing in the night club to subconsciously nodding the head or tapping the foot. Such movements occur, for instance, when listening to music at home, or when being on a jazz or maybe even on a classical concert (despite the cultural constraints related to movement during classical concerts). Nevertheless, it seems that people enjoy moving to music, as it might serve as a way to express themselves in relation to the music.

The proclivity to move with music seems to be built into human nature. The notion of groove describes this experience of wanting to move to (some aspect of) the music (Janata et al., 2012; Madison, 2006; Madison, Gouyon, Ullén, & Hörnström, 2011). Madison et al. (2011) and Janata et al. (2012) found connections between groove and rhythmic aspects of the music, while Stupacher, Hove, Novembre, Schütz-Bosbach, and Keller (2013) established links between perceived groove and motor activity in the brain. The results of these studies showed furthermore that the experience of groove in music was consistent among participants and that groove was related to familiarity and enjoyment of the music (Janata et al., 2012).

While research on groove focusses on the sensation of wanting to move (in particular in relation to music-intrinsic aspects), the experiments conducted within the studies presented in this thesis concentrated on the actual, physical movements that individuals perform when listening to music. In general, such movements appear to be synchronized to the beat of the music and associated with other music-intrinsic features. However, at the same time, they show highly individualistic properties that appear related to personality or preference. The aim of this thesis was to investigate music-induced movement of adults without a professional dance background when asked to move naturally to music. The work is of exploratory nature, since not much previous research has been conducted that could have served as a basis for formulating specific hypotheses. Besides intrinsic musical characteristics, as suggested by the groove literature, this thesis research additionally focussed on more abstract musical concepts, such as perceived emotions and genre of the music, as well as on personality as an individual factor. These four factors were chosen, because they represent a broad range of different viewpoints on music-induced movement. They are certainly not exhaustive and will not cover the entire field of music-induced movement, but they form an appropriate set of aspects as a starting point for further investigations of the field. In order to adopt a feasible approach, the studies utilized static movement and music characteristics obtained by averaging the time-series of movement and musical features to receive one descriptor value for each feature. Retaining the time-series would have allowed investigating dynamics and temporal development of the features. Such – more complex – analysis methods, however, will be implemented in future work that will follow this thesis.

In order to achieve its goal, this thesis employed an interdisciplinary approach by using motion capture technology to record movement data as well as computational feature extraction of both movement and musical features, perceptual evaluations, and a wide range of statistical methods to analyze the data. The research presented in this thesis is technology-driven and would have been impossible without the technological opportunities available at the Music Department of the University of Jyväskylä and within the Finnish Centre of Excellence in Interdisciplinary Music Research. The department offers a high-quality motion laboratory hosting a Qualisys[1] optical marker-based motion capture system with which this thesis research was carried out.

The increasing opportunities of quantitative research methods for recording and analyzing body movement have offered new insights and perspectives for studying music-related movement. Music-induced movement, especially quasi-spontaneous, free movement to music yields large amounts of unconstrained data offering a considerable amount of challenges in data processing and analysis only feasible to be tackled with appropriate equipment and applications. Several applications have been developed so far. Camurri, Lagerlöf, and Volpe (2003); Camurri, Mazzarino, Ricchetti, Timmers, and Volpe (2004), for instance, created a video analysis tool (called EyesWeb) to recognize and classify expressive and emotional gestures in professional dance performances. Jensenius (2006) devel-

1 www.qualisys.com

oped the technique of motiongrams for visualizing and analyzing movement and gestures. Nymoen (2013) presented the technique of mocapgrams for visualizing gross patterns of optical marker-based motion capture data. Furthermore, tools like the Matlab MoCap Toolbox, which is described in this thesis (see Section 5.1.2), offer excellent possibilities for visualizing and analyzing movement data on computational basis.

Five out of the seven publications included in this thesis were products of an ongoing project called Music, Movement, & Personality (MMP) within the Centre of Excellence in Interdisciplinary Music Research. The aim of the project is to investigate music-induced movement from as many angles as possible. Besides the factors addressed in the thesis (music, emotions, genre, and personality), members of the project group have been investigating synchronization between music and movement, mood of the dancers, and attraction-related influences.

This thesis summary shall give an overview of the research conducted for this thesis, including descriptions of the theoretical and empirical background, methodologies employed in the different studies, and the studies themselves.

Chapters 2 and 3 outline the background linked to this work, in particular embodied music cognition as the theoretical framework in which this thesis is embedded, and the empirical work summarizing movement-related research concerning musical features, genres, emotions, and personality.

The aim of the thesis is reported in Chapter 4, providing an overview of the organization of the thesis.

Chapter 5 describes the methodologies used in this thesis. Optical motion capture as the technology for collecting movement data will be introduced in detail, as well as musical feature extraction, experiment designs, and analysis methods employed in the thesis research.

The six empirical studies included in this thesis are outlined in Chapter 6. Aim, specific method-related aspects, and a summary of results and discussion are provided for each study.

The final chapter, Chapter 7, offers discussion, limitations, and conclusive remarks related to the main findings of this thesis research.

2 EMBODIED MUSIC COGNITION

Music has the capacity to elicit spontaneous movement in listeners. It can even make people move in an organized way – they, for example, mimic instrumentalists' gestures or rhythmically synchronize with the pulse of the music by tapping their foot, nodding their head, or moving their whole body in various manners (Leman & Godøy, 2010; Godøy et al., 2006). Moreover, Leman (2007) suggests, "Spontaneous movements [to music] may be closely related to predictions of local bursts of energy in the musical audio stream, in particular to the beat and the rhythm patterns" (p. 96).

Such utilization of the body is the core concept of embodied cognition, which claims that the body is involved in or even required for cognitive processes (e.g., Lakoff & Johnson, 1980, 1999; Shapiro, 2011; Varela, Thompson, & Rosch, 1991). The notion that human cognition is linked to the existence of the body and to corporeal experience situated in an environment has gained increased recognition in recent decades. Similar to the whole field of cognitive sciences, research on embodied cognition is interdisciplinary and combines areas such as psychology, artificial intelligence, philosophy, neuroscience, linguistics, and anthropology. However, what distinguishes embodied cognition from more traditional approaches to cognition, is that cognition in the embodied view is not seen as temporally planned (perceive, think/compute, act), and organized in mental representations, but as interactive, corporeal, and real-time. Perception and action are mutually influenced by each other, as cognition and knowledge of the world are constructed through continuous couplings of sensory (e.g., visual) information and goal-directed actions in real-time. Perception requires action provoking a new perception, resulting in a steady interplay of adjustments to adapt to the changing conditions of the environment. Bodily experience and exploration of the environment are essential for understanding the own perceptual-motor abilities and for learning and accomplishing more complex cognitive tasks. Thus, a constant interaction between mind/brain, sensorimotor capabilities, body, and environment forms the basis for human cognition (e.g., Anderson, 2003; Clark, 2011; Pfeifer & Scheier, 1999; Wilson & Foglia, 2011).

Following the idea of embodied cognition, we can approach music (or musical involvement) by linking our perception of it to our body movement (Leman, 2007). One could postulate that our bodily movements reflect, imitate, help to parse, or support understanding the structure and content of music. Leman suggests that corporeal articulations could be influenced by three (co-existing) components or concepts: "Synchronization", "Embodied Attuning", and "Empathy", which differ in the degree of musical involvement and in the kind of action-perception couplings. "Synchronization" forms the fundamental component, as synchronizing to a beat is easy and spontaneous. As the first step in engaging with the music, movements could be used for imitation and prediction of beat-related features in the music, as suggested with the term 'inductive resonance'. The second component, "Embodied Attuning", concerns the linkage of body movement to musical features more complex than the basic beat, such as melody, harmony, rhythm, tonality, or timbre. Following this idea, movement could be used to reflect, imitate, and navigate within the musical structure in order to understand it. Finally, "Empathy" is seen as the component that establishes the link to expressivity and emotions. In other words, listeners feel and identify with the emotions expressed in the music and imitate or reflect them by using body movements.

Music-induced movement involves both perception and motor activity. Connections between these two areas have been proposed in several disciplines, the most known being in speech perception and in vision. The motor theory of speech perception, for example, claims that in order to understand speech, listeners rather identify and simulate the vocal tract gestures needed to articulate the heard utterances (by activating motor areas in the brain), than decode the sound patterns of the utterance (Liberman & Mattingly, 1985). Although the theory has been controversially discussed, brain studies have provided support by finding motor-related activity in the brain during speech perception (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Hickok, Buchsbaum, Humphries, & Muftuler, 2003). These studies are closely related to the discovery of the mirror neuron system (e.g., Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Rizzolatti & Craighero, 2004) – a network of neurons that are both active when performing an action and when perceiving another one performing the same action – which has led to further support for the notion of embodied cognition, as it suggests that the motor system is active during perception as well as in the process of understanding goals and intentions (e.g., Iacoboni et al., 2005). Various music-related studies (for a review see Zatorre, Chen, & Penhune, 2007) could show that motor areas in the brain were active when musicians and non-musicians, for example, listened to different beats without the intention to move to them (with motor activity found to be related to the presence of a beat) (Bengtsson et al., 2009; Grahn & Rowe, 2009), when pianists listened to piano performances (Bangert et al., 2006), when pianists observed piano playing (Haslinger et al., 2005), or when non-musicians, after learning to play a piece on the piano, listened to it again (Lahav, Saltzman, & Schlaug, 2007). Stupacher et al. (2013) could furthermore established links between high-groove stimuli and increased motor activity in the brain during music listening. Such studies give support to the notion that music and movement are interconnected in various ways.

3 FACTORS INFLUENCING MUSIC-INDUCED MOVEMENT

Music-induced movement is a multilayered phenomenon influenced by a range of different factors. Some aspects are very obvious, such as the music itself that the listener is exposed to, whereas some other might be more concealed, such as the emotional connotation of the music. This chapter will introduce the four factors that were selected to be investigated in this thesis research in more detail: musical characteristics, genre, perceived musical emotions, and personality.

3.1 Music

Although it seems obvious that musical characteristics are strongly related to the way people move to music, this aspect of music-induced movement has been rarely investigated. So far, studies have been tackling synchronization abilities of infants and adults, spontaneous motion to short sounds as well as dancing/moving to systematically manipulated stimuli and professional dancing.

Zentner and Eerola (2010) investigated infants' ability to bodily synchronize with musical stimuli, finding that infants showed more rhythmic movement to music and metrical stimuli than to speech, suggesting a predisposition for rhythmic movement to music and other metrical regular sounds. Eerola, Luck, and Toiviainen (2006) studied toddlers' corporeal synchronization to music, finding three main periodic movement types being at times synchronized with the pulse of the music. Toiviainen, Luck, and Thompson (2010) investigated how music-induced movement exhibited pulsations on different metrical levels, and showed that eigenmovements of different body parts were synchronized with different metric levels of the stimulus. Nymoen, Caramiaux, Kozak, and Torresen (2011); Nymoen, Torresen, Godøy, and Jensenius (2012) studied sound tracing, an approach to investigate how individuals relate motion to sound by bodily reacting to short, systematically manipulated sounds. Caramiaux, Bevilacqua, and Schnell (2010) analyzed relationships between gestures and sound in music performance and listening. Witek et al. (2012) investigated spontaneous body reactions to various kinds of drum loops representing different levels of groove. Van Dyck et al. (2013) studied the effect of dynamics of the bass drum and found that participants' spontaneous movements increased with the presence of the bass drum. Naveda and Leman (2010) as well as Leman and Naveda (2010) investigated movement in samba and Charleston dancing, focusing on spatiotemporal representations of dance gestures and discovered movement periodicities corresponding to different metrical levels in the music. Casciato, Jensenius, and Wanderley (2005) conducted a study, in which three professional dancers freely moved to non-tonal acoustic music. Despite the lack of a salient pulse in most of their stimuli, rhythmic features were reflected in the movements of the dancers, such as long notes embodied with smooth and slow arm movements or rhythmic attacks associated with slashing arm movements, showing commonalities to how the sounds were produced. Stevens, Schubert, Wang, Kroos, and Halovic (2009) studied movements of professional dancers regarding time keeping with and without music. Krumhansl and Schenck (1997) conducted a study on classical ballet dance in which they found that the dance movements could convey structural boundaries of the music.

It has been argued that music and dance (or, more general, movement to music) have evolved together in most cultures (Arom, 1991; Cross, 2001) and are crucial elements of most social and collective human behavior (Brown, Merker, & Wallin, 2000). Nettl (2000) noted that most cultures have developed coordinated dance movements to rhythmically predictable music. There is neurobiological evidence for a connection between rhythmic (and beat) components of music and movement (e.g., Bengtsson et al., 2009; Chen, Penhune, & Zatorre, 2009; Grahn & Brett, 2007; Grahn & Rowe, 2009; Stupacher et al., 2013), suggesting a predisposition for movement when listening to music. Besides brain studies, also behavioral studies suggest links between movement/body and rhythm/beat aspects in music: Phillips-Silver and Trainor (2008) showed that especially head movements were found to bias metrical encoding of rhythm and meter perception. Moreover, Trainor, Gao, Lei, Lehtovaara, and Harris (2009) discovered that galvanic stimulation of the vestibular system could be used to disambiguate an ambiguous metric pattern. Todd, Cousins, and Lee (2007) found that 16% of variation in preferred beat rate could be predicted from anthropometric factors, such as weight as well as length and width of certain body segments. Such findings and the studies related to groove, as mentioned earlier, have led to the assumption that humans prefer music that facilitates entrainment and synchronization and respond to it with movement (Madison et al., 2011).

However, since systematic investigations targeting the relationships between musical features and human movement characteristics have not been conducted, this thesis research was aimed at filling this gap. Furthermore, finding dependencies of musical characteristics and body movements that are consistent between individuals would support the notion of embodied music cognition (Leman, 2007).

3.1.1 Musical features relevant for music-induced movement

Based on the neurobiological and behavioral links between movement/motor activity and beat and rhythmic components of music as described above, it appears straightforward to assume that such musical elements are relevant for inducing movement in listeners.

Rhythmic music is based on beats, which can be physically characterized as distinct energy bursts in time. If such beats occur as regular and repetitive temporal patterns, they give rise to a percept of pulse. Beat and pulse structures can be regarded as the basic metrical structure in music from which more complex temporal structures, such as rhythm, emerge. This is typically achieved by subdividing the basic metrical structure in smaller and larger units of varying lengths, constructing a metrically interlocked grid with events on different temporal levels (Parncutt, 1994). These rhythmic structures can vary, for example, in the degree of pulse clarity. Pulse clarity estimates, on a large time scale, how clearly the underlying pulsation in music is perceivable and can therefore be regarded as a measure for the underlying periodicity of the music (Lartillot, Eerola, Toiviainen, & Fornari, 2008). Another aspect of rhythmic structure is covered by spectro-temporal features, such as the sub-band spectral flux, which has been found to be among the most important features contributing to polyphonic timbre perception (Alluri & Toiviainen, 2010). Spectral flux measures spectral change, which, when taken separately for different sub-bands, reflects the strength of rhythmic elements created by instruments within the frequency range of the respective sub-band. However, besides being used in music information retrieval-related applications, such as automatic classification (see, e.g., Jiang, Lu, Zhang, Tao, & Cai, 2002 and Cai, Lu, Hanjalic, Zhang, & Cai, 2006, both slightly differing in their technical implementation), the perceptual dimension of spectral flux has so far only been studied in connection to polyphonic timbre (see Alluri & Toiviainen, 2010, 2012), but neither in relation to rhythm perception nor in relation to music-induced movement. It could be assumed that sub-band flux is a crucial feature not only in a (passive) listening situation, but also in a (active) movement situation. Furthermore, related timbral characteristics could have an influence on movement responses to music. For instance, high amounts of percussive elements in music could result in fast movements, reflecting the way such sounds are often produced. Following these notions, it could be assumed that variations in musical features, such as Pulse Clarity, Spectral Flux, or Percussiveness, not only increase or decrease the amount of movement, but also change the kinds and properties of the movements.

Besides features such as pulse clarity, tempo is an important factor contributing to the perception of rhythm (Fraisse, 1982). Tempo is the speed at which beats are repeated, the underlying periodicity of music, and is usually measured in beats per minute (bpm). A large body of research has been conducted on listeners' abilities to perceive different tempi, synchronize to them, and reproduce them in tapping tasks (for reviews see Repp, 2005; Repp & Su, 2013). Large and Palmer (2002) have developed a model on listeners' perception of temporal regularity in music performance. The model was operationalized as a system of multiple, self-sustained oscillations running at different periods corresponding to the hierarchical levels of temporal structure. The system tracked temporal structures of music performances containing expressive tempo variations and could predict and interpret such temporal deviations as musically expressive by classifying them into discrete metrical categories. With respect to tempo-synchronization and -production tasks, free tapping tasks have found that a majority of participants tapped at a rate close to 600 msec, though the individual rates differed considerably (Fraisse, 1982). Synchronizing to steady, periodic beat stimuli is possible at a wide range of tempi, however it is most regular and accurate for intervals around 400-500 msec (Collyer, Broadbent, & Church, 1992), respectively 400-800 msec (Fraisse, 1982), while with slower and faster tempo the time between two taps becomes more variable. Moelants (2002) suggested 120 bpm as the preferred tempo – the tempo where tempo perception is considered to be optimal and appears most natural. Interesting to note here is that literature often draws links between spontaneous/preferred tempo and repeated motor activities, such as walking, for which the spontaneous duration of steps is around 500-550 msec (Fraisse, 1982; MacDougall & Moore, 2005; Murray, Drought, & Kory, 1964). Walking has been suggested as "a fundamental element of human motor activity" (Fraisse, 1982, p. 152) and could therefore serve as an origin of preferred tempo perception. Following these considerations, it could furthermore be assumed that music with tempi around 110-120 bpm stimulate movement more than music with other tempi. This thesis research aimed at investigating pulse-, tempo-, and rhythm-related musical features and thereby revealing relationships between these features and various movement characteristics.

3.1.2 Genre

Musical genres are descriptions created by humans to label and categorize different pieces of music. They usually cannot be strictly defined and do not have clear boundaries, as they arise based on interactions between the community of listeners and are highly context dependent (Aucouturier & Pachet, 2003). However, musical pieces belonging to the same genre share certain musical characteristics, such as instrumentation, rhythmic structure, and harmonic and pitch content (Tzanetakis & Cook, 2002). Therefore, genre offers a description of music both in musical and in symbolic/semantic terms (Aucouturier & Pachet, 2003). The detection of audio features in order to automatically classify music by their genre is an important research applications in the field of music information retrieval (e.g., Aucouturier & Pachet, 2003; Costa, Oliveira, Koerich, Gouyon, & Martins, 2012; Hartmann, Saari, Toiviainen, & Lartillot, 2013; Silla Jr., Koerich, & Kaestner, 2009; Tzanetakis & Cook, 2002). Thus, it seems highly likely that such identifiable acoustic cues would also elicit very specific movements (e.g., Godøy & Jensenius, 2009). Particular genres are indeed commonly associated with characteristic movements, as can be seen, for instance, in the way listeners tend to nod their head or tap their foot when listening to jazz music, or remain still when being in a classical music concert. Furthermore, stereotypic and culturally-driven movements are linked with particular genres, for example, "air guitar" (Godøy et al., 2006) or "head banging" motions linked to rock music, and swaying of the hips to latin music. For their 'Latin Music Database', Silla Jr. et al. (2009) even used two professional dance teachers that labeled musical genres based on how each music was danced, indicating a strong connection between music and movement in latin music. Despite these rather obvious relationships, there appears to be a lack of empirical work examining the effect of musical genre on listeners' body movement, a topic being addressed in this thesis.

3.2 Emotions

In addition to music-intrinsic characteristics, music-induced movement might be shaped by the emotional content of the music. Emotions are an essential component of musical expression (e.g., Gabrielsson & Lindström, 2010) that can have a strong influence, for instance, on the listener's mood (e.g., Saarikallio, 2011). However, only a few studies have addressed possible connections between emotions and music-induced movement.

3.2.1 Emotion-specific nonverbal behavior

Research on emotion-specific nonverbal behavior has shown that it is possible to successfully express and distinguish emotions based on movement characteristics, and furthermore to communicate emotional states to observers through body movements. Already in 1872, Darwin assigned certain body movements and postures quite specifically to emotional states; joyful movements, for example, were described as jumping, stamping, body thrown backwards and shaking, and upright torso and head, whereas anger was characterized by trembling body, shaking fist, erected head, and expanded chest. Sad movements were described as passive and motionless with a downward directed head. Despite these finding, the topic was rather neglected for a while – probably due to lack of appropriate technologies – and research focused more on facial expression (e.g., Ekman, 1982), following the assumption that gross body movements would only communicate intensity of emotions, but not qualitative characteristics (Ekman & Friesen, 1974).

Nevertheless, several recent studies have shown that emotions can be expressed by, and successfully recognized from, body postures or movements. Wallbott (1998) conducted a study in which he used a scenario-based approach, with professional actors performing certain emotions. Besides a recognition test, he also analyzed movement features characteristic and distinctive for each emotion category. Dael, Mortillaro, and Scherer (2012) also used actors in a scenario-based approach, and manually labeled their data with a self-developed microcoding system (Body Action and Posture system; BAP) including 49 movement variables, such as "shoulder action" or "arm action towards body". A different approach was chosen by de Meijer (1989), who used actors performing several movement characteristics instead of emotions. These movements differed in general features, such as trunk or arm movement, velocity, and spatial direction. In a subsequent step, observers attributed emotional characteristics to the movements resulting in coherent links between emotions and movements. Coulson (2004) used pictures of static body postures of computer-generated figures, which he systematically manipulated in several parameters and showed that different postures could communicate different emotions. Clynes (1980, 1992), for example, investigated in the context of human-machine interaction from the early seventies onwards how communication can be made more human-like and more efficient. For this purpose, he developed a repertoire of so-called dynamic transient forms, or sentic forms, specific for each emotion.

Emotion recognition and communication to observers has also been investigated in several studies. Atkinson, Dittrich, Gemmell, and Young (2004) compared static vs. dynamic whole body expressions and full-light vs. point-light displays and found that all of them could communicate emotions, though recognition rates and misclassification patterns differed for individual emotions. Pollick, Paterson, Bruderlin, and Sanford (2001) investigated visual perception of emotions in simple arm movements shown as point-light displays, such as drinking and knocking. They found that arm movements could communicate emotions, though observers tended to confuse similar emotions. Besides in acted emotion approaches, emotion recognition has been studied in other contexts as well, such as gait. Montepare, Goldstein, and Clausen (1987), for instance, showed that happiness, anger, and sadness could be successfully recognized in walking patterns. Research in linguistics investigated integration of auditory and visual information in emotion perception using face-voice stimuli. Such studies showed that usually the visual information dominated (Collignon et al., 2008), and that bimodal stimuli could be successfully integrated even if they displayed incongruent emotion combinations (de Gelder & Vroomen 2000; Massaro & Egan, 1996).

3.2.2 Emotions in music

Emotions are a central component of music and have been investigated in a large number of music-related studies. According to Krumhansl (2002), people report that their primary motivation for listening to music is its emotional impact. Various rating experiments have shown that listeners are able to perceive emotional content in music. One of the issues that researchers in this field face concerns the choice of emotion categories and models that can cover musical emotions successfully. As Eerola and Vuoskoski (2011) noted, there are basically three different approaches that have been used in studies on music and perceived emotions: discrete (basic) emotions (e.g., Balkwill & Thompson, 1999; Gabrielsson & Juslin, 1996), domain-specific emotion models, such as GEMS (Geneva Emotion Music Scale, see Zentner, Grandjean, & Scherer, 2008), and dimensional models (e.g., Ilie & Thompson, 2006; Schubert, 1999). The discrete model describes emotions as unidimensional, independent from each other, and derived from a limited number of universal and innate basic emotions, such as happiness, anger, sadness, or fear. The advantage of using this approach is the possibility to focus on emotions that are commonly expressed and represented by music, and name them in clear terms. Inspired by this approach, domain-specific discrete scales such as GEMS have been developed to create a repertoire of emotion terms that is highly music-specific. A disadvantage of discrete emotions, according to Eerola and Vuoskoski (2011), is that they can be mixed and confused with each other in discrimination tasks. This issue might be reduced by using a dimensional approach, in which all affective states are represented as a combination of two (usually valence and arousal) or three (valence, arousal, and tension) mutually independent dimensions. A common way to illustrate dimensional models is in a spatial design, such as a circle, with two bipolar dimensions – valence forming the horizontal axis, and arousal the vertical axis (in case of a three-dimensional model, tension would form the third axis resulting in a spherical design). Examples of such approaches are the circumplex models of affect by Russell (1980) and Schlosberg (1954). Eerola and Vuoskoski (2011) compared the different models and showed that, although both dimensional and discrete models performed similarly, the discrete model performed worse when it came to ambiguous emotions. Additionally, they found that a two-dimensional model consisting of valence and arousal successfully represented music-related emotions.

Musical emotions are conveyed not only by the music itself, but also through movement. In music performance, for example, movements additional to the plain movements needed for producing the sound are used to communicate expressivity (e.g., Thompson & Luck, 2012; van Zijl & Luck, 2013; Wanderley, Vines, Middleton, McKay, & Hatch, 2005). Dahl and Friberg (2007) investigated emotion communication of a marimba, a bassoon, and a saxophone player, each of them performing a piece of music with different emotional intentions (happy, angry, sad, and fearful). Observers were presented with visual-only elements of the performances and were asked to identify movement cues used to convey the different emotions. For each of the four emotions, distinct movement patterns related to the regularity, fluency, speed, and amount of movement could be identified.

More direct links between music and emotion-specific movement have been investigated in research on dance, in which movement is the only way to convey expressivity and emotion. Camurri et al. (2003) described a study in which professional dancers were asked to perform the same dance with four different emotional intentions (anger, joy, sadness/grief, and fear). In a qualitative analysis, they identified characteristic movements for the expression of each of these emotions, mainly based on force and tempo. Camurri et al. (2003) and Camurri et al. (2004) developed a video-based analysis tool to recognize and classify expressive gestures in professional dance performances. This tool was used, for example, in Castellano, Villalba, and Camurri (2007), who studied dynamic qualities of motion cues as opposite to different movement shapes that might convey emotional qualities. They recorded the same gesture performed in four emotions (happy, angry, sad, and pleasurable) and could identify two main movement characteristics that differentiated between them: Quantity of Motion (QoM), defined as "the overall measure of the amount of detected motion, involving velocity and force" (p. 47), and Contraction Index (CI), defined as "the contraction and expansion of the body [...], i.e., the minimum rectangle surrounding the body" (p. 47) in a two-dimensional space. Furthermore, Boone and Cunningham (1998) reported results of a dance study in which actors displayed four discrete emotions (happiness, sadness, anger, and fear). For the different emotions, the authors could identify distinct movement characteristics.

Besides studying music performers' and dancers' movement characteristics when communicating emotions, emotion communication and recognition through music-related movement by observers has also been investigated. Davidson (1993) conducted a study in which observers rated expressive movements of violinists and pianists. Results indicate that visual information more clearly communicated the expressive manner of the musician than sound alone and sound and vision presented together. Vines, Krumhansl, Wanderley, and Levitin (2006) examined clarinetists' abilities of communicating tension to observers and found that auditory and visual signals evoked different perceptions of tension, with the sound dominating the judgments. Additionally, their results indicated that the audiovisual presentation increased the perceived tension compared to the audio-only and video-only conditions, suggesting that participants integrated both signals. In the previously mentioned study by Dahl and Friberg (2007), they asked the observers in the first part of the experiment to rate which emotions they perceive when watching the visual elements of the performances. They could detect the happy, angry, and sad performances successfully, but failed with the fearful ones. Burger and Bresin (2010) developed a small robot displaying different emotions based on the movement cues found by Dahl and Friberg (2007). In a perceptual experiment, all emotions were successfully recognized by the observers. Sörgjerd (2000) investigated a clarinetist and a violinist who performed a piece of music with different emotional intentions, and reported that happiness, anger, sadness, and fear were better identified than tenderness and solemnity. However, no significant differences for the presentation condition (audio-only, movement-only, audio+movement) were found. Petrini, McAleer, and Pollick (2010) found that in audiovisual presentations of musicians playing with different emotional characteristics, the sound dominated the visual signal, both in emotionally matching and in emotionally mismatching stimuli. Furthermore, when participants were asked to focus on the visual information of the audiovisual stimuli, their ability to correctly identify the emotion decreased in case of emotionally incongruent stimuli, whereas it was unaffected by the video information, when participants were asked to focus on the audio information. Kaiser and Keller (2011) conducted a study in which they asked participants to rate short clips portraying different emotions (in a two-person dialog situation) that were accompanied with either emotionally compatible or incompatible music. They found that music affected the accuracy of the emotion ratings: happiness and sadness were perceived more accurately when combined with compatible than with incompatible music, while contentedness was rated vice versa. The perception of anger was found unaffected by the music.

Several perception studies have also focussed on dance and showed that dance movement could successfully communicate emotions to observers, both in regular video and in stick figure animations. In the previously mentioned study by Boone and Cunningham (1998), professional actors portrayed four discrete emotions by means of dance movements. Observers were able to successfully identify the emotions expressed in these movements. Dittrich, Troscianko, Lea, and Morgan (1996) conducted a study using short video clips of professional dancers aiming to convey fear, anger, grief, joy, surprise, and disgust. Observers could successfully recognize the portrayed emotions. Lagerlöf and Djerf (2009) asked professional dancers to improvise on different emotions (joy, anger, fear, and sadness) as solo dances. Observers were able to successfully determine the intended emotions in the movements displayed as video clips. Walk and Homan (1984) showed point-light displays of dancers portraying emotions to observers, who could effectively identify the emotions. Krumhansl and Schenck (1997) found that expressed emotions were low in section beginnings, but increased during sections and reached a peak short before sections ends. Dance and music could further convey similar emotion judgements.

This relatively large body of research regarding emotion-specific movement demonstrates the high relevance of emotions in human movement, in particular related to music. However, most dance-related studies used professional actors or dancers, indicating a lack of research dedicated to movement of nonprofessionals. Moreover, in these studies, the portrayal of emotions was intentional (i.e., the dancers/actors were asked to display certain emotions), thus those relationships between musical emotions and movement characteristics required further investigation when (layperson) dancers freely and spontaneously move to music without the conscious attention towards displaying a certain emotion expressed in the music. Besides the actual movement features characteristic for each emotion, the abilities of emotion communication and recognition in such movement was investigated for this thesis research. Links between musical emotions and movement responses could further provide support for the notion of embodied music cognition (Leman, 2007).

3.3 Personality

Above, music-induced movement was discussed in relation to music-related aspects, such as musical features, genre, and perceived emotion. However, individual differences, such as personality, can also be assumed to have an influence on movement. Personality is “the dynamic and organized set of characteristics possessed by a person that uniquely influences his or her cognitions, motivations, and behaviors in various situations” (Ryckman, 2008, p. 4). Links between personality and movement could be established on a general level; for example, people with dissimilar personality structures have been shown to move differently (e.g., Brebner, 1985; North, 1972). Koppensteiner and Grammer (2010) asked observers to watch animated stick figure videos of speeches given by politicians and rate the politicians' personality traits. Observers could associate personality traits with different movement qualities, including amount and flow of movement. Bechinie and Grammer (2003) developed a computer system producing a personality profile based on human dance movement (although without music). Furthermore, robots and agents haven been created that decode the personality types of the user based on gestures and movements for improving the interaction between robot and human (e.g., Ball & Breese, 2000), or that are provided with behaviors that observers can perceive as personality (e.g., Kim, Kwak, & Myungsuk Kim, 2008; Neff, Wang, Abbott, & Walker, 2010). Personality disorders have also been found to be reflected in bodily movement (Kluft, Poteat, & Kluft, 1986).

In terms of evaluating personality, the five-factor model, also called Big Five (John, Naumann, & Soto, 2008; McCrae & Costa, 1987) – measuring the five general traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience – is widely accepted (John et al., 2008) and considered to describe personality at the highest level of organization (Goldberg, 1993). Furthermore, it has been validated across cultures (McCrae, 2002; McCrae & Terracciano, 2005). There are several versions of Big Five questionnaires available, varying in length (between 10 and 240 items) and style (single adjectives vs. short phrases providing more context). This thesis research used the so-called Big Five Inventory (BFI), a 44-item version measuring the five traits using short phrases (John et al., 2008; John & Srivastava, 1999). Items were rated on 5-point scales ranging from strongly disagree (1) to strongly agree (5).

The five personality traits can be described as follows (e.g., John et al., 2008; McCrae, Gaines, & Wellington, 2012): Extraversion includes traits such as energetic, enthusiastic, positive, sociable, talkative, assertive, and outgoing. Agreeableness is defined as the tendency to be friendly, compassionate, trustful, modest, emphatic, and helpful, while being less anger-prone (Kuppens, 2005). Conscientiousness is associated with self-disciplined, controlled, dutiful, dependable, efficient, planned, and organized behavior. Neuroticism is related to the tendency to easily experience negative emotions, such as anger, anxiety, depression, vulnerability, and emotional instability. Openness to Experience can be defined as inventive, creative, imaginative, and curious, with an appreciation for arts, emotion, aesthetic experiences, adventure, unusual ideas, and novelty.

The Big Five personality traits have been investigated in a variety of music-related contexts. Rentfrow and Gosling (2003), for instance, have studied relationships between the personality traits and music preference. They found that extraverts had a preference for both "upbeat and conventional" and "energetic and rhythmic" music, while open participants showed a preference for "reflective and complex" as well as "intense and rebellious" music. Vuoskoski and Eerola (2011b) conducted a study investigating relationships between personality traits and emotion perception and preference in film music excerpts. They found that open participants liked sad and fearful stimuli, whereas extravert and agreeable participants preferred happy-sounding excerpts, with agreeable participants also preferring the tender excerpts, although disliking angry- and fearful-sounding examples. In another study, Vuoskoski and Eerola (2011a) investigated personality and felt emotions to music. Here, extraversion was positively related to experienced happiness, sadness, and tenderness, while Openness was positively connected with felt tenderness and sadness, and Agreeableness with tenderness. G. Dunn (2009) related preference ratings of music to personality and audio features and could, for example, find that Excitement-Seeking (a facet of Extraversion) was positively related to music with a greater number of percussive events and a preference for rap. Aesthetics (a facet of Openness to Experience) was positively related to tonal complexity and the liking of jazz, while participants lower in Aesthetics preferred simpler tonal complexity as found in pop music. In another study, P. G. Dunn, Ruyter, and Bouwhuis (2012) supported the findings that personality traits are connected with listening behavior: results indicated that Neuroticism is positively related to a preference for classical music, and Openness to Experience is positively related to a preference for jazz music. Fink et al. (2012) conducted a study in which they gathered Big Five personality data from male participants and asked them to dance to a basic rhythm. Subsequently, women rated the stick figure animations related to dance quality, yielding positive correlations between movement of conscientious and agreeable men and women's perception of dance quality. Luck, Saarikallio, Thompson, Burger, and Toiviainen (2012) conducted a perceptual experiment on attraction on the dance floor and studied how the personality of both dancer and observer influenced the observers' attractiveness ratings of the dancers' movements presented as stick figure animations. The results of the study suggest that personality traits shape interpersonal attraction on the dance floor.

Given the connections between personality and movement, as well as between personality and various music-related settings, it seems reasonable to assume that there are relationships between personality and music-induced movement. As music is effective in eliciting body movement, it could be (besides the previously mentioned music-related characteristics) that the personality traits of a person influences the way she moves to music.

4 AIMS AND OVERVIEW OF THE THESIS

The main aim of this thesis is to investigate different factors that were considered to be influential on music-induced movement. The factors studied in this thesis included musical features, perceived emotions in music, genre, and personality traits of the dancers. A schematic outline of the topics and studies is given in Figure 1.

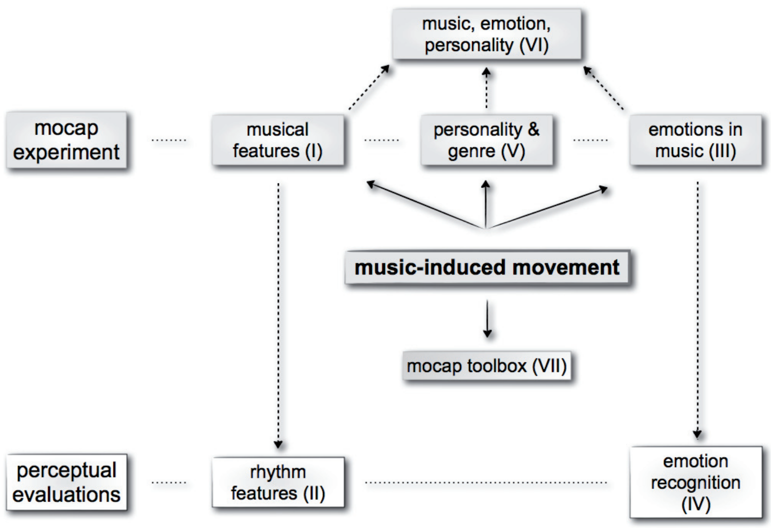


FIGURE 1 Schematic overview of the studies included in this thesis (the numbers in the brackets refer to the studies).

Two main experiment designs were employed in order to shed light on the basic research aim from different viewing angles: motion capture recordings and perceptual (rating) evaluations. The motion capture study used an infrared marker-based optical motion capture system to quantitatively study the movements that participants performed when being ask to move to 30 different popular-music stimuli of 30 seconds length each. Studies I, III, V, and VI were based on this motion capture study, covering effects of musical features (I), perceived emotions in music (III), personality, and genre (both V) on characteristics of music-induced movement. Studies I, III, and V investigated relationships between movement features and each of the factors separately. However, relationships might also exist between the different factors, thus Study VI describes two different multivariate statistical approaches to investigate more complex relationships than the dual ones described in Studies I, III, and V.

Studies II and IV feature perceptual evaluations as their main experiment design. They were conducted to investigate participants' ratings for audio stimuli regarding perceptual qualities of a certain musical feature (II) – the spectral flux – and participants' ratings for audio and video stimuli regarding emotion communication (IV). Study II evaluated the relationship of spectral flux to perceived rhythm and the desire to move, since Study I found that the spectral flux of low and high frequency bands was important for music-induced movement assuming it would be related to rhythmic properties of the music. Study IV describes a rating study regarding the possibilities of audio and movement conveying emotions to observers. This study is related to Study III and can be seen as a follow-up experiment, as Study III found movement characteristics to be related to perceived musical emotions.

Study VII does not describe a research study but a research tool, the MoCap Toolbox, a Matlab[2] toolbox for analysis and visualization of motion capture data. The author of this thesis has been involved in the advancement and support of the project for the time of her doctoral studies and has furthermore developed several features related to data analysis and visualization that were implemented in the toolbox. The extraction of the movement features used in Studies I, III, V, and VI has been carried out with MoCap Toolbox.

A secondary aim of this thesis was related to the development of computational features that describe movement in a meaningful way. This thesis tried to go beyond features based on plain velocity and acceleration calculations and aimed at developing higher-level descriptors not only applicable to music-induced movement, but also to other music-related activities, such as performing music, or even non-musical movement, for instance, related to non-verbal communication.

[2] www.mathworks.com

5 METHODOLOGIES

The purpose of this chapter is to introduce the different methodologies and approaches adopted in this thesis research. First, motion capture technology as the methodological approach employed for recording and analyzing movement data is elaborated in detail. Furthermore, information regarding the computational extraction of musical features, as well as the experiment designs and analysis methods utilized in the different studies are given.

5.1 Motion capture

In order to conduct research on human movement, it is necessary to employ methods that precisely describe the motion. This can be achieved using various approaches, ranging from gathering still pictures of body postures with a photo camera to capturing continuous movement with high-speed video or infrared cameras or non-visual techniques that consider orientation, acceleration, or force. Besides sensing technologies to capture the motion, computer technology is commonly used to process or store the data and represent the movement data in numerical form to allow quantitative analysis or real-time processing. Thus, *motion capture* (mocap) usually refers to motion representations in digital formats.

Motion data can be represented in several ways. Position/displacement data can be measured in two dimensions (i.e., on a plane) or in three dimensions (i.e., in a space). Additionally, motion data can be described in terms of three-dimensional orientation (referring to an object and its orientation in space). Such data is usually denoted in the form of six degrees of freedom (6DOF), that is 3-dimensional position and 3-dimensional orientation data.

There is a wide range of sensing and motion capture technologies available, ranging from inexpensive hand-held devices, such as mobile phones and game controllers, to high-precision tracking instruments used in film and game industries as well as in research, mainly in biomechanics and sports. The technologies mostly used in music-related research and applications comprise inertial, magnetic, mechanical, and optical (camera-based) tracking systems. The following list provides a short overview of the different systems.

- *Inertial systems:* accelerometers and gyroscopes to measure acceleration and orientation/rotation
  - Affordability: inexpensive sensors, as e.g., implemented in smart phones or game controllers; high-end systems (e.g., whole-body suits) expensive
  - Usability: no direct visibility needed (sensors can be hidden); unaffected by light conditions or magnetic fields; requires sensor for each joint of interest; body suit potentially motion-restricting, customized
  - Portability: yes, light and small sensors and self-contained system
  - Precision: the higher quality the better; drift (position error), cumulating during time/length of data capture
  - Data: 6DOF derivative data, yielding drift when integrating
- *Magnetic systems:* position and orientation of objects measured in a magnetic field
  - Affordability: expensive
  - Usability: no direct visibility needed; unaffected by light conditions; affected by other magnetic fields
  - Portability: yes
  - Precision: high accuracy in near-field; distortion of other magnetic fields and in larger areas
  - Data: 6DOF
- *Mechanical systems:* gears, potentiometers, and bend sensors measuring angles between joints
  - Affordability: the better, the more expensive
  - Usability: no direct visibility needed; unaffected by light conditions or magnetic fields; can be worn on the body (e.g., suit or glove) or detached as external controller device; require a sensor for each joint of interest; body suit potentially bulky, motion-restricting, customized
  - Portability: yes
  - Precision: high
  - Data: angular
- *Optical systems:* camera-based systems using light (usually either light in the visible or in the infrared spectrum) and different imaging techniques
  - Affordability: wide range from low-quality, inexpensive devices (e.g., a web camera) to high-end, expensive systems (e.g., infrared motion capture)
  - Usability: unaffected by magnetic fields; partly affected by light conditions; direct visibility needed; limited capture space restricted by the cameras – in case of infrared systems: reflective markers, allowing flexible configurations, easy to attach, though might constrain or fall

- Portability: yes, but restricted (since large equipment)
- Precision: the higher quality, the better
- Data: 2- or 3-dimensional position data with regular video cameras; with infrared cameras also 6DOF

In the following, infrared camera- and marker-based optical motion capture technology will be emphasized, as this was the technique used for capturing movement data in this thesis research. In the progression of this thesis summary, the term motion capture (or mocap) will therefore refer to infrared marker-based optical motion capture.

Infrared cameras work with light detectable in the infrared part of the electromagnetic spectrum. Such light has longer wavelengths and lower frequencies than visible light. Infrared optical tracking is based on an active source emitting pulses of infrared light at a very high frequency. In a passive system, as the one used for this thesis research, this emitting source is usually attached to the camera (e.g., LEDs arranged as a ring around the camera lens), which in turn captures the reflections of this light produced by small, usually spherical markers attached to the tracked object. These markers are placed on points of interest, such as characteristic joints of the human body. Section 5.1.1 will describe marker placements in more detail. The temporal resolution of the system is specified by the rate of the infrared light pulses. Standard capture speeds in music-related applications range from 60 to 240 frames per second, which is sufficient for a wide range of music-related movement, such as playing most instruments and dancing[3]. Optical systems require direct visibility of the tracked objects (i.e., if a marker is hidden, e.g., by clothes, it will not be recorded). Each camera captures a two-dimensional image, commonly with white pixels indicating the markers on a black background. If several cameras are set up in a network around the capture space, data in three dimensions can be obtained (the system can construct the 3D-location of a marker as soon as two cameras detect it). The use of several cameras also increases the field of view and thus minimizes the occurrence of hidden markers. It also reduces the measurement error (at least in some cases). More details regarding data capturing and the system that was used for conducting the studies reported in this thesis is given in the next section.

5.1.1 Motion capture data collection

The data used for the studies of this thesis were collected in the Motion Capture Lab hosted by the Music Department of the University of Jyväskylä, Finland. At the time of the data collection, the lab featured eight Qualisys ProReflex MCU120 cameras mounted on the walls around the capture space at a height of approximately 2.50 to 3 m. A depiction of the lab is provided in Figure 2.

For each camera, the orientation (based on the mounting), focus, and aperture can be adjusted to achieve an appropriate capture space in which markers

[3] Some systems can capture several thousand frames per second. However, a higher frame rate increases the amount of data without necessarily describing the movement more accurately, as human movement cannot exceed a certain tempo.

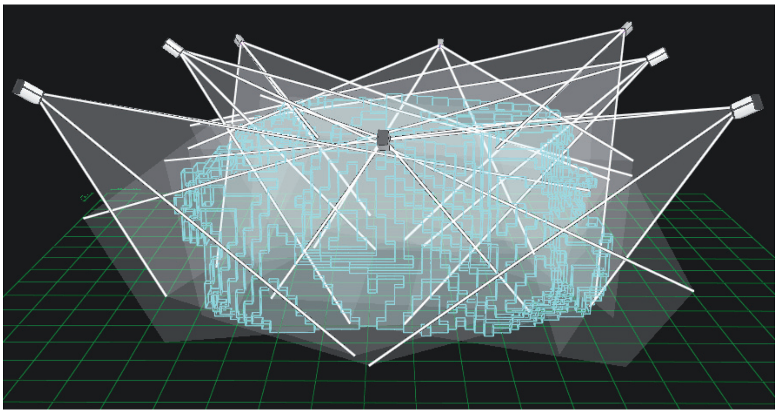


FIGURE 2 Depiction of the motion capture lab showing the camera cones of each of the eight cameras and the covered volume after calibration.

can be detected well. For capturing data in three dimensions, the first step of a data collection is to calibrate the system, so that it obtains the exact position and orientation of each camera with respect to the others and the floor in order to create a three-dimensional representation of the capture space (see covered volume depicted in Fig. 2). As a result, the calibration provides origin and orientation of the Cartesian coordinate system relative to which the locations of the markers are determined, calculated using a method called direct linear transformation (Robertson, Caldwell, Hamill, Kamen, & Whittlesey, 2004), depicted in Figure 3(A).

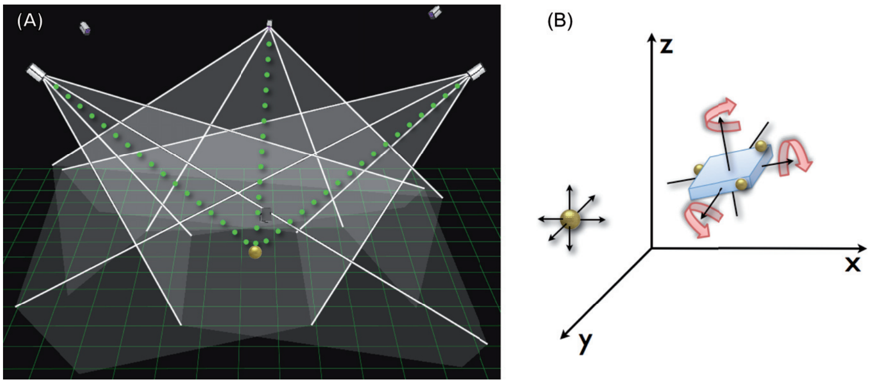


FIGURE 3 (A) As the field of view, location, and orientation of each camera is known, the system can calculate the location of a marker in 3D based on its 2-dimensional location detected by each camera, indicated by the three dotted beams. The 3D coordinates are determined as the point in which the beams intersect. (B) Marker with three degrees of freedom (yellow, left side) and rigid body (of three markers) with six degrees of freedom (blue, right side).

This approach yields time-series data representing the displacement of the markers in three dimensions (x, y, and z). If additionally the orientation of objects are of interest, so-called rigid bodies can be employed. Rigid bodies consist of at least three markers that are mounted on an object as a fixed cluster (i.e., having stable distances and angles). These markers establish their own coordinate system with the center of the marker cluster as origin and track 6DOF data (three position coordinates yielding displacement in the capture space and three rotation angles yielding orientation in the capture space). Both displacement of a marker in three dimensions as well as displacement and orientation of a rigid body with six degrees of freedom are conceptualized in Figure 3(B).

The ProReflex system can record at a maximal speed of 120 frames per second, which was used in the data collection for the studies described in this thesis. As general characteristics of music-induced movement were of greatest interest, this capture speed was sufficient and provided a good combination of amount of data and information depth.

In order to capture whole-body movement, reflective markers are placed on characteristic joints of the human body, such as head, arms, or legs. Twenty-eight markers were used in the motion capture data collection conducted for this thesis research. They were attached to the following body parts of each participant (the locations are displayed in Figure 4(A) and (B)): (L = left, R = right, F = front, B = back): 1: LF head; 2: RF head; 3: LB head; 4: RB head; 5: L shoulder; 6: R shoulder; 7: sternum; 8: spine (T5); 9: LF hip; 10: RF hip; 11: LB hip; 12: RB hip; 13: L elbow; 14: R elbow; 15: L wrist/radius; 16: L wrist/ulna; 17: R wrist/radius; 18: R wrist/ulna; 19: L middle finger; 20: R middle finger; 21: L knee; 22: R knee; 23: L ankle; 24: R ankle; 25: L heel; 26: R heel; 27: L big toe; 28: R big toe.

Optical motion capture systems basically record all reflections they detect without "knowing" what they represent. Thus, each marker needs to be correctly identified and labeled accordingly during or after the recording process. This procedure varies for different motion capture systems. In some systems, such as Vicon[4] and Optitrack[5], the user manually creates a marker model of the person to be tracked, which is subsequently applied during the capture to automatically identify the markers. In case the automatic approach fails, the user continues to label manually. With the Qualisys system, the user first captures and then annotates the data manually. A manually annotated capture can be used for creating a model to automatically label subsequent captures. This can be done either after the first capture, so that the model can be applied during the tracking of subsequent captures, or after all captures are made. Three main challenges can be mentioned regarding the process of labeling: 1) Markers might get occluded during capture, so the same marker might be split into several parts (called trajectories), which need to be combined during labeling. 2) If markers come too close to each other during the capture, the system might switch them, resulting in two markers being assigned to the same trajectory, which requires separation during labeling. This might also happen when one or both markers were hid-

[4] www.vicon.com
[5] www.naturalpoint.com/optitrack

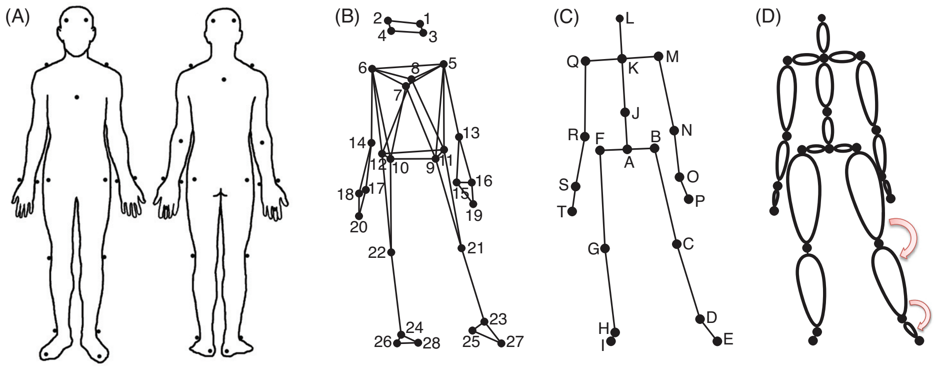


FIGURE 4 Marker and joint locations. (A) Anterior and posterior view of the marker placement on the participants' bodies; (B) Anterior view of the marker locations as stick figure illustration; (C) Anterior view of the locations of the secondary markers/joints after a marker-to-joint transformation; (D) Segment representation. Each segment is characterized by its parent segment, length, rotation, angle to the parent segment (indicated by the arrows), and a mass and moments of inertia specified by the Dempster model (Dempster, 1955; Robertson et al., 2004).

den for a short period of time. 3) Reflections might cause extra markers that the system tries to match with the actual markers. These double markers need to be removed during labeling.

If a marker was occluded during the capture, these frames will be missing, resulting in a gap. Gaps can be filled using different inter- or extrapolation methods[6]. These methods apply different algorithms to estimate the missing frames based on the data points before and after the gap. Usually, some kind of polynomial (e.g., spline) interpolation techniques are used for gap-filling. The simplest way of polynomial interpolation is linear interpolation, in which the two data points before and after the gap are connected with a straight line. More advanced methods of polynomial and spline interpolation use higher-order polynomial functions to fill gaps. Gap-filling works well for short gaps. However, with an increasing length it becomes more difficult to estimate the gap mathematically, so gap-filling might lead to inaccurate results. Gap-filling can be either applied in the motion capture software used for recording, such as the Qualisys Track Manager (QTM), or at a later stage of the analysis.

Although motion capture applications generally offer a few methods for data analysis, the data is usually exported after the labeling is completed and subsequently imported into a more generic analysis environment, such as Matlab. The movement analysis performed for the studies of this thesis was con-

[6] Interpolation refers to procedures where missing data between two points of the capture are determined, whereas extrapolation is used in cases when data is missing in the beginning or in the end of a capture (i.e., only one captured data point is available).

ducted using the *MoCap Toolbox*, which will be explained in more detail in the next section.

5.1.2 Mocap Toolbox

The tool that was used for analyzing the movement data was the MoCap Toolbox, a set of function written in Matlab for analyzing and visualizing movement data. Publication VII included in this thesis describes structure, functionality, and use of the toolbox in detail. It has been developed for analyzing music-related movement, but might be useful for other research areas as well. Users can freely adopt the functions according to their needs, as the toolbox code is written in Matlab syntax and available as open source[7]. Additionally, users can utilize the functionality that Matlab offers as a generic (scientific) computing environment, to further analyze the features extracted with MoCap Toolbox while staying in the same environment.

There are closed-source software solutions available for motion capture analysis and visualization, such as Visual3D[8] or MotionBuilder[9], and the applications that primarily used for data recording, such as QTM and Vicon Nexus[10]. However, such applications are usually too limited in their functionality, too focussed on visualization, and/or too restrictive in adapting to the needs of the researcher, such as to develop movement features useful for their individual research questions. To overcome these issues, this toolbox was implemented in Matlab as an open-source environment.

The MoCap Toolbox supports several motion capture data formats, in particular the .c3d file format (which, e.g., Vicon or OptiTrack optical motion capture systems can produce), the .tsv and the .mat format, both produced by the Qualisys motion capture system, and the .wii data format produced by the WiiDataCapture software (also developed at the Music Department of Jyväskylä University and downloadable for free from the same internet address as the MoCapToolbox – see Footnote 7).

The current version of the toolbox (version 1.4) provides 64 functions for analyzing and visualizing mocap data. The main categories can be summarized as data input and edit functions, coordinate transformation and coordinate system conversion functions, kinematic and kinetic analysis functions, time-series analysis functions, visualization functions, and projection functions. Furthermore, it uses three different data structures: the MoCap data structure, the norm data structure, and the segment (segm) data structure. To convert between the different data representations and enable certain visualizations, parameter structures are used. A MoCap data structure instance is created when mocap data is read from a file to the Matlab workspace (using the function `mcread`). A MoCap data structure contains the 3-dimensional locations of the markers as well as basic in-

7 downloadable from www.jyu.fi/music/coe/materials/mocaptoolbox
8 www.c-motion.com/products/visual3d/
9 www.autodesk.com/motionbuilder
10 www.vicon.com/products/nexus.html

formation, including the type of structure, the file name, number of frames of the recording, the number of markers in the data, the frame rate, the names of the markers, and the order of time differentiation of the data. The marker representation reflects the actual maker locations. It can be transformed into a joint representation that is related to locations derived from marker locations. A joint can consist of one marker, but it can also be derived from more than one markers. It can, for example, be used for calculating the location of a body part where it is impossible to attach a marker. The mid-point of a joint, for instance, can be then derived as the centroid of four markers around the joint. The norm data structure is similar to the MoCap data structure, except that marker/joint data contains the Euclidean norm of the data from which it was derived. If, for instance, the norm is applied to velocity data, the resulting data structure holds the magnitude of the velocity, or speed, of each marker. The third data structure, the segm data structure, is not like the other two related to points in space (markers or joints), but to segments of the body (see, e.g., Robertson et al., 2004). Body segments are commonly simplified as a chain of interconnected rigid bodies that have physical characteristics, such as segmental masses and locations of the segmental centers of gravity, and inertial properties of the segmental mass, required when performing kinetic analysis. Body-segment parameters have been quantified in various ways, for instance by measuring human cadavers as done by Dempster (1955). For modeling body segments, Dempster's parameters have been implemented in the toolbox. A segm data structure instance is generated by transforming a joint representation to a segment representation. Most fields of a segm data structure are similar to the ones of a MoCap data structure, however, the data field is replaced by four other fields. These fields store information on how the joints are connected to form segments, and about the location and orientation of the center of the body (called root) and of the segments in several ways (as Euclidean vector pointing from the proximal to the distal joint of the segment, the length of each segment, the rotation as a quaternion representation (Hanson, 2006), and the angles between two adjacent segments). The conception of the segment representation is depicted in Figure 4(D).

After importing a motion capture file into Matlab, any missing frames can be filled using linear interpolation. Visualizing and animating data is a useful approach to get a first overview of the data. Data can be visualized as time-series or as single frames. Time-series plots are provided in Figure 5, whereas single-frame marker visualizations were given in Figure 4(B) and (C). For single-frame visualizations, a parameter structure can be employed to provide settings for the visualization, such as colors, connector lines, and marker and line sizes. To create an animation from a recording, a series of single frames are produced that are afterwards compiled into a movie.

Kinematic variables, such as velocity and acceleration, are estimated using the time derivatives of the position data (with `mctimeder`). The function calculates either the differences between two successive frames and then uses a Butterworth filter for smoothing or derives the data with a Savitzky-Golay smoothing

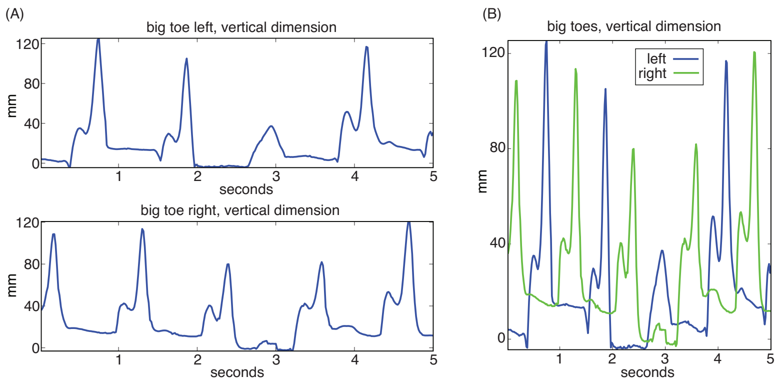


FIGURE 5 Time-series plots of mocap data. (A) Toe markers plotted as separate plots; (B) Both toe markers plotted into the same plot.

filter (Savitzky & Golay, 1964). The first implementation is considerably faster than the second, whereas the second yields slightly more accurate results.

Furthermore, MoCap Toolbox offers several functions to analyze movement data, such as calculating the cumulative distance travelled by a marker, the bounding rectangle (the area covered by the movement), the distance between two markers, or the periodicity of movement. Kinetic variables, such as the kinetic energy of different body parts, can be estimated based on the body-segment representation. A full overview of the functions can be obtained in the manual of the toolbox.

The MoCap Toolbox was intended as a tool to analyze music-related movement, but has attracted the attention of researchers working in various fields. Besides for music-related research, it has been downloaded for face recognition, sports, gait, or biomechanics research, as well as for artificial intelligence research, such as robotic motion, human-robot interaction, and machine learning. Since its first launch in 2008, MoCap Toolbox has continuously been extended by the main developers as well as by the users who provided valuable feedback. The following section describes the movement features extracted with MoCap Toolbox that were used in the studies presented in this thesis.

5.1.3 Movement feature extraction

This section describes the movement features derived for the mocap studies (I, III, V, and VI) using MoCap Toolbox. For preprocessing the mocap data, the first step after importing and gap-filling was to synchronize the data with the musical stimuli and trim the data according to the length of the stimuli. Following this, a set of 20 secondary markers or joints was derived from the original 28 markers. The markers were reduced, since some markers were redundant for the subsequent analysis (in any thesis-related analyses, there was, for instance, no interest in the orientation of the head). Furthermore, these 20 joints are required for conducting body segment modeling. The locations of the 20 joints are depicted in Figure

4(C). The locations of joints C, D, E, G, H, I, M, N, P, Q, R, and T are identical to the locations of one of the original markers, while the locations of the remaining joints were obtained by averaging the locations of two or more markers; joint A: midpoint of the four hip markers; B: midpoint of markers 9 and 11 (left hip); F: midpoint of markers 10 and 12 (right hip); J: midpoint of breastbone, spine, and the hip markers (midtorso); K: midpoint of shoulder markers (manubrium), L: midpoint of the four head markers (head); O: midpoint of the two left wrist markers (left wrist); S: midpoint of the two right wrist markers (right wrist). Joint A (middle of the hip, also called root) and joint L (head) are examples for several markers being used to calculate a virtual marker at a location impossible to place a marker. The joint representation was additionally transferred to a segment representation using Dempster's parameters.

The joint representation was next transformed to a local coordinate system with joint A located in the origin, and segment BF had zero azimuth (i.e., identical to the x-axis of the coordinate system), so that the movement was expressed relative to the center of the body. The movement features were estimated based on the local coordinate system, unless indicated differently. Next, for both the local and the global joint representations, the instantaneous velocity, acceleration, and jerk were estimated using numerical differentiation based on the Savitzky-Golay smoothing FIR filter (Savitzky & Golay, 1964) with a window length of seven samples and a polynomial order of two. These values were found to provide an optimal combination of precision and smoothness in the time derivatives. These kinematic variables were estimated for the three-dimensional data as well as for each of the dimensions separately.

From this basic data, several movement features related to different body parts and movement qualities were extracted and used in the different studies. Besides topic-related selections of features (i.e., previous research suggesting certain features), another aim of this thesis was to develop and introduce meaningful movement features and descriptors that are potentially useful and applicable in different kinds of movement-related research. The extracted features comprised postural features, global and local kinematic features, and kinetic features.

- Postural features
  - Distance between both hands (joints P and T), both elbows (joints N and R), and both feet (joints E and I).
  - Torso Tilt, the vertical tilt of the torso (joints A-K) with positive tilt being related to bending forward.
- Kinematic features
  - Magnitude of Velocity (Speed) of center of mass (joint A), head (joint L), both hands (joints P and T), and both feet (joints E and I). The center of mass used the global joint data (since the hip was set to zero in the local data).
  - Magnitude of Acceleration of center of mass (joint A), head (joint L), both hands (joints P and T), and both feet (joints E and I) (global joint

data in case of the center of mass).

- Magnitude of Jerk of center of mass (joint A), head (joint L), both hands (joints P and T), and both feet (joints E and I) (global joint data in case of the center of mass).
- Amount of Movement as based on the travelled distance of all markers.
- Area of Movement (i.e., bounding rectangle), defined as the smallest rectangle that contains the projection of the trajectory of the center of mass (joint A) on the horizontal plane (i.e., floor), averaged across a four-second analysis window with a two-second overlap. This feature used the global joint data.
- Fluidity, an overall movement fluidity/smoothness measure based on the ratio of velocity to acceleration. The combination of high velocity and low acceleration reflects fluid movement, whereas the combination of low velocity and high acceleration reflects non-fluid movement.
- Rotation Range, defined as the amount of rotation of the body (joints M and Q) around the vertical axis.
- Movement Complexity, based on principal components analysis (PCA) of the position data of all joints. PCA, in this case, can be understood as a decomposition into different (independent) movement components. The PCA was performed on individual subject level, i.e., one PCA per participant and stimulus. Subsequently, the cumulative sum of the proportion of (explained) variance contained in the first five PCs was determined. For five PCs, the cumulative variance was found to have the highest range of variation across movement recordings, so using five PCs discriminated best between complex and non-complex movement. The proportion of variance explained by the first five PCs varied between 67.0% and 99.8%, with an average of 85.5% (subject level). For an excerpt with 99.8% explained variance, almost all movement can be explained by the first five components, whereas for an excerpt with 67.0% explained variance, 33% cannot be explained by the first five components (and more components are needed). Thus, a high proportion of unexplained variance (i.e., a low cumulative sum) would mean that the underlying movement is complex, as a high number of PCs is needed to explain the movement sufficiently. A low proportion of unexplained variance (i.e., a high cumulative sum), on the other hand, implies simpler movement, as it can be sufficiently explained with a low number of PCs. Figure 6 illustrates this feature by displaying projections of the first five principal components of complex (A) and non-complex (B) movements.
- Hip Wiggle, defined as the absolute angular velocity of the hips (joints B and F) around the anteroposterior (front-back) axis (global joint data).
- Shoulder Wiggle, defined as the absolute angular velocity of the shoulders (joints M and Q) around the anteroposterior axis (global joint data).

- Kinetic features
  - Kinetic Energy, the energy contained in the body when moving (using Dempster's (1955) parameters).

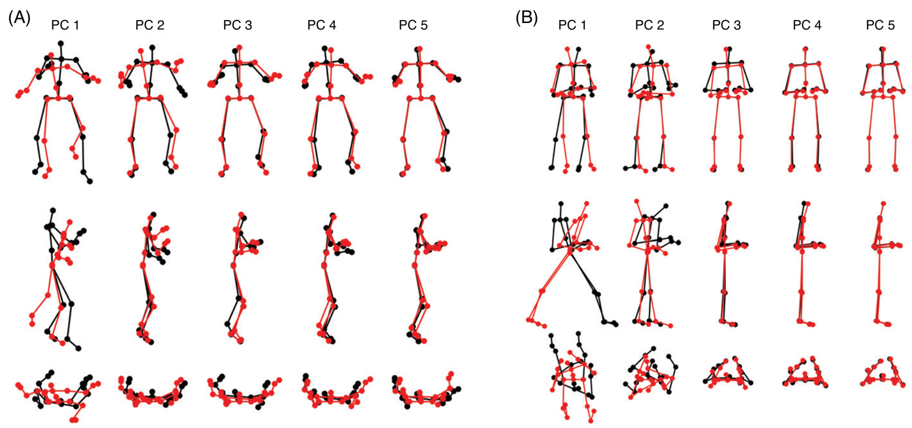


FIGURE 6 Movement Complexity illustrated by displaying the projections of the first five principal components of complex movement (A) and non-complex movement (B). The extreme deflections of the projections/dimensions are plotted from the front, from the side, and from the top. (A) High Movement Complexity, as movement is visible in all five PC projections, thus a high number of principal components is needed to explain the movement; (B) Low Movement Complexity, as most movement is found in the projection of PC 1 – a low number of PCs can sufficiently explain this movement.

In order to conduct further analysis, the instantaneous values of each movement feature were averaged across each stimulus presentation (time) yielding one value per movement feature for each participant and stimulus. In case of the kinetic energy, the standard deviation was also calculated besides the mean. Such an approach of calculating 'static' movement features characterizing a movement progression might be oversimplifying as it disregards any temporal evolution. However, emphasis was put more towards finding broad descriptors accounting for the whole movement sequence, thus one-value descriptors seemed appropriate as opposed to any time-series analysis approach.

5.2 Musical feature extraction

This section describes the musical feature extraction process used in Studies I, II, and VI. The features were obtained by employing the Matlab MIRToolbox (version 1.4) (Lartillot & Toiviainen, 2007), a Matlab toolbox for music information retrieval. The stimuli were imported to Matlab and subsequently analyzed with different functions provided by the MIRToolbox. The extracted features included pulse-, rhythm-, attack-, dynamics-, and timbre-related features. Some features were based on the waveform representation of the signal (time domain), whereas others were based on the spectral representation of the signal (frequency domain). For several features, the waveform was first divided into shorter time spans called frames, for which the features were subsequently calculated.

Pulse Clarity indicates the strength of rhythmic periodicities and pulses in the signal, estimated by the relative Shannon entropy of the fluctuation spectrum (Pampalk, Rauber, & Merkl, 2002). In this context, entropy can be understood as a measure of the degree of peakiness of the spectrum. Music with easily perceived beat has a distinct and regular fluctuation spectrum, indicating low entropy. Thus, high pulse clarity is associated with low fluctuation entropy. To illustrate this measure of Pulse Clarity, fluctuation spectra for high and low Pulse Clarity are shown in Figure 7.

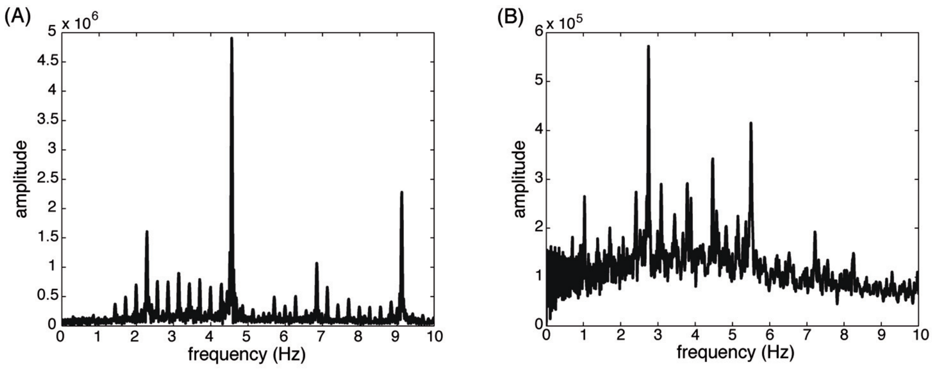


FIGURE 7 Fluctuation spectra of two stimuli used in the studies. (A) Peaks at a regular distance of 0.28 Hz, with the highest peak at 4.56 Hz and other clear peaks at 2.29, 6.85, and 9.13 Hz, suggesting clear pulses and periodicity (stimulus 1, see Appendix 1). (B) Markedly lower magnitude values, a less periodic pattern of peaks, and more noise, suggesting low Pulse Clarity (stimulus 21, see Appendix 1).

Spectral Flux indicates to which extent the spectrum changes over time, reflecting rhythmic properties of the music, such as strength and prominence. It is computed by calculating the Euclidean distances of the spectra for each pair of consecutive frames of the signal (Alluri & Toiviainen, 2010), using a frame length of 25 ms and an overlap of 50% between successive frames. For the calculation of Spectral Flux of different frequency bands, the stimulus was first divided into 10 frequency bands in the range of 0-22050 Hz, with the first band reaching from 0-50 Hz and the remaining bands containing one octave each. The Sub-band Flux, with the same procedure as described above, was then calculated for each of these ten bands separately. Three different frequency bands were used in studies included in this thesis: Studies I and VI employed sub-bands no. 2 (50-100

Hz) and no. 9 (6400-12800 Hz), whereas Study II used these two sub-bands and additionally sub-band no. 6 (800-1600 Hz). Two spectrograms of sub-band no. 2 are displayed in Figure 8 to show the difference between high and low amounts of Sub-band Flux.

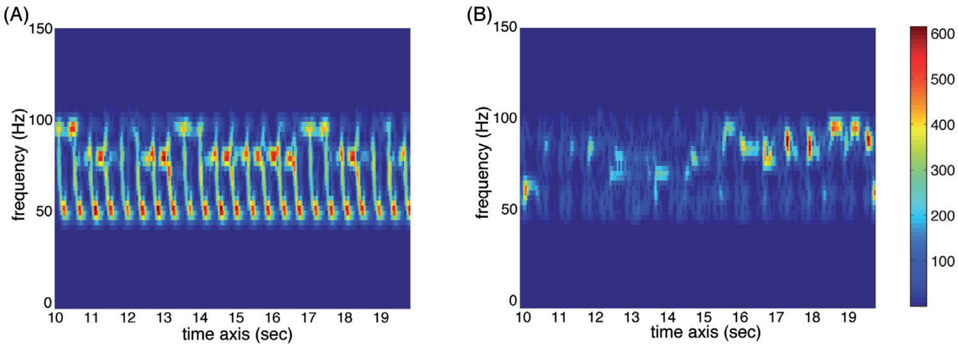


FIGURE 8 Spectrograms (sec. 10-20) of sub-band no. 2 (50-100 Hz) of two stimuli used in the study. (A) High amount of temporal change (red represents high energy at the respective time and frequency, whereas blue represents low energy; see color bar) resulting in high value for Sub-band Flux (stimulus 26, see Appendix 1). (B) Low amount of temporal change resulting in low Sub-band Flux (stimulus 5, see Appendix 1).

Percussiveness is an attack-related feature. The steeper the slope, the more percussive and accentuated the sound. To calculate it, the slopes of the amplitude envelopes at note onsets are estimated.

Low Energy is a dynamics-related feature, defined as the proportion of time during which the sound energy is below the average sound energy. Vocal music with silences, for instance, would result in high values, whereas continuous strings have low values of Low Energy (Tzanetakis & Cook, 2002). The feature was calculated for successive frames of the signal having a length of 25 ms and an overlap of 50%.

The Spectral Centroid is a timbral feature, specified as the geometric center of the amplitude spectrum. It is commonly associated with brightness of sounds (McAdams, Winsberg, Donnadieu, Soete, & Krimphoff, 1995). The feature was calculated for successive frames of the signal having a length of 25 ms and an overlap of 50%.

Number of Onsets denotes the sum of note onsets detected in the stimulus. A high number of onsets is related to high sound density.

Furthermore, the Tempo of the stimuli was assessed in a separate tapping experiment, in which ten participants tapped to the 30 stimuli. The average tempo was evaluated by taking the median value of all intertap intervals.

Apart from Tempo, Number of Onsets, and Pulse Clarity, all features resulted in time-series that were, subsequent to their calculations, averaged across time to match the movement features that were, as noted above, averaged across time as well. The musical stimuli were relatively stable in terms of their musical characteristics, so the averaged features should adequately reflect the main attributes of the stimuli.

5.3 Experiment designs

Basically, two different experiment designs were utilized for conducting the studies reported to this thesis. Studies I, III, V, and VI were based on a motion capture study, in which participants were individually recorded while moving to different musical stimuli. Studies II, III, and IV utilized perceptual experiments asking participants to rate various stimuli. In studies II and III, participants were ratings audio stimuli, whereas in Study IV, participants rated audio, silent video and audio-video stimuli. In Studies II and IV, each participant performed the experiment individually. Both play-back of the stimulus and indication of the ratings were organized in a graphical user interface created in Max/MSP 5, a graphical programming environment. Study III employed a data collection of perceptual data in group-form. The stimuli were played back through a Max/MSP 5 patch, while the ratings were given on paper. In all experiments, stimuli were presented in random order.

5.4 Analysis methods

This section summarizes the different analysis methods used in this thesis.

As a first step for the studies, in which analysis related to the musical stimuli (e.g., musical features) was performed (Studies I, II, III, and IV), the consistency between the participants' results (e.g., perceptual ratings or movement characteristics) was assessed. This was achieved by calculating intraclass correlations (Shrout & Fleiss, 1979) for the movement features and the perceptual ratings used in each of the studies. This approach not only gives a first insight into the general (rating/movement) behavior of the participants, but can also be taken as a justification for averaging across participants, as this was the main aim of assessing inter-participant consistency. However, such averaging is only reasonable, if the consistency between participants is considerably high.

In studies I, II, and III, correlations were used to investigate the strength of linear relationships or dependencies between two variables (utilizing Pearson product-moment correlation coefficients), for instance between musical and movement features or between perceptual ratings and movement features. Studies I and III consisted of a large number of correlations (32 and 60 respectively). Therefore, the rate of (potential) false discoveries was controlled by applying Bonferroni correction in Study I and Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) in Study III.

In Study I, a series of ANOVAs (analysis of variance) was performed to investigate the role of tempo on music-induced movement features. It was assumed that tempo would be non-linearly related to movement features, as research on preferred tempo rates (see Section 3.1.1) suggested that mid-range tempi would stimulate movement most. Therefore, the five slowest, the five medial, and the five fastest stimuli were selected to be included in a series of ANOVAs to examine if these three groups behave differently with respect to the movement features.

Study IV employed several repeated measures ANOVAs on perceptual rating data to investigate differences in the perception of different emotions in different presentation modalities (audio, video, and audio-video). The significance level of the subsequent post-hoc tests was adjusted using Bonferroni correction to account for multiple comparisons.

A series of linear regression analyses was conducted in Study IV to predict the audio-video ratings from the audio ratings and the video ratings to investigate which modality had greater influence on the audio-video ratings.

Study IV additionally employed principal component analysis (PCA) to reduce the four-dimensional discrete emotion space (i.e., ratings for Happiness, Anger, Sadness, and Tenderness) to a two-dimensional space. Applying Kaiser's criterion, two components were retained that could account for 77.8% of the total variance. Subsequently, the components were rotated using varimax rotation to maximize the loadings of the variables on the components (i.e., to more evenly distribute the explained variance across the components).

PCA was also employed in Study V to demonstrate a method for reducing the dimensionality of a large movement feature space containing potentially correlated features by converting them into a smaller space of linearly uncorrelated features. Using the Kaiser criterion, five principal components (PCs) were retained, accounting for 90.3% of the variance of the original features. Subsequently, the PC solution was rotated using varimax rotation to maximize the loadings of the variables on the components.

Subsequent to the PCA, Study V employed a multivariate analysis of variance (MANOVA) to investigate the effects of the level of personality trait (low scorers vs. high scorers) and genre of the music (six genres: rock, techno, latin, pop, jazz, and funk) on music-induced movement features (i.e., the five principal components).

Study VI utilizes two other multivariate approaches: mediation and moderation (Baron & Kenny, 1986; Hayes, 2009). Mediation was used to investigate if musical emotions could account for relationships between musical features and movement characteristics, whereas moderation was used to investigate whether personality traits of the dancers could affect relationships between musical features and movement characteristics.

6 STUDIES

This chapter describes the experimental studies included in this thesis (publications I to VI).

6.1 Aims of each study

In order to investigate different factors influencing music-induced movement, six studies were conducted and included in this thesis.

In Study I, relationships between musical features, in particular rhythm- and timbre-related features, and movement characteristics were studied. Following the notion of embodied music cognition (Leman, 2007), it was assumed that specific body movements reflect various aspects of the music, such as indicating beat and rhythmic structures. Main hypotheses included that clear beats would be embodied with an increased amount/speed of movement, whereas musical characteristics and structures were assumed to be reflected with hand and arm movements, as they offer the most freedom when moving. Furthermore, music-induced movement was hypothesized to be related to the tempo of the music, as mid-ranged tempi around 120 bpm were suggested to be most stimulating to move to (Fraisse, 1982; Moelants, 2002).

Study II served as a continuation of Study I, focusing on sub-band spectral flux, one of the musical features significantly contributing to music-induced movement. Results from Study I suggested that spectral flux of low and high frequency bands were important for inducing movement. As rhythmic components, in particular from kick drum and bass guitar, appeared to be prevalent in these frequency bands, it was therefore speculated that sub-band spectral flux is perceptually related to rhythm and rhythmic strength. However, perceptual aspects of sub-band spectral flux have not been investigated in great detail, so the aim of this study was to explore sub-band flux in relation to perceived rhythm, melody, fluctuation, and the propensity to move to shed further light on relationships between perception of spectral flux and music-induced movement.

Study III explored the role of perceived emotions in music in the context of music-induced movement. Emotions are an essential part of musical expression and experience (Gabrielsson & Lindström, 2010) and are furthermore an important component in the theory of embodied music cognition (Leman, 2007). Therefore, it was assumed that the use of body parts as well movement characteristics and patterns change when different emotions are expressed in the music – for instance, happy and active music was presumed to be embodied with active, fluid, and complex movement, as happy and active music might encourage listeners to move. Angry and happy music was supposed to show similar patterns with smoother movements for happiness. Inactive music, such as sad and tender songs, was assumed to result in reduced movements, since such music was imagined to be less movement-stimulating than active music (see, e.g., de Meijer, 1989; Boone & Cunningham, 1998; Burger & Bresin, 2010; Camurri et al., 2003; Dahl & Friberg, 2007; Wallbott, 1998).

Study IV is a follow-up study of Study III. Whereas results of study III revealed that perceived emotions in music and movement characteristics are strongly associated with each other, this study aimed at investigating whether music-induced movement can convey emotions to observers and how auditory and visual information would interact in this perceptual process. Three conditions of stimuli were included: audio-only, video-only, and audio-video with both emotionally congruent and emotionally incongruent combinations (i.e., combining the music representing happiness with the movements performed to the music representing sadness). The audio-video condition was assumed to receive highest recognition rates, while the auditory domain would dominate the perception of the audio-video stimuli (both suggested, for example, in Vines et al., 2006, and Petrini et al., 2010).

Study V addressed participants' personality traits and the genre of the music as factors influencing music-induced movement. Since personality traits were found to be associated to general movement characteristics (Ball & Breese, 2000) as well as to music preference (Rentfrow & Gosling, 2003) and emotions in music (Vuoskoski & Eerola, 2011a, 2011b), it was supposed that there is a connection as well with music-induced movement. A second focus of this study was put on the effect of genre on music-induced movement, assuming that the style of music considerably affects how people move, yielding, for instance, stereotypical or style-dependent (culturally-driven) movements.

The aim of Study VI was to combine musical features, perceived emotions, and personality that were separately analyzed and described in Studies I, III, and V. As it could be assumed that relationships exist between these factors and the ways they affect music-induced movement, Paper VI utilized two multivariate approaches to investigate them: mediation and moderation. Mediation was used for relationships between music, movement, and emotions; more specifically, the analysis tested if the emotions perceived in the music accounted for the relationship between musical and movement features. Moderation, on the other hand, was used to test if the level of personality trait changed the relationships between music and movement.

6.2 Methods

A motion capture data collection formed the core of this thesis. Study I and III-VI are directly related to this data collection, whereas Study II is based on the results of Study I, but features a separate collection of perceptual data. Furthermore, Study III included an additional data collection – a listening experiment – besides the motion capture data. Study IV used a subset of the motion capture recordings as audio and video stimuli in a separate perceptual experiment.

6.2.1 Participants

A total of 64 participants took part in the motion capture data collection (as used in Studies I and III-VI). Four participants were excluded from further analysis due to incomplete data. Thus, 60 participants remained for subsequent analysis. Participants were students from the different faculties of the University of Jyväskylä, and all but two were of Finnish nationality. They were recruited based on a database of 952 individuals that contained their scores of the Big Five Inventory (John et al., 2008). The aim was to recruit high- and low-scoring individuals for each of the five dimensions.

In Study II, 38 Finnish and international students from the University of Jyväskylä took part in the data collection. Thirty-four Finnish musicology students from the University of Jyväskylä participated in the additional data collection of Study III, whereas 80 university students of Finnish nationality took part in the perceptual experiment described in Study IV.

Participants of all data collections were compensated with a movie ticket.

6.2.2 Stimuli

In the motion capture data collection, participants were presented with 30 musical stimuli of different genres of popular music. All stimuli were 30 seconds long, non-vocal, and in 4/4 time, but differed in their rhythmic complexity, pulse clarity, and tempo, as well as in their perceived emotional content. The stimuli were chosen by the experimenters from a pool of 53 songs based on a perceptual evaluation with the aim of covering different degrees of rhythmic complexity and danceability. The list of stimuli used in the studies (including their tempi) can be found in Appendix 1.

The same set of stimuli was used in the listening experiment of Study III. Study II used the same stimuli in four different versions, the original and three frequency-restricted versions, which were created with MIRToolbox by dividing the stimulus into 10 octave-wide frequency bands and then choosing the 2nd (50-100 Hz), 6th (800-1600 Hz), and 9th sub-band (6400-12800 Hz) (same approach as used for the spectral flux computation described in Section 5.2). Study IV used a subset of four music stimuli representing the four discrete emotions happiness, anger, sadness, and tenderness and the respective movement data to these stimuli of four participants, resulting in 16 (silent) video stimuli. Additionally, the audio and video material was combined in all possible manners yielding 16 emotionally congruent and 48 emotionally incongruent audio-video stimuli.

6.2.3 Apparatus

For the motion capture data collection, participants' movements were recorded with an eight-camera optical motion capture system (Qualisys ProReflex), tracking at the frame rate of 120 Hz the three-dimensional positions of 28 reflective markers attached to each participant (see Section 5.1.1). The stimuli were played back via a pair of active studio monitors (Genelec 8030A) controlled by a Max/ MSP 5 patch.

For presenting the stimuli as well as for gathering the perceptual ratings for Studies II and IV, special custom-made patches were created using Max/ MSP 5. Study III used a custom-made Max/MSP 5 patch for presenting the stimuli, while the ratings were given on paper. In Studies II and III, the stimuli were played back through active studio monitors (Genelec 8030A), whereas in Study IV, studio quality headphones were used (AKG K141 Studio).

6.2.4 Procedures

In the motion capture data collection (as used in Studies I, III, V, and VI), participants were recorded individually. They were asked to move to the presented stimuli in a way it felt natural to them. The stimuli were presented in random order.

The experiment in Study II was divided into four sections: the first three sections contained the three frequency-restricted versions of the original stimuli, separated for each frequency band. The sections were presented in random order to the participants. The fourth (and last) section always comprised the original versions of the stimuli. The stimuli were randomized within each section. For each stimulus, four questions related to prominence of the rhythm, prominence of the melody, fluctuation, and desire to move were asked to be rated on seven-step scales. Participants were accomplishing the experiment individually.

For the perceptual task in Study III, two equally-sized groups were presented with the stimuli in different random orders. Each stimulus was rated according to its perceived emotional content on six seven-point scales related to arousal, valence, happiness, anger, sadness, and tenderness.

The experiment in Study IV was divided into three sections of four audio clips, 16 silent video clips, and 64 audio-video clips (16 emotionally congruent and 48 emotionally incongruent stimuli). The three sections were presented in random order, with the order of stimuli within each section being randomized as well. Tested individually, participants were instructed to rate the emotions expressed in the clips on seven-step scales for Happiness, Anger, Sadness, and Tenderness.

6.2.5 Movement features

For the analyses in Studies I, III, V, and VI, various movement features were extracted from the motion capture data. The mocap data preprocessing, the feature extraction process, and the features themselves have been explained in Section 5.1.3 of this thesis. Different sets of movement features were used in the studies, partly topic-related and partly to introduce new relevant and useful movement descriptors.

In Study I, eight movement features were used: Speed of Center of Mass, Head, Hands, and Feet, Distance of Hands, Amount of Movement, Hip Wiggle, and Shoulder Wiggle. The features were chosen, since they covered important body parts as well as relevant and distinct movement characteristics. Given the lack of research on specific movement features associated with musical features, the aim was to test and utilize a comprehensive range of different movement features.

Study III utilized ten movement features: Torso Tilt, Distance of Hands and Feet, Acceleration of Head, Hands, and Feet, Area of Movement, Fluidity, Rotation Range, and Movement Complexity. These features were chosen partly because they cover distinct body parts and movement characteristics and partly because of previous research suggesting particular features (e.g., de Meijer, 1989; Burger & Bresin, 2010; Camurri et al., 2003; Dahl & Friberg, 2007; Wallbott, 1998).

The movement features used in Study V were the results of a Principal Component Analysis (PCA) performed on 55 postural (distances between both hands, both elbow, and both feet; torso tilt), kinematic (velocity, acceleration, and jerk of center of mass, head, hands, and feet (both in three dimensions and separately for each dimension); area of movement), and kinetic movement (kinetic energy) features. The purpose of this PCA was to demonstrate a method for reducing the dimensionality of a large set of potentially correlated movement features to a smaller set of uncorrelated features. Five principal components were retained, explaining 90.3% of the variance: PC 1 was labeled Amount of Local Movement, PC 2 was tagged Amount of Global Movement, PC 3 was identified as Hand Flux, PC 4 was named Head Speed, and PC 5 was marked Hand Distance. Illustrations of the five PCs are shown in Figure 9.

Study VI used six movement features: Speed of Head and Hands, Acceleration of Head and Hands, Fluidity, and Rotation Range. These movement features formed a subset of the ones used in Studies I and III and were selected to meet the prerequisite needed for the analysis performed in the study that the musical features, the emotion ratings, and the movement features significantly correlated (pair-wise) with each other.

6.2.6 Musical features

For the analyses in Studies I, II, and VI, several musical features were extracted from the stimuli. The musical feature extraction process and the features have been described in Section 5.2.

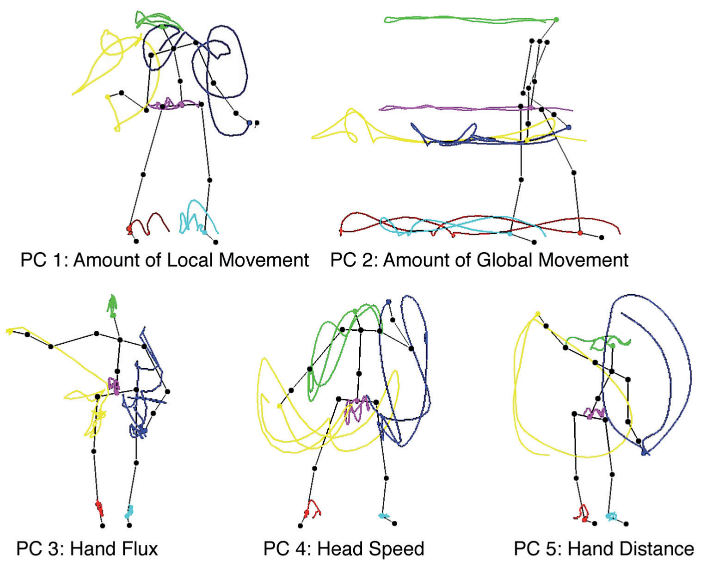


FIGURE 9 Illustration of the PCs.

Study I used five musical stimuli: Pulse Clarity, Spectral Flux of Sub-bands 2 and 9, Percussiveness, and Tempo. The first four were computationally extracted using MIRToolbox, whereas the Tempo was determined in a separate tapping experiment. The features were selected, since they covered important and diverse aspects of beat, rhythm, and timbre in music.

For Study II, the amount of Spectral Flux for the four versions of each stimulus (Sub-bands 2, 6, and 9, and whole) was used.

Studies VI utilized six musical features in the analysis: Spectral Flux of Sub-bands 2 and 9, Attack Time, Number of Onsets, Low Energy, and Spectral Centroid. These musical features were selected, since the prerequisite of the analysis performed in this study was that the musical features, the emotion ratings, and the movement features significantly correlated with each other.

6.3 Results and discussion

Next, the most important results of each study will be presented and discussed.

6.3.1 Study I – Rhythm- and timbre-related musical features and music-induced movement

Study I investigated relationships between musical characteristics and movement features. First, inter-participant consistency of the movement features was ensured by calculating intraclass correlations (r ranging from .62 to .95, $p < .001$)

to justify subsequent averaging of the features. The averaged features were then correlated with the (averaged) musical features. Results revealed significant correlations between Pulse Clarity and Center of Mass and Foot Speed, Amount of Movement, and Hip and Shoulder Wiggle (r ranging from .55 to .67, $p < .05$). Thus, with clearly perceivable beats, participants used a large amount of different movements and movement types, suggesting Pulse Clarity to influence movement on a global, whole-body level.

Spectral Flux of Sub-band 2 showed a strong correlation with Head Speed ($r = .73$, $p < .001$). It was assumed that Spectral Flux of Sub-band 2 was associated with kick drum and bass guitar, creating strong low-frequency rhythmic elements related to the beat period. The head might be most prone to move to the beat level due to its biomechanical properties (cf. nodding to the beat or head banging).

Spectral Flux of Sub-band 9 correlated significantly with Head and Hand Speed, Hand Distance, and Amount of Movement (r ranging from .57 to .65, $p < .05$). Spectral Flux of Sub-band 9 was assumed to be associated with hi-hat and cymbal sounds creating rhythmic figures and "liveliness" of the rhythm. The rationale for the upper body and hands in particular being related to high-frequency rhythmic structures could be that they are most free to move, so they reflect fast rhythmic structures best. The proposed relationship between rhythm and both sub-band fluxes provided the basis for Study II.

Percussiveness was associated with Center of Mass, Head, and Hand Speed, Hand Distance, Amount of Movement, and Shoulder Wiggle (r ranging from .55 to .67, $p < .05$). This result suggests that percussive elements were related to upper body movement, in particular to hand movement. High amounts of percussiveness seemed to be embodied with high speeds as well, reflecting possible ways such sounds are produced.

Furthermore, to test the non-linear relationships that were assumed between Tempo and the movement features, a series of ANOVAs was conducted. However, all ANOVAs showed non-significant differences between stimuli of slow, mid, and fast tempi for the movement features, thus stimuli of mid tempo (~120 bpm) failed to stimulate movement in a way different from slower or faster music. However, the variety of musical styles covered by the stimuli might have undermined the influence of tempo. Hence, this issue remains an open question for further investigation.

6.3.2 Study II – Spectral flux, perceived rhythm, and the propensity to move

Study II served as a continuation of Study I, focusing on the perceptual qualities of sub-band spectral flux. After verifying high inter-participant rating consistency by calculating intraclass correlations for each question and stimulus type separately (r ranging from .87 to .96, $p < .001$), the ratings were averaged across participants and subjected to correlational analyses between the ratings scores of the four questions and the computationally extracted spectral flux data of the respective frequency bands. Significant correlations were found between Rhythm Prominence and Flux of Sub-band no. 2 ($r = .79$, $p < .001$), Flux of Sub-band no. 9

($r = .47$, $p < .01$), and Flux of the Original Stimulus ($r = .54$, $p < .01$). The same Flux data correlated significantly with Movement Desire (Sub-band 2: $r = .65$, $p < .001$; Sub-band 9: $r = .52$, $p < .01$; Original: $r = .58$, $p < .001$). Thus, it was suggested that spectral flux of both low and high frequency ranges can be utilized as a possible measure of perceived rhythmic strength in music, and that high amounts of low and high frequency spectral flux are (at least partly) connected with the desire to move when listening to music. This conclusion was additionally supported by the significant correlations between the ratings for Rhythm Prominence and Movement Desire (Sub-band 2: $r = .86$, $p < .001$; Sub-band 6: $r = .78$, $p < .001$; Sub-band 9: $r = .72$, $p < .001$; Original: $r = .86$, $p < .001$). The results found in this study provide support for the outcome of Study I that the amount of spectral flux of low and high frequency bands is a crucial aspect for stimulating music-induced movement, as it is related to perceived rhythmic strength of the music.

6.3.3 Study III – Perceived emotions and music-induced movement

Study III investigated perceived emotions in music related to music-induced movement. First, inter-participant consistency of movement features and emotion ratings was ensured (intraclass correlations resulted in coefficients between .38 and .98, all significant at .05 level at least). Subsequently, correlational analysis between the averaged movement features and emotions ratings was conducted, which disclosed that each emotion category correlated significantly with a different set of movement features, which were therefore assumed to be emotion-specific. Significant correlations ranged from $r = |.45|$, $p < .05$, to $r = |.84|$, $p < .001$, with the highest correlations between the Arousal scores and Fluidity ($r = -.84$, $p < .001$) and Arousal and Hand and Head Acceleration ($r = .76$ and $.75$, $p < .001$). The detailed correlation results are presented in Publication III.

Accordingly, music perceived as active was associated with high acceleration of head, hands, and feet, non-fluent movement, and an upright torso. Perceived valence in the music was found related to Fluidity and Rotation Range, thus participants moved smoother and used more rotation when being exposed to pleasant music. Happy music was embodied by body rotation and complex movements, i.e., features that support the set of movement characteristics found in previous studies (Boone & Cunningham, 1998; Burger & Bresin, 2010; Camurri et al., 2003; Dael et al., 2012), but also extend the set. Music perceived as angry was expressed with irregular and jagged movements and no body rotation. These findings confirm previous literature (Boone & Cunningham, 1998; Camurri et al., 2003; Dahl & Friberg, 2007). Sad music was reflected with movements of low complexity, which as well corroborates previous research (Camurri et al., 2003; Dahl & Friberg, 2007; Wallbott, 1998). Tender music was embodied with a forward tilted torso and fluid movements of low acceleration, expressing the relaxed character of tender music. It was furthermore remarkable that Tenderness and Arousal exhibited a reverse correlation pattern suggesting that they share the same movement features, but with opposite characteristics.

Besides these discrete emotion ratings, the relationships between movement features and (discrete) emotion ratings relative to a two-dimensional space based on the valence and arousal ratings were investigated. Inspired by Russell's (1980) circumplex model of affect, an angular variable, *Circular Affect*, was derived from the valence and arousal ratings per stimulus (established as a projection of the rating combination on the unit vectors at each polar angle). Next, the newly created Circular Affect variable was correlated with the movement features and the four discrete emotion ratings and, as last step, the maximal correlation values and corresponding angles were obtained from the correlation vector. The results of this approach, the *Circular Affect Space*, are displayed in Figure 10.

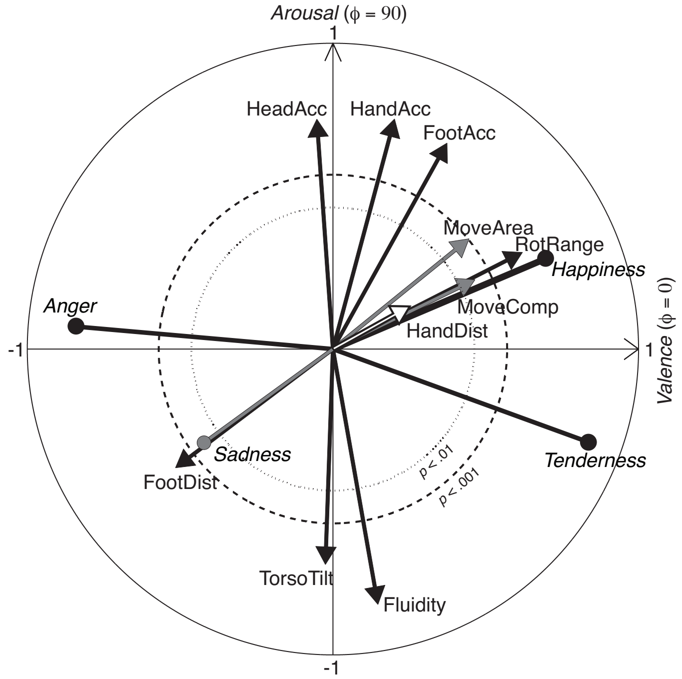


FIGURE 10 Maximal values of the correlations between *Circular Affect* and the ten movement features and the discrete emotion ratings arranged in a circular model (the *Circular Affect Space*). The length of the arrows reflects the maximal value of the correlation, and the polar angle indicates the direction of the maximal value.

Apart from Hand Distance, all movement features and the four discrete emotion ratings correlated significantly with Circular Affect (ranging from $r = .52$, $p < .01$, to $r = .93$, $p < .001$) with the highest maximal correlations between Circular Affect and Fluidity ($r = .85$, $p < .001$) in case of the movement features, and between Circular Affect and Tenderness ($r = .93$, $p < .001$) in case of the emotion ratings. The locations of the maximal correlations were spread over the whole Circular Affect Space with an accumulation in the pleasant active (upper right) quadrant, suggesting that music perceived as pleasant and active encouraged participants to move, especially with more accelerated as well as complex movement, rotation, and by using a large area.

The locations of the discrete emotions in relation to Circular Affect appeared consistent with previous results, both of this study and of Russell's (1980) circumplex models. Happiness was found most clearly associated with specific movement features (rotation, complex movement, and area covered when moving), whereas Anger was located between the inverse of the maximal correlations of Fluidity and Rotation Range. Sadness was found close to the inverse location of Movement Complexity and Foot Distance. For Tenderness, the corresponding movement features did not show a clear resemblance in the Circular Affect Space. However, this might be due to the choice of stimuli being rather active in order to stimulate movement, which might, in turn, explain the rather active location of the maximal correlation of Tenderness in the Circular Affect Space.

6.3.4 Study IV – Emotion recognition in music-induced movement

Study IV explored the capabilities of music-induced movement in conveying emotions to observers. Following an outlier detection and intraclass correlations to determine rating consistency (r ranging from .95 to .97, $p < .001$), it was first analyzed which emotions were perceived in the different conditions and if the target emotions were perceived. In the audio condition, the four emotions Happiness, Anger, Sadness, and Tenderness were successfully conveyed (for detailed results including repeated measures ANOVAs and error bar plots see Publication IV). In the video condition, only Happiness was successfully conveyed, whereas Anger and Tenderness were both confused with Happiness, and Sadness was mixed with Tenderness. In the congruent audio-video condition, Happiness and Anger could be successfully recognized, while Sadness was confused with Tenderness and Tenderness with Happiness. Such confusions have been reported in the literature (e.g., Atkinson et al., 2004; Eerola & Vuoskoski, 2011; M. M. Gross, Crane, & Fredrickson, 2010; Pollick et al., 2001), though in case of this study, they might be also due to the nature of the stimuli (containing a recognizable beat that induced movement). Conveying negative emotions (like anger and sadness) through dance movement might also be a challenging task that is only possible if the dancers are explicitly asked to move accordingly, as dancing is commonly rather associated with happiness and positive attitudes.

The audio condition was most effective in conveying the four target emotions (highest average ratings, conditions differed significantly), followed by the audio-video condition. The results – error bar plots and repeated measures ANOVAs – can be found in Publication IV. This result failed to confirm both the hypothesis and previous studies that suggested the audio-video condition to receive highest recognition (e.g., Collignon et al., 2008; Davidson, 1993; Petrini et al., 2010; Vines et al., 2006). However, it might be related to the use of non-professional dancers and the instructions not being focussed on emotions.

A series of linear regression analyses for each of the target emotions (audio ratings predicting the audio-video ratings, video ratings predicting the audio-video ratings, and both audio and video ratings predicting the audio-video ratings) revealed that for all target emotions the ratings for the audio condition could explain a larger amount of the variance of the audio-video ratings than the video ratings (for statistics, see Publication IV). It was therefore concluded that the participants were more attentive to the auditory than to the visual part of the stimuli, which is partly in line with previous research (Petrini et al., 2010; Vines et al., 2006). Furthermore, the audio and video ratings were found to be additive in the regression model: Between 91 and 96% of the variance of the audio-video ratings could be explained by the audio and the video ratings, thus the audio-video ratings could be almost perfectly predicted from the single-modality ratings.

To investigate the distribution and alignment of the ratings in the audio-video condition, the dimensionality of the ratings was reduced by using principal components analysis (PCA). Two components were retained after applying Kaiser's criterion (accounting for 77.6% of the total variance – PC 1: 41.25% and PC 2: 36.55%). Subsequently, the components were varimax rotated and then averaged across dancers and participants yielding 16 values per PC (one per audio-video combination). These values are displayed as coordinates in Figure 11.

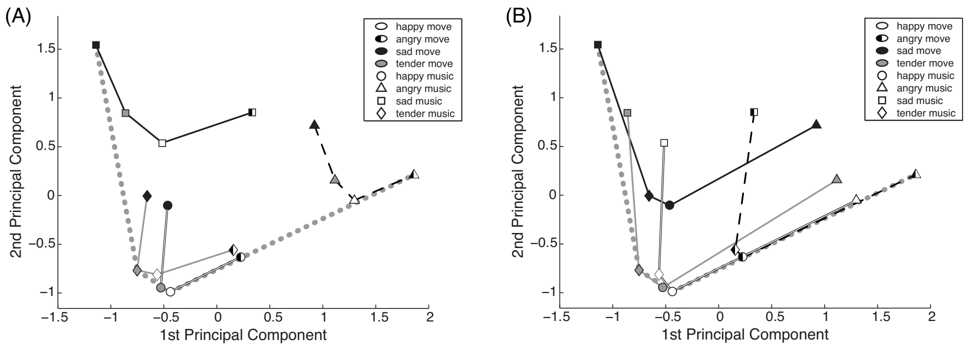


FIGURE 11 Principal components solution of audio-video ratings. (A) Same audio stimuli connected (same symbol). (B) Same video stimuli connected (same color/shade). The dotted lines connect the congruent stimuli.

Figure 11A shows the solution with the same audio stimuli being connected, whereas Figure 11B shows the solution with the same video stimuli being connected. The stimuli having the same audio are grouped closer to each other than the stimuli having the same video, as there is no overlap of the clusters in the same-audio solution (if the happiness and tenderness clusters are taken together, since they were rated similarly as described previously) as opposed to the same-video solution. This supports the finding that the auditory domain was more influential for the audio-video ratings. Furthermore, all the clusters show a self-similar structure with each other as well as with the alignment of the congruent stimuli (dotted lines in Fig. 11), indicating that the participants rated very consistently and coherently despite the contradictory information in the incongruent stimuli. Both congruent stimuli and music might have guided the perception of the incongruent stimuli and caused a homogeneous rating scheme that participants applied during this part of the experiment.

6.3.5 Study V – Personality and music-induced movement

Study V investigated effects of personality and musical genre on music-induced movement. Previous to the analysis, the participants were divided into three groups of low, mid, and high scorers per personality trait, retaining the low and high groups. A 2 (personality: low-scorers vs. high-scorers) x 6 (genre) Multivariate Analysis of Variance (MANOVA) was conducted for each of the five personality traits and movement component (for statistical results see Publication V). The significant main effects for Openness, Conscientiousness, Extraversion, and Agreeableness on several movement characteristics as represented by the five principal components showed positive relationships; in other words, high-scorers moved, for instance, more and faster than low-scorers. For Neuroticism, this (positive) result was only found for Local Movement, whereas this trait was negatively related to the other PCs (i.e., low-scorers moved more than high-scorers). Extraversion and Neuroticism were the only traits that showed significant effects on all five movement components, suggesting that these two traits might be most clearly associated with typical movement characteristics. This could be because these two traits are strongly linked with emotional expressivity (J. J. Gross & John, 1995), which is linked with music (e.g., Sloboda, 2010) and movement (e.g., Atkinson et al., 2004; Pollick et al., 2001). Extraverts tend to be energetic, positive, and stimulation-seeking, which could account for the positive relationships with the five movement components. Neurotics, on the other hand, tend to be anxious and stressed (especially in a lab situation), so they might have reacted with reduced levels of hand and head movements and moved in a local rather than in a space-intensive, global area.

Regarding genre, Rock was associated with high levels of Head Speed, in particular for extravert and agreeable participants, and with low levels of Global Movement, especially for extraverts and neurotics. This result hints at stereotypical movements related to Rock, such as “head banging” and movements covering a small area. Techno was related to high levels of Local Movement, especially for open, extravert, and agreeable participants, thus eliciting high levels of stationary limb movements. Latin music was linked to high levels of Global Movement, in particular for participants high on Openness and Agreeableness, and to low levels of Head Speed, especially for open participants. Hence, Latin music caused participants to use the existing space while keeping their head still – both characteristics of Latin American dances, so participants probably associated the music with typical dancing figures and established conventions.

6.3.6 Study VI – Music, emotion, personality, and music-induced movement

This study describes two multivariate analysis approaches, mediation and moderation, that allow investigation of relationships between more than two variables. Mediation analysis was used here to study whether the perceived emotional content accounted for relationships between musical features and movement characteristics. Mediation was found to occur in case of Arousal, Valence, Happiness, and Tenderness with several movement characteristics and musical features, indicating that the emotional content could account for relationships between music and participants' movement. This result suggests that participants' movements were rather guided by the emotional connotation of the music than by the musical features themselves.

Moderation analysis was utilized to investigate whether the level of personality trait affected relationships between music and movement. However, the resulting moderation models failed to show significant interactions between personality traits and musical features, suggesting that the personality level did not change the relationship between music and the movements of the participants. This could indicate that musical characteristics and personality traits are independent factors that are connected to movement in different, non-interactive ways.

7 GENERAL DISCUSSION AND CONCLUSION

This thesis research investigated various factors influencing music-induced movement, in particular musical features, perceived emotions in music, genre, and personality. With all four factors, distinct relationships could be established: in general, we found that music with clear beats and strong rhythmic components, such as high spectral flux in low and high frequency components, encouraged participants to move along with it. In addition to the musical features, the emotional characteristics of the music affected the movement properties, suggesting that the body reflected emotional qualities of the music. A novel approach of mapping the movement features into a 2-dimensional valence-arousal space was found to be meaningful and consistent with previous research (e.g., Russell, 1980). Furthermore, it was found that certain genres elicited stereotypic and style-dependent movement, such as head banging to rock or dancing figures typical for Latin dance. Besides these rather music-intrinsic factors, individual differences, such as the personality traits of the dancers, were also included in this research and found to shape music-induced movement in distinctive ways. Two personality traits, extraversion and neuroticism, were found to be particularly related to movement characteristics. This might be because these two traits are strongly connected to emotional expressivity (J. J. Gross & John, 1995), thus emotions both expressed and induced by music might be especially linked to body movement for such participants. Statistical methods that integrate more than two variables showed that some of the factors related to music-induced movement might be independent from each other, such as personality and musical features, whereas others, such as emotions and music, are more intertwined. Moreover, perceptual experiments indicated that emotions were perceived consistently and congruently in dance movements presented as animated stick figure clips, even when their content (the audio and video material) was emotion-wise incongruent. The audio content was generally influencing the ratings more strongly than the video content, suggesting that the participants' ratings were guided by the audio content.

The studies and their results show that relationships between music and music-induced movement characteristics are multilayered and manifold, being related to a variety of factors that not only include musical characteristics. The results also suggest extending, for instance, the notion on groove, as such research has mainly focussed on musical characteristics so far (e.g., Madison et al., 2011). Investigating relationships between music-induced movement and groove would be beneficial as well as relating groove to aspects other than music-intrinsic ones.

The multidisciplinary approach of integrating motion capture technology, computational feature extraction of movement and musical features, perceptual experiments, and different kinds of statistical analysis enabled investigating music-induced movement from various angles and view points providing the research area with novel contributions as well as more comprehensive understanding. Additionally, it demonstrates that research in this field greatly benefits from current technological advances and analysis methods, as this research would not have been possible without, for example, a high-quality motion capture system. Technological and computational approaches chosen for conducting the studies made handling large amounts of rather unconstrained data feasible and meaningful.

When comparing the plain correlation coefficients obtained in Studies I (musical features) and III (emotions), it can be noticed that they were higher in the emotion study than in the musical feature study. This finding might shed further light on the relation and weighting between musical characteristics and emotions; as suggested by the results from Study VI, participants' movements might have been more guided by the emotional content than by the music and musical characteristics themselves. "Empathy" as the more abstract description in terms of the embodied music cognition framework seems to be of greater importance than the more music-related concept of "Embodied Attuning".

The studies contribute to the field of embodied cognition, and more specifically to the field of embodied music cognition. Movements and bodily expressions were found to be goal-directed, as they reflected musical structure and characteristics as well as emotional content. Participants' movement responses were suggested to be related to the three concepts of corporeal articulations introduced by Leman (2007), "Synchronization", "Embodied Attuning", and "Empathy". However, they were shown to be difficult to disentangle, as they occurred at the same time, as, for instance, observed in Study III. This would suggest that individuals reacted to various cues communicated by the music simultaneously and on a subconscious level, as they were asked to move freely and naturally to the music (and not to display emotions or aspects of the musical structure). The studies further showed that embodied music cognition covers essential everyday phenomena of human behavior, since individuals of the general populations were used as participants.

Interesting to note is that individual differences, such as personality, do not appear in the theory of corporeal articulations. However, it seems likely to assume that movements can be seen as gestural expressions related to individual factors of a performer/listener/dancer, just as they are assumed to embody, for example, emotions. Thus, the integration of individual factors into approaches of body-related musical behavior might be a fruitful strategy to formulate a holistic theory of music-related and -induced corporeal articulations.

Regarding musical characteristics exhibiting relationships with movement responses, the spectral flux of different frequency ranges was demonstrated to be an important feature for inducing movement. However, only the spectral flux of low and high frequency ranges were found significant, whereas the flux of mid frequencies was found not significant. The low and high frequency ranges are assumed to carry information regarding the rhythmic structure, while the mid frequencies would rather convey melody-related aspects. Thus, it could be presumed that music-induced movement (at least the kinds of movement collected in this study) rather reflected the rhythmic aspects of music, or that the extracted movement features were more related to rhythmic than to melodic aspects, as the features were somewhat gross and broad. The process of averaging the time-series of the feature values might have additionally contributed to that result. It could therefore be that melody is embodied through movement that evolves over a longer period of time, whereas rhythm would be embodied with repetitive movement (covered well with the averaged features).

Moreover, static timbre features, such as brightness, failed to show notable relationships with movement characteristics. Although such features were found relevant for the perception of polyphonic timbre (Alluri & Toiviainen, 2010), the results from the investigations presented in this thesis imply that dynamic timbre characteristics, such as the spectral flux, are more pertinent for music-induced movement than static timbre features.

Study I and Study VI partly included different musical features, since the analyses chosen in both studies had different requirements. Forty-three musical features were extracted in the beginning and entered into subsequent analyses: Study I attempted to control for false discoveries by decreasing the significance level using Bonferroni correction. This approach caused some of the correlation coefficients between the musical and movement features to be too low to be significant, which were thus excluded. Study VI, however, required that the pairwise correlations between all variables (i.e., musical features with emotion ratings, musical features with movement features, and emotion ratings with movement features) that were subjected to the analyses were significant, which only a small number of the possible combinations did.

The secondary aim of this thesis was related to the development of movement features applicable to music-related movement and, at best, to movement analysis in general. The intention was to cover in particular higher-level features that are more profound than plain velocity-, acceleration-, and distance-related features in describing movement in terms of shapes and patterns. Such features would thus comprise more conceptual and abstract aspects of movements than the lower-level features. Features such as body rotation, fluidity, and movement complexity have been developed and applied in this thesis research to fulfill this aim. It was shown that these features covered relevant and distinctive movement characteristics arising in music-induced movement, contributing to the pool of existing movement features and to the field in general. Some of the features developed for the thesis research, such as the bounding rectangle, have already been integrated in the MoCap Toolbox, while more, such as body wiggle and rotation, will be implemented in the near future.

7.1 Limitations and future work

This research explored four factors influencing music-induced movement (musical features, musical emotions, genre, and personality). However, movements induced by music could also be affected by other factors, such as preference, familiarity, motivation, empathy, or self-esteem. The influence of preference (Lamont & Greasley, 2009; Rentfrow & Gosling, 2003) might reveal interesting insights into relationships between musical characteristics and induced movement, as it could be assumed on one hand that music is more likely to induce movement, if the individual likes the music. On the other hand however, certain music features could induce movement regardless of the individual liking the music. Identifying such musical characteristics would advance research in both music-induced movement and in groove. It could be further presumed that movement patterns would change the more familiar individuals are with certain musical pieces or a certain genre. Besides finding characteristic movement features associated with familiarity, such research could also disclose changes in movement characteristics over time (i.e., how movement characteristics change when the music gets more familiar).

Personality was shown to be important in shaping music-induced movement. In order to obtain a wider understanding how personality characteristics affect movement properties, alternatives to the Big Five personality traits could be employed, such as the Myers-Briggs Type Indicator (Briggs Myers & Myers, 1980), or also the assessment of temperament. The link between personality and emotions related to the neurotic and extravert participants (as suggested by the results of Study V) would deserve further attention, as it could be supposed that certain music-related emotions are embodied differently depending on the personality of the dancers. Additionally, other individual difference measures could be worth investigating as well, for instance empathy. The investigation of mood, as already pursued within the MMP project, has revealed insights (Saarikallio, Luck, Burger, Thompson, & Toiviainen, 2013). However, the investigation could be extended by measuring how music-induced movement can affect and change mood.

The studies presented in this thesis have not considered any background variables related to the participants, although they might reveal valuable insights into habits and behaviors related to music-induced movement. To mention a few examples, the preference for dancing could be investigated as well as the effect of dance training on movement characteristics. Besides general musical preferences and dance habits, participants were also asked about familiarly, danceability, and preference of the stimuli used in the mocap study. These data will be analyzed in the near future. In subsequent studies, the Goldsmiths Musical Sophistication Index (Gold-MSI) could be employed to measure musical attitudes, behaviors, and skills (Müllensiefen, Gingras, Stewart, & Musil, 2012). Additionally, commonalities and differences between professional and non-professional dancers and musicians would serve as an interesting path to gather more knowledge about music-related movement. Such an approach also gives raise to the question of skill, the measurement of different levels of skill, and if skill is related to certain musical styles and characteristics (i.e., if there are styles to which people can dance regardless of skill).

The motion capture study conducted during this thesis research was based on movement performed individually. However, such dance movements are more commonly performed in groups (e.g., in a night club), thus a more ecological and interactive setting is to be achieved in future studies. Interactions are a crucial part of embodied (music) cognition, in particular the mutual influences between action and perception could be studied. Aspects such as turn-taking and leader-follower relationships (Goebl & Palmer, 2009) as well as joint action (Keller, 2008) and shared representations (Tomasello & Carpenter, 2007) could be studied. Such behaviors have been found in various kinds of collaborative music activities, for instance in music performance: coordinating the own playing actions with those of the co-performers to reach the common goal of a successful performance; or in ballroom dancing: coordinating the actions of both partners to achieve correct dance figures and expressive movements. For dancing in a group, this might mean: mimicking others, coordinating gestures, sharing ideas and intentions, having fun. Therefore, studying mutual interaction on the dance floor gives access to a wide range of unexplored research questions.

The choice of stimuli could be seen as a limitation. All music stimuli were of western style with a beat to induce movement. However, follow-up studies could use a wider range of genres and different kinds of music, such as world and classical music or music without a beat. Another approach would be to use very clear examples of specific genres to address the effects of genre. In particular, extreme genres, such as death metal, would provide insights into relations between music, movement, and preference, as such genres are likely to be more connected to preference (i.e., only people who like such music would enjoy moving to it). Moreover, stimuli that very clearly express certain emotions could be used to further tackle the effects of emotions. Such approaches would also provide the opportunity to constrain the stimulus material and to give more distinct participant instructions to test more specific hypothesis. This stream of thought could be also continued in relation to musical features. In particular the tempo of the musical stimuli could be investigated in a more controlled experimental approach with regards to the preferred beat period (as mentioned earlier).

The studies only used Finnish participants. However, to gain more understanding of commonalities and differences between different cultures, such experiments should be conducted with participants from various cultural backgrounds.

Mediation and moderation analysis proved to be an interesting and fruitful approach for investigating relationships between several variables. However, the approach presented in Study VI only used three variables. The aim for further approaches would be to include more variables (e.g., genre or synchronization-related features) to gain a more holistic view on the data and reveal more general results than the present analysis. Furthermore, methods of path modeling and structural equation modeling (SEM) might yield more comprehensive models including several observed variables and the possibility to build latent factors from observed variables (Schumacker & Lomax, 2010). Such an approach would allow investigating higher-level relationships between, for instance, "music" and "movement" with intervening factors "emotions" and "personality".

This thesis proposed several movement features that were shown to describe relevant movement characteristics. However, they were of static nature, as only one value (e.g., average, standard deviation, or entropy) was used to describe each feature. One can, however, assume the relationship between musical features and movement characteristics to be more complex (e.g., non-linear), as the movement develops in time according to the music. The next steps will therefore include more complex movement pattern analysis and time-series analysis to reveal information about the dynamic processes of music-induced movement. Such an analysis approach might also give further insight into the embodiment of melodies as well as provide the possibility to formulate specific hypotheses that test theories related to embodied (music) cognition. Additionally, movement features that, for instance, characterize how participants synchronize to music would describe goal-directed actions (e.g., interactions, shared goals). Such developments will be beneficial for the research area and will furthermore advance the field of (applied) music embodied cognition.

MoCap Toolbox will be developed further, especially with respect to preprocessing the data. More advanced gap-filling methods will be implemented, such as higher-level polynomial filling methods and methods that predict the gap based on the course of data preceding and following the gap. As mentioned above, the features used in this thesis will be implemented in the toolbox to make them available to other users.

Future studies could address links between music-induced movement and groove, as the research conducted for this thesis mainly dealt with actual movement, whereas groove rather considers the sensation of wanting to move. However, bringing both together by, for instance, asking participants to physically move to stimuli with varying amounts of groove would yield interesting insights into the embodiments of groove.

A very interesting approach would be to combine movement research and brain research to further investigate links between auditory and motor areas and other parts that could be important for music-induced movement, such as pleasure- and preference-related areas in the brain.

The outcome of this research could be used in applications related to music information retrieval. Gesture-based search-engines, with which users would get suggestions of songs based on their finger movement on the touch screen or based on their body movement recorded with the camera of the device could be one objective. Furthermore, a visualizer application displaying animated stick-figures dancing to music played on the device would be an entertainment implementation of this research. Such an approach would furthermore include modeling the stick-figure movement based on musical or emotional characteristics. Moreover, DJs could use the information gained in this research when selecting music to be played in the club.

7.2 Concluding remarks

In an interdisciplinary way, this thesis research presented various angles of approaching music-induced movement and could identify several factors (musical features, perceived emotions, genre, and personality) that affected such movements in different ways. Both a motion capture study and perceptual studies could reveal important insights into the ways individuals move quasi-spontaneously to music. High-level movement features were developed and demonstrated to be applicable to meaningfully describe movement characteristics.

The studies show that even though quasi-spontaneous whole-body movement to music yields a large amount of rather unconstrained data, relevant movement characteristics and commonalities could still be obtained and linked to musical behavior. Nevertheless, this research field is still in its beginning phase and requires more exploration. The four factors investigated in this thesis as well as other aspects, such as preference, would need further exploration in relation to music-induced movement. Such approaches would advance the field of embodied music cognition and would help creating a holistic, integrative theory including factors that shape music-induced movement.

In conclusion, several aspects investigated in this work suggest paths that would bring research in music-related, and in particular music-induced, movement considerably forward in the near future. These aspects include tasks such as appropriately constraining the stimulus material and formulating participant instructions that would enable to test specific hypotheses. Moreover, aspects related to more advanced analysis methods that consider synchronization and the dynamic nature of movement and music, as well as more realistic dance situations (e.g., interactive settings) and links to brain mechanism will provide further provide insights into the research field in particular and into embodied music cognition research in general.

YHTEENVETO (FINNISH SUMMARY)

Musiikin kuunteleminen saa meidät liikkumaan. Liike saattaa vaihdella lähes huomaamattomasta jalan polkemisesta tanssiliikkeisiin, joissa koko keho on mukana. Tällaiset liikkeet vaikuttavat olevan ajallisesti yhdenmukaisia musiikin rytmin ja muiden musiikille luonteenomaisten piirteiden kanssa. Kehon liikkumisessa musiikin kuuntelun aikana on myös paljon yksilöllisiä eroja. Tämän tutkimuksen tarkoituksena on tarkastella tekijöitä, jotka vaikuttavat musiikin aikaansaamiin kehon liikkeisiin. Näitä ovat esimerkiksi musiikin ominaispiirteet, subjektiivinen kokemus musiikin välittämistä tunteista, musiikin tyylilaji sekä musiikin tahtiin liikkuvan henkilön persoonallisuuspiirteet. Tässä liikkeenkaappaustekniikkaa hyödyntävässä tutkimuksessa tarkasteltiin 60 koehenkilön vapaata liikehdintää 30 eri musiikkikatkelman mukana. Saadusta havaintoaineistosta erotettiin MATLAB MoCap Toolbox -ohjelmistoa käyttäen useampia muuttujia, jotka liittyivät kehon asentoihin, kinematiikkaan ja liikkeisiin. Musiikin ominaispiirteistä erityisesti pulssin selkeys ja spektrin vaihtelu tietyillä taajuusalueilla korreloivat vahvasti erinäisten liikemuuttujien kanssa. Arviointikoe vahvisti yhteyden taajuusalueiden spektrin vaihtelun ja selkeästi havaittavan rytmin välillä. Taajuusalueiden spektrin vaihtelu oli yhteydessä myös liikkumisen todennäköisyyteen. Arvioijan musiikissa havaitsemat tuntemukset olivat läheisessä yhteydessä erinäisiin liikemuuttujiin. Arviointikokeessa, jossa animoitujen ‘tikku-ukkojen’ liikettä arvioitiin suhteessa liikkeen ilmentämiin tunteisiin, havaittiin, että musiikin mukana liikkuminen itsessään ilmaisi arvioijalle tuntemuksia, vaikka äänen ja kuvan välittämä informaatio oli ristiriitaista. Koehenkilöiden persoonallisuuspiirteillä oli vahva vaikutus siihen kuinka he liikkuivat musiikin mukana. Monimuuttuja-analyysi paljasti, että musiikissa havaitut tunteet välittivät huomattavissa määrin musiikin ominaispiirteiden ja liikemuuttujien välistä yhteyttä. Tanssijoiden persoonallispiirteillä ei ollut välittävää roolia havaituissa musiikin ominaispiirteiden ja liikemuuttujien välisissä yhteyksissä. Tämä tutkimus esittelee uusia ja innovatiivisia menetelmiä musiikkiin liittyvien liikkeiden tutkimiseen. Musiikin aikaansaaman liikkumisen ja siihen vaikuttavien tekijöiden välisiä yhteyksiä voidaan tunnistaa ja tarkastella hyödyntämällä nykyaikaista liikkeenkaappaustekniikkaa ja tilastollisia monimuuttuja-analyysimenetelmiä. Tämä näkökulma tarjoaa lähtökohtia ja sovelluksia niin kehollisen musiikillisen kognition kuin ei-kielellisen, jokapäiväisen inhimillisen käyttäytymisenkin tutkimukseen.

REFERENCES

Alluri, V., and Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception*, *27*(3), 223–242.

Alluri, V., and Toiviainen, P. (2012). Effect of enculturation on the semantic and acoustic correlates of polyphonic timbre. *Music Perception*, *29*(3), 297–310.

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, *149*(1), 91–130.

Arom, S. (1991). *African polyphony and polyrhythm: Musical structure and methodology*. Cambridge, UK: Cambridge University Press.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., and Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, *33*(6), 717–746.

Aucouturier, J.-J., and Pachet, F. (2003). Representing musical genre: A state of the art. *Journal of New Music Research*, *32*(1), 83–93.

Balkwill, L.-L., and Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, *17*(1), 43–64.

Ball, G., and Breese, J. (2000). Relating personality and behavior: Posture and gestures. In A. Paiva (Ed.), *Affective interactions: Towards a new generation of computer interfaces. Lecture Notes in Computer Science, 1814* (pp. 196–203). Berlin/Heidelberg, Germany: Springer.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., et al. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, *30*(3), 917–926.

Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bechinie, M., and Grammer, K. (2003). Charisma Cam: A prototype of an intelligent digital sensory organ for virtual humans. In T. Rist, R. S. Aylett, D. Ballin, and J. Rickel (Eds.), *4th International Workshop on Intelligent Virtual Agents, Kloster Irsee, Germany. Lecture Notes in Computer Science, 2792* (pp. 212–216). Berlin/Heidelberg, Germany: Springer.

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., et al. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, *45*, 62–71.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Boone, R. T., and Cunningham, J. G. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology*, *34*(5), 1007–1016.

Brebner, J. (1985). Personality theory and movement. In B. Kirkcaldy (Ed.), *Individual differences in movement* (pp. 27–41). Lancaster, UK: MTP Press.

Briggs Myers, I., and Myers, P. B. (1980). *Gifts differing: Understanding personality type*. Palo Alto, CA: Consulting Psychology Press.

Brown, S., Merker, B., and Wallin, N. L. (2000). An introduction to evolutionary musicology. In N. L. Wallin, B. Merker, and Brown (Eds.), *The origins of music* (pp. 3–24). Cambridge, MA: MIT Press.

Burger, B., and Bresin, R. (2010). Communication of musical expression by means of mobile robot gestures. *Journal on Multimodal User Interfaces*, *3*, 109–118.

Cai, R., Lu, L., Hanjalic, A., Zhang, H.-J., and Cai, L.-H. (2006). A flexible framework for key audio effects detection and auditory context inference. *IEEE Transactions on Audio, Speech and Language Processing*, *14*(3), 1026–1039.

Camurri, A., Lagerlöf, I., and Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, *59*(1-2), 213–225.

Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., and Volpe, G. (2004). Multimodal analysis of expressive gesture in music and dance performances. In A. Camurri and G. Volpe (Eds.), *Proceedings of the 5th International Gesture Workshop: Gesture-based Communication in Human-Computer Interaction, Genova, Italy. Lecture Notes in Computer Science, 2915* (pp. 20–39). Berlin/Heidelberg, Germany: Springer.

Caramiaux, B., Bevilacqua, F., and Schnell, N. (2010). Towards a gesture-sound cross-modal analysis. In S. Kopp and I. Wachsmuth (Eds.), *Proceedings of the 8th International Gesture Workshop: Gesture in Embodied Communication and Human-Computer Interaction, Bielefeld, Germany. Lecture Notes in Artificial Intelligence, 5934* (pp. 158–170). Berlin/Heidelberg, Germany: Springer.

Casciato, C., Jensenius, A. R., and Wanderley, M. M. (2005). Studying free dance movement to music. In *Proceedings of the International Conference in Practical, Psychological, Philosophical, and Educational Issues in Music Performance* (pp. 1–4). Porto, Portugal.

Castellano, G., Villalba, S. D., and Camurri, A. (2007). Recognising human emotions from body movement and gesture dynamics. In A. Paiva, R. Prada, and R. W. Picard (Eds.), *Proceedings of the 2nd International Conference in Affective Computing and Intelligent Interaction, Lisbon, Portugal. Lecture Notes in Computer Science, 4738* (pp. 71–82). Berlin/Heidelberg: Springer.

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *The Neurosciences and Music III – Disorders and Plasticity: Annals of the New York Academy of Science*, *1169*, 15–34.

Clark, A. (2011). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York, NY: Oxford University Press.

Clynes, M. (1980). The communication of emotion. In R. Plutchik and H. Kellerman (Eds.), *Emotion: Theory, research, and experience. Vol. 1: Theories of emotion* (pp. 270–300). New York, NY: Academic Press.

Clynes, M. (1992). Time-forms, nature's generators and communicators of emotion. In *Proceedings of the IEEE International Workshop on Robot and Human Communication* (pp. 18–31). Tokyo, Japan.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135.

Collyer, C. E., Broadbent, H. A., and Church, R. M. (1992). Categorical time production: Evidence for discrete timing in motor control. *Perception & Psychophysics*, *51*(2), 134–144.

Costa, Y., Oliveira, L., Koerich, A., Gouyon, F., and Martins, J. (2012). Music genre classification using LBP textural features. *Signal Processing*, *92*(11), 2723–2737.

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117–139.

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences*, *930*, 28–42.

Dael, N., Mortillaro, M., and Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, *12*(5), 1085–1101.

Dahl, S., and Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, *24*(5), 433–454.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London, UK: John Murray.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*, 103–113.

De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*(3), 289–311.

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, *13*(4), 247–268.

Dempster, W. T. (1955). *Space requirements of the seated operator: Geometrical, kinematic, and mechanical aspects of the body with special reference to the limbs.* WADC Technical Report 55-159, Wright-Patterson Air Force Base, Ohio.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., and Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*(6), 727–738.

Dunn, G. (2009). Music preferences based on audio features and its relation to personality. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, and P.-S. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Soci-*

ety for the Cognitive Sciences of Music (pp. 83–92). Jyväskylä: University of Jyväskylä.

Dunn, P. G., Ruyter, B. de, and Bouwhuis, D. G. (2012). Toward a better understanding of the relation between music preference, listening behavior, and personality. *Psychology of Music*, *40*(4), 411–428.

Eerola, T., Luck, G., and Toiviainen, P. (2006). An investigation of pre-schoolers' corporeal synchronization with music. In M. Baroni, A. R. Addessi, R. Caterina, and M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition* (pp. 472–476). Bologna, Italy: University of Bologna.

Eerola, T., and Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49.

Ekman, P. (Ed.). (1982). *Emotions in the human face*. Cambridge, UK: Cambridge University Press.

Ekman, P., and Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, *29*(3), 288–298.

Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.

Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*(6), 2608–2611.

Fink, B., Weege, B., Flügge, J., Röder, S., Neave, N., and McCarty, K. (2012). Men's personality and women's perception of their dance quality. *Personality and Individual Differences*, *52*(2), 232–235.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York, NY: Academic Press.

Gabrielsson, A., and Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, *24*, 68–91.

Gabrielsson, A., and Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin and J. Sloboda (Eds.), *Handbook of music and emotion. Theory, research, applications* (pp. 367–400). Oxford, UK: Oxford University Press.

Godøy, R. I., Haga, E., and Jensenius, A. R. (2006). Playing "air instruments": Mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty, and J.-F. Kamp (Eds.), *Gesture in human-computer interaction and simulation. Lecture Notes in Computer Science, 3881* (pp. 256–267). Berlin/Heidelberg, Germany: Springer.

Godøy, R. I., and Jensenius, A. R. (2009). Body movement in music information retrieval. In K. Hirata, G. Tzanetakis, and K. Yoshii (Eds.), *Proceedings of the*

10th International Conference on Music Information Retrieval (pp. 45–50). Kobe, Japan: International Society for Music Information Retrieval.

Goebl, W., and Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, *26*(5), 427–438.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26–34.

Grahn, J. A., and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Grahn, J. A., and Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*(23), 7540–7548.

Gross, J. J., and John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences*, *19*(4), 555–568.

Gross, M. M., Crane, E. A., and Fredrickson, B. L. (2010). Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior*, *34*(4), 223–248.

Hanson, A. J. (2006). *Visualizing quaternions*. San Francisco, CA: Elsevier/Morgan Kaufmann.

Hartmann, M., Saari, P., Toiviainen, P., and Lartillot, O. (2013). Comparing timbre-based features for musical genre classification. In R. Bresin (Ed.), *Proceedings of the International Conference on Sound and Music Computing* (pp. 707–714). Stockholm, Sweden: KTH Royal Institute of Technology.

Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, *17*(2), 282–293.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420.

Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*(5), 673–682.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*(3), 529–535.

Ilie, G., and Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*(4), 319–330.

Janata, P., Tomic, S. T., and Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75.

Jensenius, A. R. (2006). Using motiongrams in the study of musical gestures. In *Proceedings of the International Computer Music Conference* (pp. 499–502). New Orleans, LA: Tulane University.

Jiang, D.-N., Lu, L., Zhang, H.-J., Tao, J.-H., and Cai, L.-H. (2002). Music type classification by spectral contrast feature. In *Proceedings of the IEEE International Conference on Multimedia and Expo* (pp. 113–116). Lausanne, Switzerland: IEEE.

John, O. P., Naumann, L. P., and Soto, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, and L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.

John, O. P., and Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin and O. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). New York, NY: Guilford Press.

Kaiser, R., and Keller, P. E. (2011). Music's impact on the visual perception of emotional dyadic interactions. *Musicae Scientiae*, *15*(2), 270–287.

Keller, P. E. (2008). Joint action in music performance. In F. Morganti, A. Carassa, and G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions* (pp. 205–221). Amsterdam, the Netherlands: IOS Press.

Keller, P. E., and Rieger, M. (2009). Special issue - Musical movement and synchronization. *Music Perception*, *26*(5), 397–400.

Kim, H., Kwak, S. S., and Myungsuk Kim. (2008). Personality design of sociable robots by control of gesture design factors. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 494–499). Munich, Germany: IEEE.

Kluft, E. S., Poteat, J., and Kluft, R. P. (1986). Movement observations in multiple personality disorder: A preliminary report. *American Journal of Dance Therapy*, *9*(1), 31–46.

Koppensteiner, M., and Grammer, K. (2010). Motion patterns in political speech and their influence on personality ratings. *Journal of Research in Personality*, *44*(3), 374–379.

Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, *11*(2), 45–50.

Krumhansl, C. L., and Schenck, D. L. (1997). Can dance reflect the structural and expressive qualities of music. *Musicae Scientiae*, *1*(1), 63–85.

Kuppens, P. (2005). Interpersonal determinants of trait anger: Low agreeableness, perceived low social esteem, and the amplifying role of the importance attached to social relationships. *Personality and Individual Differences*, *38*(1), 13–23.

Lagerlöf, I., and Djerf, M. (2009). Children's understanding of emotion in dance. *European Journal of Developmental Psychology*, *6*(4), 409–431.

Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, *27*(2), 308–314.

Lakoff, G., and Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., and Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic Books.

Lamont, A. M., and Greasley, A. E. (2009). Musical preferences. In S. Hallam, I. Cross, and M. Thaut (Eds.), *The Oxford handbook of music psychology.* Oxford, UK: Oxford University Press.

Large, E. W., and Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*(1), 1–37.

Lartillot, O., Eerola, T., Toiviainen, P., and Fornari, J. (2008). Multi-feature modeling of pulse clarity: Design, validation, and optimization. In J. P. Bello, E. Chew, and D. Turnbull (Eds.), *Proceedings of the 9th International Conference on Music Information Retrieval* (pp. 521–526). Philadelphia, PA: Drexel University.

Lartillot, O., and Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. In *Proceedings of the 10th International Conference on Digital Audio Effects* (pp. 1–8). Bordeaux, France: University of Bordeaux.

Leman, M. (2007). *Embodied music cognition and mediation technology*. Cambridge, MA/London, UK: MIT Press.

Leman, M., and Godøy, R. I. (2010). Why study musical gesture? In R. I. Godøy and M. Leman (Eds.), *Musical gestures. Sound, movement, and meaning* (pp. 3–11). New York, NY: Routledge.

Leman, M., and Naveda, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in samba and Charleston. *Music Perception*, *28*(1), 71–92.

Lesaffre, M., Voogdt, L. de, Leman, M., Baets, B. de, Meyer, H. de, and Martens, J.-P. (2008). How potential users of music search and retrieval systems describe the semantic quality of music. *Journal of the American Society for Information Science and Technology*, *59*(5), 695–707.

Liberman, A. M., and Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

Luck, G., Saarikallio, S., Thompson, M. R., Burger, B., and Toiviainen, P. (2012). Do opposites attract? Personality and seduction on the dance floor. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, and K. Pastiadis (Eds.), *Proceedings of the 12th International Conference in Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences for*

Music (pp. 626–629). Thessaloniki, Greece: Aristotle University of Thessaloniki.

MacDougall, H. G., and Moore, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology*, *99*, 1164 –1173.

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, *24*(2), 201–208.

Madison, G., Gouyon, F., Ullén, F., and Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1578–1594.

Massaro, D. W., and Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*, *3*(2), 215–221.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G. D., and Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, *58*(3), 177–192.

McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae and J. Alik (Eds.), *The Five-Factor model of personality across cultures* (pp. 105–125). New York, NY: Kluwer Academic Publisher.

McCrae, R. R., and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90.

McCrae, R. R., Gaines, J. F., and Wellington, M. A. (2012). The five-factor model in fact and fiction. In I. Weiner, H. A. Tennen, and J. M. Suls (Eds.), *Handbook of psychology: Personality and social psychology.* Somerset, NJ: Wiley.

McCrae, R. R., and Terracciano, A. (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology*, *89*(3), 407–425.

Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, Australia* (pp. 580–583). Adelaide, Australia: Causal Productions.

Montepare, J. M., Goldstein, S. B., and Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behavior*, *11*(1), 33–42.

Müllensiefen, D., Gingras, B., Stewart, L., and Musil, J. J. (2012). *Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0: Technical report and documentation.* Goldsmiths, University of London, UK.

Murray, M. P., Drought, A. B., and Kory, R. C. (1964). Walking patterns of normal men. *Journal of Bone and Joint Surgery*, *46*(2), 335–360.

Naveda, L., and Leman, M. (2010). The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA). *Music Perception*, *28*(1), 93–112.

Neff, M., Wang, Y., Abbott, R., and Walker, M. (2010). Evaluating the effect of gesture and language on personality perception in conversational agents. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents, Philadelphia, PA. Lecture Notes in Computer Science, 6356* (pp. 222–235). Berlin/Heidelberg, Germany: Springer.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker, and S. Brown (Eds.), *The origins of music* (pp. 463–472). Cambridge, MA: MIT Press.

North, M. (1972). *Personality assessment through movement*. London, UK: MacDonald and Evans Limited.

Nymoen, K. (2013). *Methods and technologies for analysing links between musical sound and body motion*. PhD Dissertation, University of Oslo, Norway.

Nymoen, K., Caramiaux, B., Kozak, M., and Torresen, J. (2011). Analyzing sound tracings: A multimodal approach to music information retrieval. In C. C. S. Liem, M. Müller, D. Eck, and G. Tzanetakis (Eds.), *Proceedings of the 1st International ACM Workshop on Music Information Retrieval with User-centered and Multimodel Strategies* (pp. 39–44). Scottsdale, AZ: ACM Press.

Nymoen, K., Torresen, J., Godøy, R. I., and Jensenius, A. R. (2012). A statistical approach to analyzing sound tracings. In S. Ystad, M. Aramaki, R. Kronland-Martinet, K. Jensen, and S. Mohanty (Eds.), *Speech, sound and music processing: Embracing research in India, Lecture Notes in Computer Science, 7172* (pp. 120–145). Berlin/Heidelberg, Germany: Springer.

Pampalk, E., Rauber, A., and Merkl, D. (2002). Content-based organization and visualization of music archives. In *Proceedings of the 10th ACM International Conference on Multimedia, Juan-les-Pins, France* (pp. 570–579). New York, NY: ACM Press.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*(4), 409–464.

Petrini, K., McAleer, P., and Pollick, F. (2010). Audiovisual integration of emotional signals from music improvisation does not depend on temporal correspondence. *Brain Research*, *1323*, 139–148.

Pfeifer, R., and Scheier, C. (1999). *Understanding intelligence*. Cambridge, MA, London, UK: MIT Press.

Phillips-Silver, J., and Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*, 94–102.

Pollick, F. E., Paterson, H. M., Bruderlin, A., and Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*, B51–B61.

Rentfrow, P. J., and 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.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.

Repp, B. H., and Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Robertson, D. G. E., Caldwell, G. E., Hamill, J., Kamen, G., and Whittlesey, S. N. (2004). *Research methods in biomechanics*. Champaign, IL/Leeds, UK: Human Kinetics.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Ryckman, R. M. (2008). *Theories of personality*. Belmont, CA: Thomson Wadsworth.

Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music*, *39*(3), 307–327.

Saarikallio, S., Luck, G., Burger, B., Thompson, M. R., and Toiviainen, P. (2013). Dance moves reflect current affective state illustrative of approach-avoidance motivation. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(3), 296–305.

Savitzky, A., and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review*, *61*(2), 81–88.

Schubert, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, *51*(3), 154–165.

Schumacker, R. E., and Lomax, R. G. (2010). *A beginner's guide to structural equation modeling*. London, UK: Routledge Academic.

Shapiro, L. (2011). *Embodied cognition*. London, UK/New York, NY: Routledge.

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Silla Jr., C. N., Koerich, A. L., and Kaestner, C. A. A. (2009). A feature selection approach for automatic music genre classification. *International Journal of Semantic Computing*, *3*(2), 183–208.

Sloboda, J. A. (2010). Music in everyday life. In P. N. Juslin and J. A. Sloboda (Eds.), *Handbook of music and emotion. Theory, research, applications* (pp. 493–514). Oxford, UK: Oxford University Press.

Sörgjerd, M. (2000). *Auditory and visual recognition of emotional expression in performances of music*. Unpublished thesis, University of Uppsala, Sweden.

Stevens, C. J., Schubert, E., Wang, S., Kroos, C., and Halovic, S. (2009). Moving with and without music: Scaling and lapsing in time in the performance of contemporary dance. *Music Perception*, *26*(5), 451–464.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., and Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, *82*(2), 127–136.

Thompson, M. R., and Luck, G. (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, *16*(1), 19–40.

Todd, N. P. M., Cousins, R., and Lee, C. S. (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Musicology Review*, *2*(1), 1–13.

Toiviainen, P., Luck, G., and Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, *28*(1), 59–70.

Tomasello, M., and Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*(1), 121–125.

Trainor, L. J., Gao, X., Lei, J.-j., Lehtovaara, K., and Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex*, *45*(1), 35–43.

Tzanetakis, G., and Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, *10*(5), 293–302.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., and Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception*, *30*(4), 349–359.

Van Zijl, A. G. W., and Luck, G. (2013). Moved through music: The effect of experienced emotions on performers' movement characteristics. *Psychology of Music*, *41*(2), 175–197.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., and Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*, 80–113.

Vuoskoski, J. K., and Eerola, T. (2011a). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159–173.

Vuoskoski, J. K., and Eerola, T. (2011b). The role of mood and personality in the perception of emotions represented by music. *Cortex*, *47*(9), 1099–1106.

Walk, R. D., and Homan, C. P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society*, *22*, 437–440.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*(6), 879–896.

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., and Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, *34*(1), 97–113.

Wilson, R. A., and Foglia, L. (2011). Embodied cognition. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* Stanford, CA: Stanford University. [http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition/].

Witek, M. A. G., Clarke, E. F., Wallentin, M., Hans, M., Kringelbach, M. L., and Vuust, P. (2012). "...and I feel good!" Ratings, fMRI-recordings and motion-capture measurements of body-movements and pleasure in response to groove. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, and K. Pastiadis (Eds.), *Oral presentation at the 12th International Conference in Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences for Music* (pp. 152–153). Thessaloniki, Greece: Aristotle University of Thessaloniki.

Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews. Neuroscience*, *8*(7), 547–558.

Zentner, M., and Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 5768–5773.

Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521.

APPENDIX 1 SONG LIST

	ARTIST	SONG (ALBUM)	PART	BPM
1.	Alice Deejay	Better Off Alone (Who Needs Guitars Anyway?)	2:40-2:54*	137
2.	Andre Visior	Speed Up	1:15-1:45	140
3.	Antibalas	Who is this America Dem Speak of Today? (Who Is This America?)	1:00-1:29	121
4.	Arturo Sandoval	A Mis Abuelos (Danzon)	1:53-2:21	108
5.	Baden Powell	Deixa (Personalidade)	1:11-1:41	100
6.	Brad Mehldau	Wave/Mother Nature's Son (Largo)	0:00-0:29	143
7.	Clifford Brown & Max Roach	The Blues walk (Verve Jazz Masters, Vol. 44: Clifford Brown & Max Roach)	2:01-2:31	133
8.	Conjunto Imagen	Medley-Esencia de Guaguanco/ Sonero (Ayer, Hoy y Manana)	2:18-2:48	87
9.	Dave Hillyard & The Rocksteady 7	Hillyard Street (Playtime)	0:15-0:45	135
10.	Dave Weckl	Mercy, Mercy, Mercy (Burning for Buddy)	0:10-0:40	105
11.	Dave Weckl	Tower of Inspiration (Master Plan)	0:00-0:30	125
12.	DJ Shadow	Napalm Brain/Scatter Brain (Endtroducing...)	3:29-3:58	73
13.	Gangster Politics	Gangster Politics (Guns & Chicks)	1:00-1:29	192
14.	Gigi D'Agostino	Blablabla (L'Amour Toujours)	0:00-0:30	133
15.	Herbie Hancock	Watermelon man (Cantaloupe Island)	0:00-0.30	132
16.	Horace Silver	The Natives Are Restless	0:00-0:29	139
17.	In Flames	Scream (Come Clarity)	0:00-0:30	100
18.	Jean Roch	Can You Feel it (Club Sounds Vol. 35)	0:33-1:01	126
19.	Johanna Kurkela	Hetki hiljaa (Hetki hiljaa)	3:22-3:52	122
20.	Juana Molina	Tres cosas (Tres Cosas)	0:00-0:30	110
21.	Kings of Leon	Closer (Only by the Night)	3:17-3:47	83
22.	Lenny Kravitz	Live (5)	3:02-3:30	84
23.	Martha & The Vandellas	Heat Wave (Heat Wave)	1:40-2:10	82
24.	Maynard Ferguson	Fireshaker (Live From San Francisco)	0:00-0:28	91
25.	MIA	20 Dollar (Kala)	0:17-0:45	120
26.	Nick Beat	Techno Disco	2:26-2:56	138
27.	Panjabi MC	Mundian To Bach Ke (Legalised)	0:47-1:06*	98
28.	Patrick Watson	Beijing (Wooden Arms)	2:30-2:59	154
29.	The Rippingtons	Weekend in Monaco (Weekend in Monaco)	1:13-1:42	113
30.	Yuri Buenaventura	Salsa (Salsa Movie Soundtrack)	2:17-2:45	102

* loop

ORIGINAL PAPERS

I

INFLUENCES OF RHYTHM- AND TIMBRE-RELATED MUSICAL FEATURES ON CHARACTERISTICS OF MUSIC-INDUCED MOVEMENT

by

Birgitta Burger, Marc R. Thompson, Suvi Saarikallio, Geoff Luck & Petri Toiviainen, 2013

Frontiers in Psychology vol. 4:183, 1-10

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 12 April 2013
doi: 10.3389/fpsyg.2013.00183

Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement

Birgitta Burger*, Marc R. Thompson, Geoff Luck, Suvi Saarikallio and Petri Toiviainen

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

Edited by:
Chris Muller, Ghent University, Belgium

Reviewed by:
Henning Scheich, Leibniz Institute for Neurobiology Magdeburg, Germany
Psyche Loui, Harvard Medical School, USA

***Correspondence:**
Birgitta Burger, Department of Music, Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, P.O. Box 35, FI-40014 Jyväskylä, Finland.
e-mail: birgitta.burger@jyu.fi

Music makes us move. Several factors can affect the characteristics of such movements, including individual factors or musical features. For this study, we investigated the effect of rhythm- and timbre-related musical features as well as tempo on movement characteristics. Sixty participants were presented with 30 musical stimuli representing different styles of popular music, and instructed to move along with the music. Optical motion capture was used to record participants' movements. Subsequently, eight movement features and four rhythm- and timbre-related musical features were computationally extracted from the data, while the tempo was assessed in a perceptual experiment. A subsequent correlational analysis revealed that, for instance, clear pulses seemed to be embodied with the whole body, i.e., by using various movement types of different body parts, whereas spectral flux and percussiveness were found to be more distinctly related to certain body parts, such as head and hand movement. A series of ANOVAs with the stimuli being divided into three groups of five stimuli each based on the tempo revealed no significant differences between the groups, suggesting that the tempo of our stimuli set failed to have an effect on the movement features. In general, the results can be linked to the framework of embodied music cognition, as they show that body movements are used to reflect, imitate, and predict musical characteristics.

Keywords: music-induced movement, dance, motion capture, musical feature extraction, pulse clarity, spectral flux

INTRODUCTION

Music makes us move. While listening to music, we often move our bodies in a spontaneous fashion. Keller and Rieger (2009), for example, stated that simply listening to music can induce movement, and in a self-report study conducted by Lesaffre et al. (2008), most participants reported moving when listening to music. Janata et al. (2011) reported a study in which they asked participants to tap to the music and found that participants not only moved the finger/hand, but also other body parts, such as feet and head. Additionally, the tapping condition (isochronous versus free tapping) influenced the amount of movement: the more "natural" the tapping condition, the more movement was exhibited.

In general, people tend to move to music in an organized way by, for example, mimicking instrumentalists' gestures, or rhythmically synchronizing with the pulse of the music by tapping their foot, nodding their head, or moving their whole body in various manners (Godøy et al., 2006; Leman and Godøy, 2010). Moreover, Leman (2007, p. 96) suggests, "Spontaneous movements [to music] may be closely related to predictions of local bursts of energy in the musical audio stream, in particular to the beat and the rhythm patterns." Such utilization of the body is the core concept of embodied cognition, which claims that the body is involved in or even required for cognitive processes (e.g., Lakoff and Johnson, 1980, 1999, or Varela et al., 1991). Human cognition is thus highly influenced by the interaction between mind/brain, sensorimotor capabilities, body, and environment. Following this, we can approach music (or musical involvement) by linking our perception of it to our body movement (Leman, 2007). One could postulate that our bodily movements reflect, imitate, help to parse, or support understanding the structure and content of music. Leman suggests that corporeal articulations could be influenced by three (co-existing) components or concepts: "Synchronization," "Embodied Attuning," and "Empathy," which differ in the degree of musical involvement and in the kind of action-perception couplings employed. "Synchronization" forms the fundamental component, as synchronizing to a beat is easy and spontaneous. The beat serves as the basic musical element, from which more complex structures emerge. Leman furthermore suggests the term "inductive resonance," referring to the use of movements for active control, imitation, and prediction of beat-related features in the music (the opposite of passively tapping to a beat) as the first step in engaging with the music. The second component, "Embodied Attuning," concerns the linkage of body movement to musical features more complex than the basic beat, such as melody, harmony, rhythm, tonality, or timbre. Following this idea, movement could be used to reflect, imitate, and navigate within the musical structure in order to understand it. Finally, "Empathy" is seen as the component that links musical features to expressivity and emotions. In other words, the listener feels and identifies with the emotions expressed in the music and imitates and reflects them by using body movement.

It has been argued that music and dance have evolved together in most cultures (Arom, 1991; Cross, 2001) and are crucial elements of most social and collective human behavior (Brown et al., 2000). Furthermore, most cultures have developed coordinated dance movements to rhythmically predictable music (Nettl, 2000). There is neurobiological evidence for a connection between rhythmic components of music and movement (e.g., Grahn and Brett, 2007; Bengtsson et al., 2009; Chen et al., 2009; Grahn and Rowe, 2009), which has led to the assumption that humans prefer music that facilitates synchronization and respond to it with movement (Madison et al., 2011). Phillips-Silver and Trainor (2008) showed in their study that especially head movements were found to bias metrical encoding of rhythm and meter perception.

The increasing opportunities of quantitative research methods for recording and analyzing body movement have offered new insights and perspectives for studying such movements. A number of studies have investigated (professional) dance movements using quantitative methods. Camurri et al. (2003, 2004), for instance, developed a video analysis tool to recognize and classify expressive and emotional gestures in professional dance performances. Jensenius (2006) developed the technique of "motiongrams" for visualizing and analyzing movement and gestures. Stevens et al. (2009) studied movements of professional dancers regarding time keeping with and without music using optical motion capture recordings. Optical motion capture was also employed by Naveda and Leman (2010) and Leman and Naveda (2010) who investigated movement in samba and Charleston dancing, focusing on spatiotemporal representations of dance gestures as movement trajectories.

Besides professional dance, several studies were dedicated to more general tasks and behaviors involving music-induced movement, such as movement of infants or laymen dancers. Zentner and Eerola (2010) investigated infants' ability to bodily synchronize with musical stimuli, finding that infants showed more rhythmic movement to music and metrical stimuli than to speech suggesting a predisposition for rhythmic movement to music and other metrical regular sounds. Eerola et al. (2006) studied toddlers' corporeal synchronization to music, finding three main periodic movement types being at times synchronized with pulse of the music. Toiviainen et al. (2010) investigated how music-induced movement exhibited pulsations on different metrical levels, and showed that eigenmovements of different body parts were synchronized with different metric levels of the stimulus. Luck et al. (2010) studied the influence of individual factors such as personality traits and preference on musically induced movements of laypersons' dancing, finding several relationships between personality traits and movement characteristics. Van Dyck et al. (2010) found that an increased presence of the bass drum tends to increase listeners' spontaneous movements. However, systematic investigations targeting the relationships between musical features, particularly rhythm-related features, and human movement characteristics have not been conducted. Such an investigation would reveal additional information as to how music shapes movement. Additionally, finding dependencies of musical structure and body movements that are consistent between individuals would support the notion of embodied music cognition (Leman, 2007).

Rhythmic music is based on beats, which can be physically characterized as distinct energy bursts in time. If such beats occur as regular and repetitive temporal patterns, they give rise to a percept of pulse. Beat and pulse structures can be regarded as the basic metrical structure in music from which more complex temporal structures, such as rhythm, emerge. This is typically achieved by subdividing the basic metrical structure in smaller and larger units of varying lengths, constructing a metrically interlocked grid with events on different temporal levels (Parncutt, 1994). These rhythmic structures can vary, for example, in the degree of pulse clarity. Pulse clarity estimates, on a large time scale, how clearly the underlying pulsation in music is perceivable and can therefore be regarded as a measure for the underlying periodicity of the music (Lartillot et al., 2008). Another aspect of rhythmic structure is covered by spectro-temporal features, such as the sub-band spectral flux, which has been found to be among the most important features contributing to polyphonic timbre perception (Alluri and Toiviainen, 2010). Spectral flux measures spectral change, which, when taken separately for different sub-bands, is related to certain elements of the rhythm (i.e., rhythmic elements created by instruments within the frequency range of the respective sub-band). It could be that sub-band flux is a crucial feature not only in a (passive) listening situation, but also in a (active) movement situation. Furthermore, other timbral characteristics, such as percussiveness, could have an influence on movement responses to music. For instance, high amounts of percussive elements in music could result in fast movements, reflecting the way such sounds are often produced. Following these notions, it could be assumed that variations in such musical features not only increase or decrease the amount of movement, but also change the kinds and properties of the movements. In line with the embodied music cognition approach (Leman, 2007), the movements could reflect and imitate the rhythmical and timbral structure of the music.

Besides features such as pulse clarity, tempo is an important factor contributing to the perception of rhythm (Fraisse, 1982). Tempo is the speed at which beats are repeated, the underlying periodicity of music. Tempo is usually measured in beats per minute (bpm). A large body of research has been conducted on listeners' abilities to perceive different tempi, synchronize to them, and reproduce them in tapping tasks (for a review see Repp, 2005). Free tapping tasks have found that a majority of participants tapped at a rate close to 600 ms, though the individual rates differed considerably (Fraisse, 1982). Synchronizing to steady, periodic beat stimuli is possible at a wide range of tempi, however it is most regular and accurate for intervals around 400–500 ms (Collyer et al., 1992), respectively 400–800 ms (Fraisse, 1982), while with slower and faster tempo the time between two taps becomes more variable. Moelants (2002) suggested 120 bpm as the preferred tempo – the tempo where tempo perception is considered to be optimal and appears most natural. Interesting to note here is that literature often draws links between spontaneous/preferred tempo and repeated motor activities, such as walking, for which the spontaneous duration of steps is around 500–550 ms (Murray et al., 1964; Fraisse, 1982; MacDougall and Moore, 2005). Walking has been suggested as "a fundamental element of human motor activity" (Fraisse, 1982, p. 152) and could therefore serve as an origin of preferred tempo perception. Following these considerations, it could furthermore be assumed that music with tempi around 110–120 bpm stimulate movement more than music with other tempi.

In order to investigate relationships between rhythmic and timbral aspects of music and movements that such music induces, we conducted a motion capture experiment, in which participants were asked to move to music that differed in tempo and in rhythm- and timbre-related features, such as pulse clarity, percussiveness, and spectral flux of low and high frequency ranges. We were interested in two main aspects: first, whether movement and musical characteristics are somehow related, and more particular, which body parts and movement types reflect different musical characteristics. Second, whether the tempo of music has an influence on the movement. Following the notion of embodied (music) cognition we assumed the movement to reflect aspects of the music. First, we expected movements to indicate the beat structure, in particular that a clear beat would be embodied by increased speed of movements. Second, we anticipated that hand and arm movements, having the most freedom when moving/dancing, would play an important role in reflecting musical characteristics. Finally, we assumed movement features to resemble the preferred tempo insofar, as tempi around 120 bpm might encourage participants to move differently than tempi slower or faster than 120 bpm.

MATERIALS AND METHODS

PARTICIPANTS

A total of 64 participants took part in the study. Four participants were excluded from further analysis due to incomplete data. Thus, 60 participants remained for subsequent analyses (43 female, 17 male, average age: 24, SD of age: 3.3). Six participants had received formal music education, while four participants had a formal background in dance tuition. Participation was rewarded with a movie ticket. All participants gave their informed consent prior to their inclusion in the study and they were free to discontinue the experiment at any point. Ethical permission for this study was not needed, which was according to the guidelines stated by the university ethical board.

STIMULI

Participants were presented with 30 randomly ordered musical stimuli representing the following popular music genres: techno, Pop, Rock, Latin, Funk, and Jazz. An overview of the stimuli can be found in the Appendix. All stimuli were 30 s long, non-vocal, and in 4/4 time, but differed in their rhythmic complexity, pulse clarity, mode, and tempo. The stimulus length was chosen to keep the experiment sufficiently short while having stimuli that were long enough to induce movement.

APPARATUS

Participants' movements were recorded using an eight-camera optical motion capture system (Qualisys ProReflex), tracking, at a frame rate of 120 Hz, the three-dimensional positions of 28 reflective markers attached to each participant. The locations of the markers can be seen in **Figures 1A,B**. The location of the markers were as follows (L, left; R, right; F, front; B, back): 1, LF head; 2, RF head; 3, LB head; 4, RB head; 5, L shoulder; 6, R shoulder; 7, sternum; 8, spine (T5); 9, LF hip; 10, RF hip; 11, LB hip; 12, RB hip; 13, L elbow; 14, R elbow; 15, L wrist/radius; 16, L wrist/ulna; 17, R wrist/radius; 18, R wrist/ulna; 19, L middle finger; 20, R middle finger; 21, L knee; 22, R knee; 23, L ankle; 24, R ankle; 25, L heel; 26, R heel; 27, L big toe; 28, R big toe. The musical stimuli were played back via a pair of Genelec 8030A loudspeakers using a Max/MSP patch running on an Apple computer. The room sound was recorded with two overhead microphones positioned at a height of approximately 2.5 m. This microphone input, the direct audio signal of the playback, and the synchronization pulse transmitted by the Qualisys cameras when recording, were recorded using ProTools software in order to synchronize the motion capture data with the musical stimulus afterward. Additionally, four Sony video cameras were used to record the sessions for reference purposes.

PROCEDURE

Participants were recorded individually and were asked to move to the presented stimuli in a way that felt natural. Additionally, they were encouraged to dance if they wanted to, but were requested to remain in the center of the capture space indicated by a 115 × 200 cm carpet.

MOVEMENT FEATURE EXTRACTION

In order to extract various kinematic features, the MATLAB Motion Capture (MoCap) Toolbox (Toiviainen and Burger, 2011)

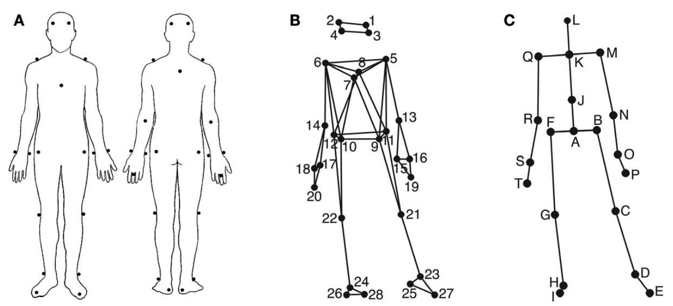


FIGURE 1 | Marker and joint locations. (A) Anterior and posterior view of the marker placement on the participants' bodies; **(B)** Anterior view of the marker locations as stick figure illustration; **(C)** Anterior view of the locations of the secondary markers/joints used in the analysis.

was used to first trim the data to the duration of each stimulus and, following this, derive a set of 20 secondary markers – subsequently referred to as joints – from the original 28 markers. The locations of these 20 joints are depicted in **Figure 1C**. The locations of joints C, D, E, G, H, I, M, N, P, Q, R, and T are identical to the locations of one of the original markers, while the locations of the remaining joints were obtained by averaging the locations of two or more markers; joint A, midpoint of the four hip markers (called root in the further analysis); B, midpoint of markers 9 and 11 (left hip); F, midpoint of markers 10 and 12 (right hip); J, midpoint of breastbone, spine, and the hip markers (midtorso); K, midpoint of shoulder markers (manubrium), L, midpoint of the four head markers (head); O, midpoint of the two left wrist markers (left wrist); S, midpoint of the two right wrist markers (right wrist). Thereafter, eight movement variables were extracted from the joint location data to cover different movement characteristics, body parts, and movement types.

- Speed of Center of Mass, Head, both Hands, and both Feet: for calculating the speed of the movement, we applied numerical differentiation based on the Savitzky and Golay (1964) smoothing FIR filter with a window length of seven samples and a polynomial order of two. These values were found to provide an optimal combination of precision and smoothness in the time derivatives. For the Speed of Head, Hands, and Feet, the data were transformed into a local coordinate system, in which joint A was located at the origin, and segment BF had zero azimuth.
- Hand Distance: this feature relates to the distance between both hands.
- Amount of Movement: this feature is based on the traveled distance of all markers and gives a measure for the total amount of movement (data were also transformed into the local coordinate system).
- Hip Wiggle: this feature is defined as the mean absolute angular velocity of the hips around the anteroposterior axis.
- Shoulder Wiggle: this feature is defined as the mean absolute angular velocity of the shoulder around the anteroposterior axis.

Subsequently, the instantaneous values of each variable were averaged for each stimulus presentation. This yielded a total of eight statistical movement features for each of the 60 participants and 30 stimuli.

MUSICAL FEATURE EXTRACTION

To investigate the effect of rhythm- and timbre-related features on music-induced movement, we performed computational feature extraction analysis of the stimuli used in the experiment. To this end, four musical features were extracted from the stimuli using the MATLAB MIRToolbox (version 1.4) (Lartillot and Toiviainen, 2007).

- Pulse Clarity: this feature indicates the strength of rhythmic periodicities and pulses in the signal, estimated by the relative Shannon entropy of the fluctuation spectrum (Pampalk et al., 2002). In this context, entropy can be understood as a measure of the degree of peakiness of the spectrum. Music with easily perceived beat has a distinct and regular fluctuation spectrum, which has low entropy. Thus, high pulse clarity is associated with low fluctuation entropy. To illustrate this measure of Pulse Clarity, the fluctuation spectra for high and low Pulse Clarity are shown in **Figure 2**.
- Sub-Band Flux: this feature indicates to which extend the spectrum changes over time. For the calculation, the stimulus is divided into 10 frequency bands, each band containing one octave in the range of 0–22050 Hz. The Sub-Band Flux is then calculated for each of these ten bands separately by calculating the Euclidean distances of the spectra for each two consecutive frames of the signal (Alluri and Toiviainen, 2010), using a frame length of 25 ms and an overlap of 50% between successive frames and then averaging the resulting time-series of flux values. For the current analysis, we used two frequency bands: band no. 2 (50–100 Hz) and band no. 9 (6400–12800 Hz). High flux in the low frequency bands is produced by instruments such as kick drum and bass guitar, whereas high flux in the high frequency bands is produced by instruments such as cymbals or

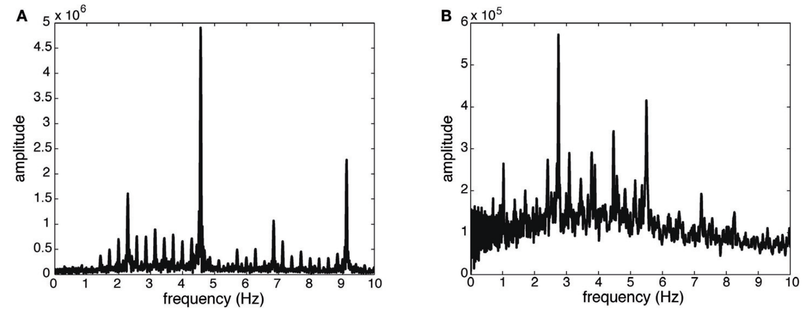


FIGURE 2 | Fluctuation spectra of two stimuli used in the study. (A) Peaks at a regular distance of 0.28 Hz, with the highest peak at 4.56 Hz and other clear peaks at 2.29, 6.85, and 9.13 Hz, suggesting clear pulses and periodicity (stimulus 1, see Appendix). **(B)** Markedly lower magnitude values, a less periodic pattern of peaks, and more noise, suggesting low pulse clarity (stimulus 21, see Appendix).

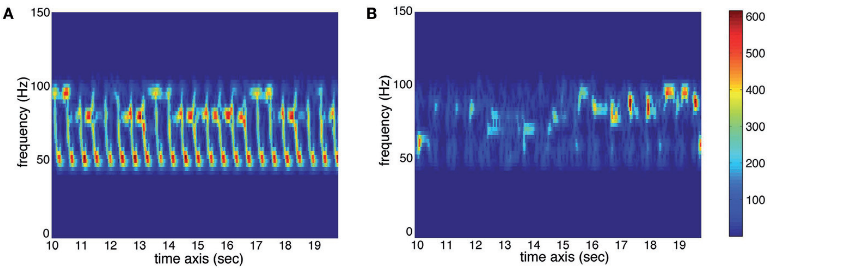


FIGURE 3 | Spectrograms (sec. 10–20) of sub-band no. 2 (50–100 Hz) of two stimuli used in the study. (A) High amount of temporal change (red represents high energy at the respective time and frequency, whereas blue represents low energy; see color bar) resulting in high value for Sub-Band Flux (stimulus 26, see Appendix). **(B)** Low amount of temporal change resulting in low Sub-Band Flux (stimulus 5, see Appendix).

hihats. Two spectrograms of sub-band no. 2 are displayed in **Figure 3** to show the difference between high and low amounts of Sub-Band Flux.

- Percussiveness: for this feature, the slopes of the amplitude envelopes at note onsets are estimated and then averaged across all slopes of the signal. The steeper the slope, the more percussive and accentuated the sound.

Additionally, the Tempo of the stimuli was assessed in a separate tapping experiment, in which ten participants tapped to the 30 stimuli. The average tempo was evaluated by taking the median value of all intertap intervals.

RESULTS

In order to perform further analysis, we first checked for consistency between participants by calculating Intraclass Correlations (cf., Shrout and Fleiss, 1979) for each movement feature separately. Intraclass correlations operate on data that are structured in groups (as opposed to standard correlation procedures that assume data to be structured as paired observations) and indicate how strongly units of the same group (in this case the values of one movement feature for all participants and songs) resemble each other. The values are presented in **Table 1**.

All intraclass correlations resulted in highly significant coefficients. Therefore, we concluded that participants' movements for each song were similar enough to justify averaging the movement features across participants, yielding one value for each stimulus presentation.

Subsequently, we correlated the movement features with Pulse Clarity, Low and High Frequency Spectral Flux, and Percussiveness. Due to the relatively large number of correlations (32), we applied Bonferroni correction to avoid false positive errors in case of multiple comparisons. The results are displayed in **Table 2**.

As can be seen, Pulse Clarity shows significant positive correlations with Speed of Center of Mass [$r(30) = 0.67$, $p < 0.01$], Speed of Feet [$r(30) = 0.55$, $p < 0.05$], Amount of Movement [$r(30) = 0.62$, $p < 0.01$], Hip Wiggle [$r(30) = 0.58$, $p < 0.05$], and Shoulder Wiggle [$r(30) = 0.67$, $p < 0.01$]. Thus, for music with a clear pulse, participants tended to move the center of the body and feet faster, wiggled more with hips and shoulders, and used an increased amount of movement related to the whole body. For an illustration of these features, an animation of movements performed by one participant for the stimulus with the highest value for Pulse Clarity (stimulus 1, see Appendix) can be found in the

Table 1 | Results of the intraclass correlations for each of the movement features.

Movement feature	*r*
Speed of center of mass	0.90***
Speed of head	0.89***
Speed of hands	0.91***
Speed of feet	0.90***
Hand distance	0.62***
Amount of movement	0.94***
Hip wiggle	0.95***
Shoulder wiggle	0.88***

***$p < 0.001$.

Table 2 | Results of the correlation between movement and musical features.

	Pulse clarity	Spectral flux SB 2	Spectral flux SB 9	Percussiveness
Speed of center of mass	**0.67****	0.51	0.52	**0.56***
Speed of head	0.52	**0.73*****	**0.64****	**0.67****
Speed of hands	0.54	0.46	**0.65****	**0.57***
Speed of feet	**0.55***	0.31	0.46	0.45
Hand distance	0.34	0.29	**0.57***	**0.55***
Amount of movement	**0.62****	0.44	**0.57***	**0.56***
Hip wiggle	**0.58***	0.25	0.47	0.46
Shoulder wiggle	**0.67****	0.39	0.48	**0.61***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

"PulseClarity.mov" in Supplementary Material. The selected participant exhibited high values for all the movement features that were shown to be characteristic for high Pulse Clarity.

Spectral Flux of Sub-band no. 2 exhibited a high positive correlation with Speed of Head [$r(30) = 0.73$, $p < 0.001$], suggesting that stimuli with strong spectral flux in the low frequencies are related to fast head movements. For illustration, an animation of movements performed by one participant for the stimulus with the highest value for Spectral Flux of Sub-band no. 2 (stimulus 26, see Appendix) can be found in the "SubBandFluxNo2.mov" in Supplementary Material.

Spectral Flux of Sub-band no. 9 showed significant positive correlations with Speed of Head [$r(30) = 0.64$, $p < 0.01$], Speed of Hands [$r(30) = 0.65$, $p < 0.01$], Hand Distance [$r(30) = 0.57$, $p < 0.05$], and Amount of Movement [$r(30) = 0.57$, $p < 0.05$], indicating that, for stimuli with strong spectral flux in the high frequencies, participants moved their head and hands faster, had a larger distance between the hands, and used an increased amount of movement. To illustrate these features, an animation of movements performed by one participant for the stimulus with the highest value for Spectral Flux of Sub-band no. 9 (stimulus 18, see Appendix) can be found in the "SubBandFluxNo9.mov" in Supplementary Material.

Percussiveness exhibited positive correlations with Speed of Center of Mass [$r(30) = 0.56$, $p < 0.05$], Speed of Head [$r(30) = 0.67$, $p < 0.01$], Speed of Hands [$r(30) = 0.57$, $p < 0.05$], Hand Distance [$r(30) = 0.55$, $p < 0.05$], Amount of Movement [$r(30) = 0.56$, $p < 0.05$], and Shoulder Wiggle [$r(30) = 0.61$, $p < 0.01$]. Thus, for stimuli containing a high amount of percussive elements, participants moved their center of mass, head, and hands faster, had a larger distance between their hands, used an increased amount of movement, and wiggled more with their shoulders. For an illustration of these features, an animation of movements performed by one participant for the stimulus with the highest value for Percussiveness (stimulus 3, see Appendix) can be found in the "Percussiveness.mov" in Supplementary Material.

As we assumed a non-linear relationship between Tempo and the movement features, we excluded Tempo from the correlational analysis. Instead, we rank-ordered the stimuli based on their tempi obtained from the tapping experiment and picked the five slowest, the five medial, and the five fastest stimuli to perform a series of ANOVAs on these three groups. The tempi of the stimuli in these groups ranged from 73 to 87 bpm, from 113 to 125 bpm, and from 138 to 154 bpm. The results of the ANOVAs are shown in **Table 3**.

As none of the ANOVAs showed significant differences between the groups, the tempo of our set of stimuli failed to have any significant effect on the movement features.

Table 3 | Results from the series of ANOVAs, testing for differences in the movement features between the three tempo groups.

	$F(2, 897)$	p
Speed of center of mass	0.36	0.70
Speed of head	0.55	0.58
Speed of hands	0.58	0.56
Speed of feet	0.40	0.67
Hand distance	0.06	0.94
Amount of movement	1.35	0.26
Hip wiggle	1.64	0.19
Shoulder wiggle	0.42	0.66

DISCUSSION

We investigated music-induced movement, focusing on the relationship between rhythm- and timbre-related musical features, such as Pulse Clarity and Spectral Flux, and movement characteristics of different body parts.

Pulse Clarity seemed to be embodied by using various movement types of different body parts, such as Speed of Center of Mass and Feet, Amount of Movement, and Hip and Shoulder Wiggle. Participants used an increasing amount of different movements and movement types of the whole body when the music contained easily and clearly perceivable pulses. Pulse Clarity might, therefore, influence body movement on a somewhat global, whole body level.

Low frequency (50–100 Hz) Spectral Flux was positively correlated to the Speed of Head. We observed that a high amount of low frequency Spectral Flux was associated with the presence of kick drum and bass guitar, creating strong low frequency rhythmic elements that are usually related to the beat period. A reason for the head being involved could be that, based on biomechanical properties, the head might be most prone to move to the beat level (cf. nodding to the beat or "head banging" in certain musical genres). Interesting to note here is that we found only one significant correlation, so it could be concluded that spectral changes in the low frequency ranges influences body movement on a rather local, specific level.

Spectral Flux in the high frequencies (6400–12800 Hz) was found to relate to Speed of Head and Hands, Hand Distance, and Amount of Movement. We observed that high frequency Spectral Flux was mainly influenced by the presence of hi-hat and cymbal sounds, creating the high frequency rhythmical figures and hence the "liveliness" of the rhythm. That could explain the dominance of hand-related movement characteristics, as the hands have the biggest movement freedom of the whole body and can therefore reflect fast rhythmic structures best.

Percussiveness correlated with Speed of Center of Mass, Head, and Hands, Hand Distance, Amount of Movement, and Shoulder Wiggle, suggesting that percussive elements in the music were related to movement of the upper body, especially to the hands as the most flexible body parts. Together with the findings for high frequency Spectral Flux, these results could indicate that timbral features not only affect movement responses to music, but more particular that the embodiment of these features is to a large proportion related to upper body and hand movement. Furthermore, high amounts of percussive elements seemed to be embodied with fast movement, supporting the assumption that the movement reflects the way such sounds are often produced.

The Tempo of the music failed to show any relationship with the movement features and could therefore not confirm our hypothesis that music around 120 bpm would stimulate movement differently than faster or slower music. However, this might be due to our selection of the stimuli. The variety of musical styles that was covered by the stimuli might have undermined the influence of the tempo. This issue thus remains an open question that requires further investigation.

The results can be linked to the framework of embodied music cognition and could provide support for Leman's (2007) concepts of Synchronization, Inductive Resonance, and Embodied Attuning. The concepts of Synchronization and Inductive Resonance can be related to our findings regarding Pulse Clarity: clearly perceivable and strong beats might have a resonating effect on the overall body/torso movement and amount of movement, reflecting and imitating the clear beat structure, and making participants synchronize to the pulse. With less clear pulses, people might be less affected by the resonating effect, and are thus less encouraged to move. Concerning the correlations with Pulse Clarity, it is interesting to note that Speed of Feet correlated significantly with this feature, but not with any other musical feature. A possible explanation for this connection could be that foot movement is usually related to walking, with steps at a regular pace. In a musical context, the pulse of the music is related to a regular pace, so it could be argued that a clear and regular pulse of the music was embodied with using more extensive and synchronized foot movement, that is, stepping to the pulse (see Styns et al. (2007) for further connections between walking and beat patterns). However, our movement features do not reveal information about the actual synchronization to the pulse, and thus further analysis has to be performed to justify this interpretation. The torso movement (Center of Mass Speed, Hip Wiggle, Shoulder Wiggle) could also be part of the stepping-/walking-type movement of legs/feet, which would explain the significant correlations with Pulse Clarity. Furthermore, the feet are required to provide stability and an upright position, so they cannot move as free as, for instance, the hands. Thus, a connection of these body parts to the fundamental components of musical engagement, such as synchronization and resonance, makes sense, whereas the upper body parts (e.g., hands and head) could be expected to be more related to timbral and rhythmical structures of music (cf., Embodied Attuning, see next paragraph).

Moreover, the results could serve as an example for the concept of Embodied Attuning – movement-based navigation through the musical/rhythmic structures of the stimuli created by Sub-Band Flux and Percussiveness. It could be suggested that participants attune to strong spectral changes in low and high frequency ranges and to percussive elements with mainly head and hand movements, as these musical characteristics (related to rhythm) might be reflected and imitated best by using the upper extremities of the body (hand, head, and shoulder movement). Especially the hands (together with arms and shoulders) could be used to navigate through the music and to parse and understand the rhythmic/musical structure better.

Besides having the biggest freedom in movement as mentioned previously, the relation between hand/arm-related movement and several musical features might also occur due to knowledge and imagination of playing a musical instrument (Leman, 2007). One could postulate that participants have applied their own experience in playing an instrument to convert this knowledge into spontaneous hand and arm movement; in other words, such movements could have been used as instrumental gestures to reflect and imitate certain musical features (Leman, 2007). Both high frequency Sub-band Flux and Percussiveness were found to be related to hand movements. Such movement characteristics could reflect the way these sound qualities are produced. For instance, participants could have imagined playing the drums during moving to the music (cf., "Air Instruments," Godøy et al., 2006). A follow-up study including an appropriate background questionnaire (covering use and experience of musical instruments) should be conducted to investigate this relationship in more detail.

Furthermore, the link between head movement and rhythm-related musical features might be based on the tendency to use head movements to spontaneously follow the rhythm of the music. This could be seen as another example of the concept of embodied attuning (Leman, 2007). An additional interpretation of such head movements could be related to the results by Phillips-Silver and Trainor (2008), who found head movement to bias meter perception, as well as by Trainor et al. (2009), who discovered the vestibular system to play a primal role in rhythm perception: movements of the head could therefore support rhythm perception and understanding.

Movements are used not only in musical contexts, but also in speech. Hand and head gestures are widely utilized to accompany speech and convey additional information. Hadar (1989) noted that such gestures are used, for example, to clarify or emphasize messages. Consequently, our participants could have used their hand and head movements in a similar fashion, i.e., to emphasize and elaborate the musical (e.g., rhythmic) structure or certain musical events.

We aimed at designing an ecological study, as far as this is possible with an optical motion capture system and a laboratory situation. To this end, we chose real music stimuli (pre-existing pop songs), accepting the downside that they were less controlled, very diverse, and more difficult to analyze, as computational analysis of complex musical stimuli is not yet as sufficiently developed as for simpler, i.e., monophonic, stimuli. Furthermore, the stimuli might contain relevant musical features that we missed to extract. However, this approach made it possible to present the participants with music that they were potentially familiar with, and that is played in dance clubs. One could assume that this kind of music would make them move more and in a more natural fashion than more artificial stimuli.

The movement characteristics chosen in this study cover only a small part of the possible movement characteristics and types. There are certainly stereotypical genre- and style-dependent movements that are rather culturally developed than intrinsic to the music. Examples of these kinds of movements would be head banging in rock music or swaying hips to Latin music. To get more insight into such movement patterns, gesture-based computational analysis of the movement could be performed in the future.

As Desmond (1993) noted, dance is related to cultural identity. Since the musical styles used in this study can all be characterized as popular music in the cultural region in which the data collection took place, different movement characteristics might be found with participants that are not as familiar with such musical styles as our participants were. Comparative studies would give insights into cultural differences and commonalities of music-induced movement characteristics.

The present study revealed relationships between rhythm- and timbre-related musical features and movement characteristics. In the future, we will investigate periodicity and synchronization of music-induced movement, as well as further relationships of movement characteristics and musical features, such as tonality features, as they play an important role for the perception of musical emotions. Additionally, the results obtained in this study regarding tempo call for further investigation. A new set of stimuli could be created to control for tempo and related styles/genres to investigate the relationship of tempo, musical style, and resulting movement features.

ACKNOWLEDGMENTS

We thank Mikko Leimu for help and technical assistance during the motion capture data collection. This study was funded by the Academy of Finland (projects 118616, 125710, 136358).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Auditory_Cognitive_Neuroscience/10.3389/fpsyg.2013.00183/abstract

REFERENCES

Alluri, V., and Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Percept.* 27, 223–242.

Arom, S. (1991). *African Polyphony and Polyrhythm: Musical Structure and Methodology*. Cambridge: Cambridge University Press.

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., et al. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex* 45, 62–71.

Brown, S., Merker, B., and Wallin, N. L. (2000). "An introduction to evolutionary musicology," in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown (Cambridge, MA: MIT Press), 3–24.

Camurri, A., Lagerlöf, I., and Volpe, G. (2003). Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *Int. J. Hum. Comput. Stud.* 59, 213–225.

Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., and Volpe, G. (2004). "Multimodal analysis of expressive gesture in music and dance performances," in *Gesture-Based Communication in Human-Computer Interaction. Lecture Notes in Computer Science, 2915*, eds A. Camurri and G. Volpe (Berlin: Springer), 20–39.

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Ann. N. Y. Acad. Sci.* 1169, 15–34.

Collyer, C. E., Broadbent, H. A., and Church, R. M. (1992). Categorical time production: evidence for discrete timing in motor control. *Percept. Psychophys.* 51, 134–144.

Cross, I. (2001). Music, cognition, culture, and evolution. *Ann. N. Y. Acad. Sci.* 930, 28–42.

Desmond, J. C. (1993). Embodying difference: issues in dance and cultural studies. *Cult. Crit.* 26, 33–63.

Eerola, T., Luck, G., and Toiviainen, P. (2006). "An investigation of preschoolers' corporeal synchronization with music," in *Proceedings of the 9th International Conference on Music Perception and Cognition*, eds M. Baroni, A. R. Addessi, R. Caterina, and M. Costa (Bologna: University of Bologna), 472–476.

Fraisse, P. (1982). "Rhythm and tempo," in *The Psychology of Music*, ed. D. Deutsch (New York, NY: Academic Press), 149–180.

Godøy, R. I., Haga, E., and Jensenius, A. R. (2006). "Playing 'Air instruments': mimicry of sound-producing gestures by novices and experts," in *Gesture in Human-Computer Interaction and Simulation, Lecture Notes in Computer Science, 3881*, eds S. Gibet, N. Courty, and J.-F. Kamp (Berlin: Springer), 256–267.

Grahn, J. A., and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* 19, 893–906.

Grahn, J. A., and Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 29, 7540–7548.

Hadar, U. (1989). Two types of gesture and their role in speech production. *J. Lang. Soc. Psychol.* 8, 221–228.

Janata, P., Tomic, S. T., and Haberman, J. M. (2011). Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol. Gen.* 141, 54–75.

Jensenius, A. R. (2006). "Using motiongrams in the study of musical gestures," in *Proceedings of the International Computer Music Conference* (New Orleans, LA: Tulane University), 499–502.

Keller, P., and Rieger, M. (2009). Special issue – musical movement and synchronization. *Music Percept.* 26, 397–400.

Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books.

Lartillot, O., Eerola, T., Toiviainen, P., and Fornari, J. (2008). "Multi-feature modeling of pulse clarity: design, validation, and optimization," in *Proceedings of the 9th International Conference on Music Information Retrieval*, eds J. P. Bello, E. Chew, and D. Turnbull (Philadelphia, PA: Drexel University), 521–526.

Lartillot, O., and Toiviainen, P. (2007). "A matlab toolbox for musical feature extraction from audio," in *Proceedings of the 10th International Conference on Digital Audio Effects* (Bordeaux: University of Bordeaux), 1–8.

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA: MIT Press.

Leman, M., and Godøy, R. I. (2010). "Why Study Musical Gesture?" in *Musical Gestures. Sound, Movement, and Meaning*, eds R. I. Godøy and M. Leman (New York, NY: Routledge), 3–11.

Leman, M., and Naveda, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in Samba and Charleston. *Music Percept.* 28, 71–92.

Lesaffre, M., De Voogdt, L., Leman, M., De Baets, B., De Meyer, H., and Martens, J.-P. (2008). How potential users of music search and retrieval systems describe the semantic quality of music. *J. Am. Soc. Inf. Sci. Technol.* 59, 695–707.

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., and Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *J. Res. Pers.* 44, 714–720.

MacDougall, H. G., and Moore, S. T. (2005). Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *J. Appl. Physiol.* 99, 1164–1173.

Madison, G., Gouyon, F., Ullén, F., and Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: first correlations with low-level audio descriptors across music genres. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1578–1594.

Moelants, D. (2002). "Preferred tempo reconsidered," in *Proceedings of the 7th International Conference on Music Perception and Cognition*, eds C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (Adelaide: Causal Productions), 580–583.

Murray, M. P., Drought, A. B., and Kory, R. C. (1964). Walking patterns of normal men. *J. Bone Joint Surg.* 46, 335–360.

Naveda, L., and Leman, M. (2010). The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA). *Music Percept.* 28, 93–112.

Nettl, B. (2000). "An ethnomusicologist contemplates universals in musical sound and musical culture," in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown (Cambridge, MA: MIT Press), 463–472.

Pampalk, E., Rauber, A., and Merkl, D. (2002). "Content-based organization and visualization of music archives," in *Proceedings of the 10th ACM International Conference on Multimedia, Juan-les-Pins* (New York, NY: ACM Press), 570–579.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Percept.* 11, 409–464.

Phillips-Silver, J., and Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain Cogn.* 67, 94–102.

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* 12, 969–992.

Savitzky, A., and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639.

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428.

Stevens, C. J., Schubert, E., Wang, S., Kroos, C., and Halovic, S. (2009). Moving with and without music: scaling and lapsing in time in the performance of contemporary dance. *Music Percept.* 26, 451–464.

Styns, F., van Noorden, L., Moelants, D., and Leman, M. (2007). Walking on music. *Hum. Mov. Sci.* 26, 769–785.

Toiviainen, P., and Burger, B. (2011). *MoCap Toolbox Manual*. Jyväskylä: University of Jyväskylä.

Toiviainen, P., Luck, G., and Thompson, M. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Percept.* 28, 59–70.

Trainor, L. J., Gao, X., Lei, J.-J., Lehtovaara, K., and Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex* 45, 35–43.

Van Dyck, E., Moelants, D., Demey, M., Coussement, P., Deweppe, A., and Leman, M. (2010). "The impact of the bass drum on body movement in spontaneous dance," in *Proceedings of the 11th International Conference in Music Perception and Cognition*, eds S. M. Demorest, S. J. Morrison, and P. S. Campbell (Seattle, WA: University of Washington), 429–434.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Zentner, M., and Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5768–5773.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 November 2012; accepted: 26 March 2013; published online: 12 April 2013.

Citation: Burger B, Thompson MR, Luck G, Saarikallio S and Toiviainen P (2013) Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. Front. Psychol. ***4****:183. doi: 10.3389/fpsyg.2013.00183*

This article was submitted to Frontiers in Auditory Cognitive Neuroscience, a specialty of Frontiers in Psychology.

APPENDIX

SONG LIST

1. Alice Deejay: better Off Alone (Who Needs Guitars Anyway?) 2:40–2:54 (loop).
2. Andre Visior: speed Up 1:15–1:45.
3. Antibalas: who is this America Dem Speak of Today? (Who Is This America?) 1:00–1:29.
4. Arturo Sandoval: a Mis Abuelos (Danzon) 1:53–2:21.
5. Baden Powell: deixa (Personalidade) 1:11–1:41.
6. Brad Mehldau: wave/Mother Nature's Son (Largo) 0:00–0:29.
7. Clifford Brown & Max Roach: the Blues walk (Verve Jazz Masters, Vol. 44: Clifford Brown & Max Roach) 2:01–2:31.
8. Conjunto Imagen: medley-Esencia de Guaguanco/Sonero (Ayer, Hoy y Manana) 2:18–2:48.
9. Dave Hillyard & The Rocksteady 7: Hillyard Street (Playtime) 0:15–0:45.
10. Dave Weckl: mercy, Mercy, Mercy (Burning for Buddy) 0:10–0:40.
11. Dave Weckl: tower of Inspiration (Master Plan) 0:00–0:30.
12. DJ Shadow: napalm Brain/Scatter Brain (Endtroducing . . .) 3:29–3:58.
13. Gangster Politics: gangster Politics (Guns & Chicks) 1:00–1:29.
14. Gigi D'Agostino: blablabla (L'Amour Toujours) 0:00–0:30.
15. Herbie Hancock: watermelon man (Cantaloupe Island) 0:00–0.30.
16. Horace Silver: the Natives Are Restless 0:00–0:29.
17. In Flames: scream (Come Clarity) 0:00–0:30.
18. Jean Roch: can You Feel it (Club Sounds Vol. 35) 0:33–1:01.
19. Johanna Kurkela: Hetki hiljaa (Hetki hiljaa) 3:22–3:52.
20. Juana Molina: tres cosas (Tres Cosas) 0:00–0:30.
21. Kings of Leon: closer (Only by the Night) 3:17–3:47.
22. Lenny Kravitz: live (5) 3:02–3:30.
23. Martha & The Vandellas: heat Wave (Heat Wave) 1:40–2:10.
24. Maynard Ferguson: fireshaker (Live From San Francisco) 0:00–0:28.
25. MIA: 20 Dollar (Kala) 0:17–0:45.
26. Nick Beat: techno Disco 2:26–2:56.
27. Panjabi MC: mundian To Bach Ke (Legalized) 0:47–1:06 (loop).
28. Patrick Watson: Beijing (Wooden Arms) 2:30–2:59.
29. The Rippingtons: weekend in Monaco (Weekend in Monaco) 1:13–1:42.
30. Yuri Buenaventura: salsa (Salsa Movie Soundtrack) 2:17–2:45.

II

RELATIONSHIPS BETWEEN SPECTRAL FLUX, PERCEIVED RHYTHMIC STRENGTH, AND THE PROPENSITY TO MOVE

by

Birgitta Burger, Riikka Ahokas, Aaro Keipi & Petri Toiviainen, 2013

Proceedings of the Sound and Music Computing Conference (pp. 179–184), KTH
Royal Institute of Technology, Stockholm, Sweden

RELATIONSHIPS BETWEEN SPECTRAL FLUX, PERCEIVED RHYTHMIC STRENGTH, AND THE PROPENSITY TO MOVE

Birgitta Burger **Riikka Ahokas** **Aaro Keipi** **Petri Toiviainen**
Department of Music, University of Jyväskylä, Jyväskylä, Finland
{birgitta.burger@, j.riikka.ahokas@student., aaro.j.keipi@student., petri.toiviainen@}jyu.fi

ABSTRACT

The tendency to move to music seems to be built into human nature. Previous studies have shown a relationship between movement and the degree of spectral flux in music, particularly in the lower sub-bands. In this study, listeners' perceptions of a range of frequency-restricted musical stimuli were investigated in order to find relationships between perceived musical aspects (rhythm, melody, and fluctuation) and the spectral flux in three different frequency bands. Additionally, the relationship between the perception of features in specific frequency bands and participants' desire to move was studied. Participants were presented with clips of frequency-restricted musical stimuli and answered four questions related to musical features. Both perceived strength of the rhythm and the propensity to move were found to correlate highly with low-frequency spectral flux. Additionally, a lower but still significant correlation was found between these perceived musical features and high-frequency spectral flux. This suggests that the spectral flux of both low and high frequency ranges can be utilized as a measure of perceived rhythm in music, and that the degree of spectral flux and the perceived rhythmic strength in high and low frequency bands are at least partly responsible for the extent to which listeners consciously desire to move when listening to music.

1. INTRODUCTION

When listening to rhythmic music we tend to move our bodies with it. Movements induced by music might be subconscious, with almost indistinguishable trappings, or deliberate, strong and intentional. The proclivity to move with music seems to be built into human nature, which is described in the literature as groove (see, e.g., [1]-[3]). These studies propose that the functional role of rhythmic music and the construct of groove are related to the evolution of entrainment and social behavior and state that synchronizing is the simplest form of entrainment from a psychological point of view. "Synchronization" is also the concept that Leman [4] suggests as the most fundamental component in bodily engagement with music. He proposed three concepts of (co-existing) corporeal articulations – "Synchronization", "Embodied Attuning", and "Empathy" – that differ in the degree of musical involvement and in the kind of action-perception couplings involved. "Synchronization" forms the fundamental component, as synchronizing to a beat is easy and spontaneous. As the first step in engaging with the music, movements could be used for imitation and prediction of beat-related features in the music. The second component, "Embodied Attuning", concerns the linkage of body movement to musical features more complex than the basic beat, such as melody, harmony, rhythm, tonality, or timbre. Following this idea, movement could be used to reflect, imitate, and navigate within the musical structure. Finally, "Empathy" is seen as the component that links musical features to expressivity and emotions.

Thus, music-induced movements seem to be associated with rhythmic features of music, such as periodic and regular patterns of beats and pulses. In basic western popular music settings the rhythm section (the drummer and bass player) are responsible for providing the rhythm. Van Dyck et al. [5] studied the effect of the dynamics of the bass drum on dancers in order to find if the bass drum is a feature that dominates music-induced movements. The authors concluded that the dynamic changes of the bass drum have an underlying effect on the intensity of movement while dancing. Burger and colleagues (see [6] and [7]) conducted a motion capture study, in which participants were asked to move to various pop music stimuli. They performed computational feature extraction on both the movement and the music data and found several relationships between movement characteristics and rhythm- and timbre-related musical features. Their results indicate that clear pulses in the music encouraged participants to move their whole body with low spatial complexity, while spectral flux in the low and high frequency ranges was more distinctly related to certain body parts. With an increasing amount of flux in the low and high frequencies, the authors discovered an increase in head and hand movement as well as an increase in temporally regular movement synchronized to different metrical levels, whereas more complex, irregular rhythmic structures resulted in temporally less regular movement. The authors concluded that spectral flux was related to the perception of rhythm of the music – flux of the low frequencies being associated with kick drum and bass guitar and high frequency flux being influenced by hi-hat and cymbal sounds – and therefore considered important for inducing movement. However, the perceptual dimension of spectral flux of restricted frequency bands has only been studied so far in connection to polyphonic timbre (see [8] and [9]) and music information retrieval related applications, such as automatic classification (see [10] and [11], both slightly differing in their technical implementation), but not strictly in relation to rhythm perception.

The spectral flux of restricted frequency bands, or sub-band flux, is a computational measure indicating the extent to which the spectrum changes over time. When computing this feature (see [8]), the stimulus is divided into 10 frequency bands, each band containing one octave in the range of 0 to 22050 Hz. After that the sub-band flux is calculated for each of these ten bands by taking the average of the Euclidean distances of the spectra for each pair of two consecutive frames of the signal (for more information about the derivation of the feature, see [8]). Two spectrograms of sub-band no. 2 (50-100 Hz) are displayed in Figure 1 to show the difference between high and low amounts of sub-band flux.

The purpose of this study was to investigate listener's perception of a range of frequency-restricted musical stimuli. The original versions of the stimuli have already been used in the movement studies cited previously (see [6] and [7]), however the present study included both the original version and three different frequency restricted versions of the original stimuli (low, mid, and high frequencies). We aimed to find relationships between specific musical aspects (such as rhythm, melody, and fluctuation) and the spectral flux in the different frequency bands. Additionally, we were interested in the relationship between the perception of musical features in specific frequency bands and participants' desire to move. We hypothesized that the perceived strength of the rhythm is positively correlated with the spectral flux, especially for the stimuli restricted to low frequencies (sub-band 2), so participants would perceive the low-frequency sub-band flux as being related to rhythm. Furthermore, we assumed positive correlations between the desire to move and the spectral flux for sub-band 2 and 9, as the spectral flux in these bands was found to be related to several characteristics of human movement (see [6] and [7]).

2. METHOD

2.1 Participants

A total of 38 participants (26 females; average age: 26.42, SD of age: 4.95) took part in the experiment. Participants were international students from the University of Jyväskylä, Finland. Participants were compensated with a movie ticket.

2.2 Stimuli

The stimuli consisted of 30-second segments from 30 different popular songs from various genres including Techno, Pop, Rock, Latin, Funk, and Jazz (the same musical stimuli as in [6] and [7] – a list of stimuli is included in these publications). They were all non-vocal and in 4/4 time, but differed in their rhythmic complexity and pulse clarity. In order to present participants with frequency-restricted stimuli, each clip was modified using MATLAB MIRToolbox 1.4 (see [12]): The clip was first divided into ten frequency bands, each band containing one octave in the range of 0 to 22050 Hz. Then the sub-bands of interest (sub-band 2: 50-100 Hz, sub-band 6: 800-1600 Hz, and sub-band 9: 6400-12800 Hz) were extracted and saved as .wav files.

2.3 Apparatus

To gather the perceptual ratings, a special patch was created in Max/MSP 5, a graphical programming environment, running on Max OS X. The setup enabled the participants to repeat excerpts as often as they wished and to move forward at their own speed. The stimuli were played back through active studio monitors (Genelec 8030A). The participants could themselves adjust the volume to a preferred level.

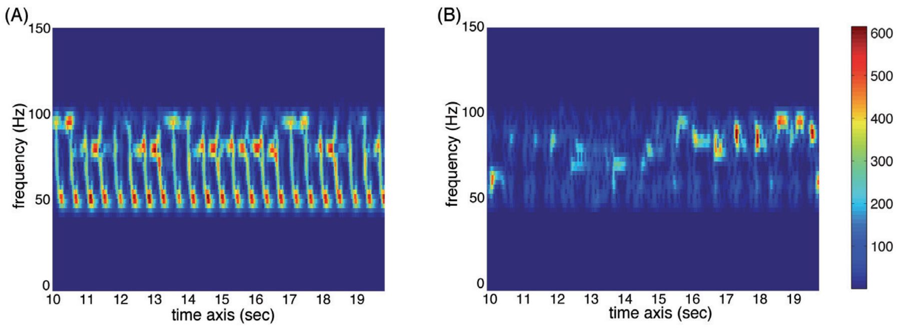


Figure 1. Spectrograms of sub-band no. 2 (50-100 Hz) (sec. 10 to 20 of two stimuli used in the study presented). (A) High amount of temporal change (red represents high energy at the respective time and frequency, whereas blue represents low energy; see color bar) resulting in high value for Sub-Band Flux. (B) Low amount of temporal change resulting in low Sub-Band Flux.

2.4 Procedure

The experiment was divided into four sections: one section containing the stimuli restricted to sub-band 2, a second section containing the stimuli restricted to sub-band 6, a third section containing the stimuli restricted to sub-band 9, and a fourth section containing the original stimuli. Each section was presented separately; first, the three sub-band restricted sections in random order, followed by the section containing the original clips. This section was always presented last to avoid a biased rating due to knowing the whole stimuli. The stimuli were also randomized within each section. Participants accomplished the experiment individually. They were asked to answer four questions, rating each on a seven-step scale (from "not at all" to "very much"):

1. *How prominent is the rhythm?*
2. *How prominent is the melody?*
3. *How much fluctuation is there in the music (How much is "going on" in the music)?*
4. *How strongly does it make you want to move?*

Preceding the experiment, there was a practice section with one example to allow participants to become familiar with the interface and the questions. The average duration of the experiment was 75 minutes.

2.5 Spectral flux extraction

For each of the four versions of each stimulus (three frequency-restricted and the original stimuli), the spectral flux was computed (using MATLAB MIRToolbox 1.4 [12]) by calculating the Euclidean distances of the spectra for each pair of consecutive frames of the signal, using a frame length of 25 ms and an overlap of 50% between successive frames. Subsequently, we averaged across the resulting time-series of flux values to receive one value for each of the four versions of the stimuli.

3. RESULTS

The first step of the analysis comprised checking the consistency of the ratings of the participants by calculating intraclass correlations (cf., [13]) for each question and stimulus type separately. The results are presented in Table 1.

	SB 2	SB 6	SB 9	Orig.
Question 1: ***rhythm?***	.95 ***	.97 ***	.96 ***	.96 ***
Question 2: ***melody?***	.94 ***	.95 ***	.93 ***	.95 ***
Question 3: ***fluctuation?***	.87 ***	.90 ***	.88 ***	.93 ***
Question 4: ***movement?***	.93 ***	.93 ***	.91 ***	.94 ***

*** $p < .001$

Table 1. Intraclass correlations for each question and stimulus type.

As these correlation coefficients indicate sufficiently high inter-participant consistency, we averaged the ratings across participants to receive one value per stimulus. Such high intraclass correlations, especially for question 1 ("*How prominent was the rhythm?*"), also suggest that the concepts of rhythm and melody were understood in a coherent way by the participants (despite findings related to cultural dependencies of rhythm perception, see [14] and [15]). Worth noting is that the correlation coefficient for question 3 ("*How much fluctuation is there in the music*") showed the lowest value for each stimulus type.

To investigate the relationship between the spectral flux data (calculated for each of the four versions of the stimuli as described in section 2.5) and the perceptual evaluations of the stimuli, we correlated the rating scores of the four questions for the music clips per stimulus type (averaged across participants) with the respective flux data. The results of the correlations are displayed in Table 2. Correlations with significance values less than p < .01 are indicated with asterisks.

	SB 2	SB 6	SB 9	Orig.
Question 1: ***rhythm?***	.79 ***	.20	.47 **	.54 **
Question 2: ***melody?***	-.01	-.28	.27	.03
Question 3: ***fluctuation?***	-.15	.01	.37	.11
Question 4: ***movement?***	.65 ***	.20	.52 **	.57 ***

** $p < .01$, *** $p < .001$

Table 2. Correlations between spectral flux and ratings on questions 1-4 for each stimulus type.

The strongest correlation ($r(30) = .79$, $p < .001$) for question 1 ("*How prominent is the rhythm?*") was found for the sub-band 2 stimuli. Question 4 ("*How strongly does it make you want to move?*") was also relatively highly correlated ($r(30) = .65$, $p < .001$) to these stimuli. Not quite as strong – though still significant – were the correlations between the same two questions and sub-band 9 flux ($r(30) = .47$, $p < .01$, for question 1, and $r(30) = .52$, $p < .01$, for question 4, respectively) and between these two questions and the flux of the original stimuli ($r(30) = .54$, $p < .01$, for question 1, and $r(30) = .57$, $p < .001$, for question 4, respectively). As all correlations were positive, these results suggest that participants rated stimuli with an increasing amount of flux in both low and high frequency ranges and overall flux with higher prominence of rhythm and with higher desire to move to the presented stimuli.

Meanwhile, the values for question 2 ("*How prominent is the melody?*") showed non-significant correlations with flux data of all stimulus types, suggesting that there is no relationship between the perceived melody prominence and the amount of (sub-band) flux in the stimuli.

Interestingly, the values for question 3: "*How much fluctuation is there in the music (How much is "going on" in the music)?*" showed no significant correlation to the (sub-band) flux data. This suggests that there was no relation between the fluctuation participants perceived in the stimuli and the computationally extracted flux in the sub-bands and the original clips.

Subsequently, we performed correlations between the ratings, segregated by each stimulus type. The results are shown in Figure 2.

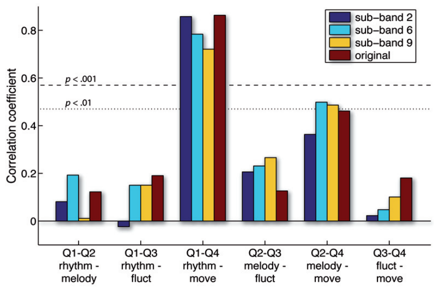


Figure 2. Correlations between the four ratings for each stimulus type separately.

The correlations between the questions show a similar pattern across the different stimulus types: for all stimulus types, question 1 ("*How prominent is the rhythm?*") correlated positively with question 4 ("*How strongly does it make you want to move?*") (sub-band 2: $r(30) = .86$, $p < .001$; sub-band 6: $r(30) = .78$, $p < .001$; sub-band 9: $r(30) = .72$, $p < .001$; original: $r(30) = .86$, $p < .001$) suggesting that the perception of a prominent rhythm was related to participants' eagerness to move to such stimuli regardless of its frequency range. For sub-band 6 and 9, additionally, question 2 ("*How prominent is the melody*") correlated moderately high with question 4 ("*How strongly does it make you want to move?*") (sub-band 6: $r(30) = .50$, $p < .01$; sub-band 9: $r(30) = .49$, $p < .01$). Thus, in the stimuli restricted to both the mid and high frequencies, the melody also appears to contribute to the willingness to move to such stimuli. The remaining correlations were non-significant.

4. DISCUSSION

We conducted an experiment to investigate participants' perceptions of rhythm, melody, fluctuation, and the desire to move in full-frequency and frequency-restricted musical stimuli. The participants' answers were consistent with our hypothesis: for stimuli restricted to the low frequency band (sub-band 2), stimuli having a higher amount of sub-band flux were perceived as being stronger related to the rhythm of the music than stimuli with a lower amount of sub-band flux, as suggested by the high correlations of flux data in sub-band 2 with the question "*How prominent is the rhythm?*". These correlations are likely due to the frequency range of specific rhythmic instruments in this sub-band, such as kick drum and low bass notes. Additionally, a greater amount of low frequency spectral flux would induce the desire of movement in participants, as was suggested by the positive correlation with the question "*How strongly does it make you want to move?*".

In addition, these two questions were also highly correlated to the spectral flux in sub-band 9. These correlations are likely due to the frequency range of specific rhythmic instruments in sub-band 9, such as hi-hat and some snare drum partials. These findings corroborate the results reported in [6] and [7], which showed higher amounts of specific bodily movements related to the amount of spectral flux in sub-bands 2 and 9. Spectral flux in sub-bands 2 and 9 may therefore be potentially more effective than spectral flux in other sub-bands in encouraging people to move to music.

The low correlations for question 3 ("*How much fluctuation is there in the music (How much is "going on" in the music)?*") with all sub-bands, however, showed that participants could not perceive the amount of fluctuation in the stimuli. There are two possible explanations for this: 1) the participants were simply unable to accurately and consistently hear the amount of fluctuation in the individual sub-bands; or 2) participants did not understand the concept of fluctuation in this context. The intraclass correlation results for this question (see Table 2) showed that the answers for this question were less consistent compared to the other questions, so the latter explanation seems likely. Interestingly, none of the participants asked about the term during the data collection, thus it could be assumed that the participants understood the term fluctuation, but that this notion differed across participants. Future studies on a similar topic should make certain that the concept of fluctuation is clearly understood by participants.

Rhythmic content has been found to be strongly related to movement (see [5]-[7]). The high correlations between the questions "*How prominent is the rhythm?*" and "*How strongly does it make you want to move?*" for all the four stimulus types suggests that the perception of a prominent rhythm is related to participants' eagerness to move to stimuli regardless of its frequency range. Future studies could analyze the actual differences of participants' movement for specific frequency ranges, for instance in a motion capture setting.

The relationship found between perceived rhythm and desire to move (see previous paragraph) could also serve as support for Leman's theory of corporeal articulations [4]. Stimuli that participants rated to contain a strong rhythm, were also rated high on desire to move, suggesting a connection between both. That could be seen as being in line with the concepts of "Synchronization" and "Embodied Attuning", in which beat/musical features, such as the rhythm, are proposed to induce body movement.

There was a weaker – but still apparent – correlation between the two questions "*How prominent is the melody*" and "*How much does it make you want to move?*", which points to a relationship between melodic strength and music-induced desire to move. The effect of melodic content on movement could also be the subject of future studies.

It could be argued that some of the participants' answers (especially to question 4: "*How much does it make* *you want to move?*") may have been skewed by prior experience with the particular stimuli since the stimuli used were clips from western popular music. However, it could be assumed that the wide range of backgrounds of participants, as well as the modification of the sub-band clips to within a certain frequency range (which often made the source music difficult to distinguish), helped to minimize the possible effects of familiarity on the ratings. Nevertheless, it might be valuable for future data acquisitions to include collecting both familiarity and preference ratings of the stimuli. This would give more insight into relationships between music characteristics and movement propensity, as it could be assumed that participants still rate certain stimuli high on "desire to move" although they do not like them.

We excluded extreme genres in our stimuli selection – such as death metal for example – as such music might be more prone to familiarity and preference than other popular music. Although musical styles such as death metal could contain high spectral flux and be considered rhythmic, not everybody would probably feel the urge to dance to such music. Thus, if such extreme genres were considered in the stimuli selection, relationships between spectral flux and movement propensity might be less linear than presented in this paper.

The presented analysis utilized one value – the average of the flux time-series – as measure for the spectral flux. As such, this could be regarded as an over-generalization, since taking the mean disregards information about the temporal regularity that the flux series should exhibit in order to induce movement. In general, a random use of, for instance, the kick drum would also result in a high amount of low-frequency flux, but it would fail to evoke a sensation of rhythm or movement in the listener. However, our stimuli were throughout the whole stimulus duration all metrically regular, had a sensation of pulse, and were steady in most of the musical characteristics. Thus, we believe that temporal averaging of the flux time-series was a suitable way to receive a relevant measure of spectral flux for each stimulus.

In conclusion, the results of this study show that for stimuli being restricted to low frequencies and, to a lesser extent, for stimuli being restricted to high frequencies, a high amount of spectral flux was perceived as having a more prominent rhythm. This suggests that the sub-band flux of both low and high frequency ranges can be utilized as a possible measure of perceived rhythm in music. Furthermore, the significant correlation between the answers to the questions "*How prominent is the rhythm*" and "*How much does it make you want to move*" to the spectral flux in sub-bands 2 and 9 point to an important role of the spectral content in these sub-bands; in essence, it suggests that the degree of flux and the perceived rhythmic strength in sub-bands 2 and 9 are at least partly responsible for the extent to which listeners consciously desire to move when listening to music. This is consistent with previous research (see [6] and [7]) that identified spectral flux in these particular sub-bands as correlating with various characteristics of bodily movements.

Acknowledgments

This study was supported by the Academy of Finland (project 118616).

5. REFERENCES

[1] G. Madison, "Experiencing groove induced by music: Consistency and phenomenology," in Music Perception 24(2), 2006, pp. 201-208.

[2] G. Madison, F. Gouyon, F. Ullén, and K. Hörnström, "Modeling the tendency for music to induce movement in humans: first correlations with low-level audio descriptors across music genres," in Journal of Experimental Psychology: Human Perception and Performance 37(5), 2011, pp. 1578-1594.

[3] P. Janata, S. T. Tomic, and J. M. Haberman, "Sensorimotor coupling in music and the psychology of the groove", in Journal of Experimental Psychology: General 141(1), 2011, pp. 54-75.

[4] M. Leman, Embodied Music Cognition and Mediation Technology. Cambridge, MA/London, UK: MIT Press, 2007.

[5] E. van Dyck, D. Moelants, M. Demey, P. Coussement, A. Deweppe, and M. Leman, "The impact of the bass drum on body movement in spontaneous dance," in Proceedings of the 11th International Conference in Music Perception and Cognition, Seattle, WA, 2010, pp. 429-434.

[6] B. Burger, M. R. Thompson, G. Luck, S. Saarikallio, and P. Toiviainen, "Music moves us: Beat-related musical features influence regularity of music-induced movement," in Proceedings of the 12th International Conference in Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences for Music, Thessaloniki, Greece, 2012, pp. 183-187.

[7] B. Burger, M. R. Thompson, S. Saarikallio, G. Luck, and P. Toiviainen, "Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement," in Frontiers in Psychology, 4:183, 2013.

[8] V. Alluri, and P. Toiviainen, "Exploring perceptual and acoustical correlates of polyphonic timbre," in Music Perception 27(3), 2010, pp. 223-242.

[9] V. Alluri, and P. Toiviainen, "Effects of enculturation of the semantic and acoustic correlates of polyphonic timbre," in Music Perception 29(3), 2012, pp. 297-310.

[10] D.-N. Jiang, L. Lu, H.-J. Zhang, J.-H. Tao, and L.-H. Cai, "Music type classification by spectral contrast feature," in *Proceedings of the IEEE International*

Conference on Multimedia and Expo, 2002, pp. 113-116.

[11] R. Cai, L. Lu, A. Hanjalic, H.-J. Zhang, L.-H. Cai, "A flexible framework for key audio effects detection and auditory context inference," in IEEE Transactions on Audio, Speech and Language Processing 14(3), 2006, pp. 1026-1039.

[12] O. Lartillot, and P. Toiviainen, "A Matlab toolbox for musical feature extraction from audio," in Proceedings of the 10th International Conference on Digital Audio Effects, Bordeaux, France, 2007, 1-8.

[13] P. E. Shrout, and J.L. Fleiss, "Intraclass correlations: uses in assessing rater reliability," in Psychological Bulletin 86(2), 1979, pp. 420-428.

[14] A. D. Patel, J. R. Iversen, and K. Ohgushi, "Cultural differences in rhythm perception: what is the influence of native language?" In Proceedings of the 8th International Conference on Music Perception and Cognition, Evanston, IL, 2004, 88-89.

[15] G. Soley and E. E. Hannon, "Infants Prefer the Musical Meter of Their Own Culture: A Cross-Cultural Comparison," in Developmental Psychology 46(1), 2004, 286-292.

III

RELATIONSHIPS BETWEEN PERCEIVED EMOTIONS IN MUSIC AND MUSIC-INDUCED MOVEMENT

by

Birgitta Burger, Suvi Saarikallio, Geoff Luck, Marc R. Thompson & Petri Toiviainen, 2013

Music Perception vol. 30(5), 519-535

Relationships Between Perceived Emotions in Music and Music-induced Movement

Birgitta Burger, Suvi Saarikallio, Geoff Luck,
Marc R. Thompson, & Petri Toiviainen
University of Jyväskylä, Jyväskylä, Finland

Listening to music makes us move in various ways. Several factors can affect the characteristics of these movements, including individual factors and musical features. Additionally, music-induced movement may also be shaped by the emotional content of the music, since emotions are an important element of musical expression. This study investigates possible relationships between emotional characteristics of music and music-induced, quasi-spontaneous movement. We recorded music-induced movement of 60 individuals, and computationally extracted features from the movement data. Additionally, the emotional content of the stimuli was assessed in a perceptual experiment. A subsequent correlational analysis revealed characteristic movement features for each emotion, suggesting that the body reflects emotional qualities of music. The results show similarities to movements of professional musicians and dancers, and to emotion-specific nonverbal behavior in general, and could furthermore be linked to notions of embodied music cognition. The valence and arousal ratings were subsequently projected onto polar coordinates to further investigate connections between the emotions of Russell's (1980) circumplex models and the movement features

Received: May 4, 2012, accepted January 18, 2013.

Key words: music, music-induced movement, motion capture, emotion, circumplex model of affect

Listening to music makes us move in various ways (Leman, 2007; Leman & Godøy, 2010). Several factors can affect the characteristics of these movements: Individual factors such as personality and preference have been found to influence movability to music (Luck, Saarikallio, Burger, Thompson, & Toiviainen, 2010), as have music-intrinsic features such as beat strength and pulse clarity (Burger, Thompson, Saarikallio, Luck, & Toiviainen, 2010; Van Dyck et al., 2010). Besides such factors, music-induced movement may be shaped by the emotional content of the music. Emotions are an essential component of musical expression (e.g., Gabrielsson & Lindström, 2001) that can have a strong influence, for instance, on the listener's mood (e.g., Saarikallio, 2011). However, only a few studies have addressed the connection between emotions and music-induced movement.

Research on emotion-specific nonverbal behavior has nonetheless shown that it is possible to successfully express and communicate emotional states to observers through body movements. Already in 1872, Darwin assigned certain body movements and postures quite specifically to emotional states; joyful movements, for example, were described as jumping, stamping, body thrown backwards and shaking, and upright torso and head, whereas anger was characterized by trembling body, shaking fist, erected head, and expanded chest. Sad movements were described as passive, motionless, and a downward directed head. Despite these finding, the topic was rather neglected for a while–probably due to lack of appropriate technologies–and research focused more on facial expression (e.g., Ekman, 1982) following the assumption that gross body movements would only communicate intensity of emotions, but not qualitative characteristics (Ekman & Friesen, 1974).

Nevertheless, several recent studies have shown that emotions can be expressed by, and successfully recognized from, body postures or movements. From the early seventies onwards, Clynes (1980, 1992) investigated in the context of human-machine interaction how communication can be made more human-like and more efficient. For this purpose he developed a repertoire of so-called dynamic transient forms, or sentic forms that are specific for each emotion.

Wallbott (1998) conducted a study in which he used a scenario-based approach, with professional actors performing certain emotions. Besides a recognition test, he also analyzed movement features characteristic and distinctive for each emotion category; for instance, sadness was expressed with a collapsed body, whereas joy and anger were associated with a more upright torso. Anger was the most active emotion, followed

Music Perception, volume 30, issue 5, pp. 517–533, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2013.30.5.517

by joy and sadness. Lateral arm movement and dynamics/power were mostly related to anger, less for joy, and even less for sadness. Spatially expansive movements were used to express anger and joy, whereas sad movements covered less space.

Dael, Mortillaro, and Scherer (2012) also used actors in a scenario-based approach, and manually labeled their data with a self-developed microcoding system (Body Action and Posture system; BAP) including 49 movement variables, such as "shoulder action" or "arm action towards body." Their analysis revealed that angry movements were characterized by forward movement, joy was communicated with repetitive hand and symmetric arm movement, pleasure was characterized by the head tilted up and asymmetrical arm movement, whereas sad movements were characterized by both arms at rest and down.

A different approach was chosen by De Meijer (1989), who used actors performing several movement characteristics instead of emotions. These movements differed in general features, such as trunk or arm movement, velocity, and spatial direction. In a subsequent step, observers attributed emotional characteristics to the movements. The results suggested that fast, active, open, light, and upward directed movements with raised arms and a stretched trunk were perceived as happy, whereas strong and fast movements with a bowed torso and a high force were perceived as angry. In contrast, sad movements were described as slow, light, downward directed, with arms close to the body.

Coulson (2004) used pictures of static body postures of computer-generated figures, which he manipulated systematically in several parameters. He found that a forward bended and directed body with hands in front of the chest and bent elbows was associated with anger, whereas a forward leaning and downwards bent/crouched body with downwards directed hands was characterized as sad, and a backwards bent body with upraised hands and arms was associated with happiness.

Furthermore, emotions are a central element of music and have been investigated in a large number of music-related studies. According to Krumhansl (2002), people report that their primary motivation for listening to music is its emotional impact. Various rating experiments have shown that listeners are able to perceive emotional content in music. One of the issues that researchers in this field face concerns the choice of emotion categories and models that can cover musical emotions successfully. As Eerola and Vuoskoski (2011) noted, there are basically three different approaches that have been used in studies on music and perceived emotions: discrete (basic) emotions (e.g., Balkwill & Thompson, 1999; Gabrielsson & Juslin, 1996), domain-specific emotion models, such as GEMS (Geneva Emotion Music Scale, see Zentner, Grandjean, & Scherer, 2008), and dimensional models (e.g., Ilie & Thompson, 2006; Schubert, 1999). The discrete model describes emotions as unidimensional, independent from each other, and derived from a limited number of universal and innate basic emotions, such as happiness, anger, sadness, or fear. The advantage of using this approach is the possibility to focus on emotions that are commonly expressed and represented by music, and name them in clear terms. Inspired by this approach, domain-specific discrete scales such as GEMS have been developed to create a repertoire of emotion terms that is highly music-specific. A disadvantage of discrete emotions is, according to Eerola and Vuoskoski (2011), that they can be mixed and confused with each other in discrimination tasks. This issue might be reduced by using a dimensional approach, in which all affective states are represented as a combination of two (usually valence and arousal) or three (valence, arousal, and tension) mutually independent dimensions. A common way to illustrate dimensional models is in a spatial design, such as a circle, with two bipolar dimensions–valence forming the horizontal axis, and arousal the vertical axis (in case of a three-dimensional model, tension would form the third axis resulting in a spherical design). Examples of such approaches are the circumplex models of affect by Russell (1980) and Schlosberg (1954). Eerola and Vuoskoski (2011) compared the different models and showed that, although both dimensional and discrete models performed similarly, however, the discrete model performed worse when it came to ambiguous emotions. Additionally, they found that a two-dimensional model consisting of valence and arousal successfully represented music-related emotions.

Musical emotions are conveyed not only by the music itself, but also through movement. While movements are required, for example, to produce sounds when playing a musical instrument, studies have shown that there are certain additional movements that are not used for the actual sound production, but for conveying emotions and expressivity (e.g., Wanderley, Vines, Middleton, McKay, & Hatch, 2005). Dahl and Friberg (2007) investigated a marimba, a bassoon, and a saxophone player, who each performed a piece of music with different emotional intentions (happy, angry, sad, and fearful). Observers could, when presented with only visual elements of the performance, detect the happy, angry, and sad performances, though failed for the fearful performances. In a second task, they were asked to identify movement cues used to convey the different emotions. Happy intention was found to be communicated by medium regularity and fluency, high speed, and high amount of movement of the whole body, whereas sad intention was conveyed by very regular and fluent movement, low in speed and quantity. Angry intention was communicated with medium regularity, low fluency, high speed, and somewhat high amount of movement. Fearful intention was perceived as being characterized by medium regular and fluid, rather fast, and rather small amounts of movement. Burger and Bresin (2010) developed a small robot displaying different emotions based on the movement cues found by Dahl and Friberg (2007) and used, for example, large, fluent, and circular movements to convey happiness, large, irregular, and jerky movements for anger, and slow, regular, and reduced movements to convey sadness. In a perceptual experiment, all emotions were successfully recognized by the observers.

More direct links between music and emotion-specific movement have been investigated in research on dance, in which movement is the only way to convey expressivity and emotion. Camurri, Lagerlöf, and Volpe (2003) describe a study in which professional dancers were asked to perform the same dance with four different emotional intentions. In a qualitative analysis, they identified characteristic movements for the expression of these emotions; anger was, for instance, characterized by short movements, frequent tempo changes, short stops between changes, and dynamic and tense movement, whereas happiness was associated with frequent tempo changes, longer stops between changes, dynamic movement, and changes between low and high tension. Sadness was portrayed by long and smooth movements, few tempo changes, and low tension, whereas fear was represented by frequent tempo changes, long stops between changes, high tension, and movements close the center of the body.

Camurri et al. (2003) and Camurri, Mazzarino, Ricchetti, Timmers, and Volpe (2004) developed a video-based analysis tool to recognize and classify expressive gestures in professional dance performances. This tool was used, for example, in Castellano, Villalba, and Camurri (2007), who studied dynamic qualities of motion cues as opposite to different movement shapes that might convey emotional qualities. They recorded the same gesture performed in four emotions (happy, angry, sad, and pleasurable) and could identify two main movement characteristics that differentiated between them: Quantity of Motion (QoM), defined as "the overall measure of the amount of detected motion, involving velocity and force" (p. 47), and Contraction Index (CI), defined as "the contraction and expansion of the body [. . .], i.e., the minimum rectangle surrounding the body" (p. 47) in a two-dimensional space. CoM could discriminate between high and low arousal (anger and joy being associated with high QoM, and sadness and pleasure with low QoM), while CI could distinguish between positive and negative emotions (joy and pleasure with low CI vs. anger and sadness with high CI).

Furthermore, Boone and Cunningham (1998) reported results of a dance study in which actors displayed basic emotions. They found that angry movements were associated with a great number of directional changes of the torso, as well as tempo changes, whereas happy movements were characterized by upward arm movements away from the torso. Sad movements were portrayed with downward gaze, low muscle tension, and a slackened body, and fearful movements were described as more rigid with a head-up "alert" posture.

Besides being an important cue in music performance and professional dance, movement also plays a considerable role in every-day music behavior. Keller and Rieger (2009), for example, stated that simply listening to music can induce movement, and in a self-report study conducted by Lesaffre et al. (2008), most participants reported moving when listening to music. Janata, Tomic, and Haberman (2011) reported in a study on groove that the wish to move along with the music is related to positive affect. They furthermore asked participants to tap to the music and found that, with increasing groove in the music, participants not only moved the finger/hand, but also other body parts, such as feet and head. Additionally, the tapping condition (isochronous versus free tapping) influenced the amount of movement: the more "natural" the tapping condition, the more movement was exhibited.

In general, people tend to move to music in an organized way; for example, by mimicking instrumentalists' gestures or rhythmically synchronizing with the pulse of the music by tapping the foot, nodding the head, or moving the whole body in various manners (Leman & Godøy, 2010). Moreover, Leman (2007) suggests, "Spontaneous movements [to music] may be closely related to predictions of local bursts of energy in the musical audio stream, in particular to the beat and the rhythm patterns" (p. 96). Such utilization of the body is the core concept of embodied cognition, which claims that the body is involved in or even required for cognitive processes (e.g., Lakoff & Johnson, 1980, 1999, or Varela, Thompson, & Rosch, 1991). Human cognition is thus highly influenced by the interaction and interplay of mind/brain, sensorimotor capabilities, body, and environment, and determined by the body with its specific shape, including particular perceptual and motor capabilities. Following this, we can approach music (or musical involvement) by linking our perception of it to our body movement (Leman, 2007). One could postulate that our bodily movements might reflect, imitate, help parse, or support the understanding of the content of music, be it musical features, such as beat, rhythm, melody, or tonality, or emotional characteristics. Leman suggests that corporeal articulations could be influenced by three (coexisting) components or concepts: Synchronization, Embodied Attuning, and Empathy, which differ in the degree of musical involvement and the kind of action-perception couplings. Synchronization forms the fundamental component, as synchronizing to a beat is easy, spontaneous, and possible even when paying minimal attention to the music. The beat serves as the basic musical element from which more complex structures emerge. Leman thus suggests the term *inductive resonance*, which refers to more active control, imitation, and prediction of movements to display beat-related features in the music (the opposite of passively tapping to a beat) as the first step in engaging with the music. The second component, Embodied Attuning, concerns the linkage of body movement to musical features more complex than the basic beat, such as melody, harmony, rhythm, tonality, or timbre. Following this idea, movement could be used to reflect and imitate the musical structure in order to understand it. Finally, Empathy is seen as the component that links music, or rather musical features, with expressivity and emotions. In other words, the listener feels and identifies with the emotions expressed in the music and imitates and reflects them by using body movement.

The aim of this study was to investigate relationships between the emotional content of music and characteristics of music-induced, quasi-spontaneous movement. We first conducted a motion-capture experiment (Experiment 1) and computationally extracted various movement characteristics from these data. Additionally, the perceived emotional content of the music used in the first part was gathered in a rating experiment (Experiment 2). Based on the considerations mentioned above, we chose to represent the emotional content with a two-dimensional model containing valence and arousal, and the four basic emotions happiness, anger, sadness, and tenderness (corresponding to the four quadrants of the dimensional model). We restricted this study to the use of perceived emotions as opposed to felt emotions, since previous literature has mostly focused on the former. It is important to distinguish between perceived and felt emotions, as these two kinds of responses have been found to differ from each other (e.g., Evans & Schubert, 2008; Gabrielsson, 2002).

In this study we were mainly interested in two aspects. First, we investigated whether movement and perceived emotional content are related, and more particularly, which body parts and movement characteristics can be associated with different emotions expressed in music. Second, we attempted to link our results to Leman's (2007) framework on corporeal articulations.

Based on the studies mentioned above (emotion-specific nonverbal behavior, music performance, and professional dance), we predicted that different emotions would be associated with different combinations of movement features as follows:

1. Music perceived as active would result in a high amount of movement occupying a large area, as active music might be encouraging listeners to move.
2. Active emotions, such as happiness and anger, would show similar patterns, with smoother movements observed for happy music than for angry music.
3. Inactive music, including sad and tender music, would result in a small amount of movement occupying a small area, as inactive music is less movement-stimulating than active music.
4. Happy, and probably pleasant, music would be related to body rotation (around the vertical axis).
5. Active emotions, such as happiness and anger, would lead to a more upward directed and stretched torso, whereas for inactive emotions (sadness and tenderness), the torso would be more directed forwards.
6. Happy (and pleasant) stimuli would be related to large hand distance.

Based on the first four predictions, we further assumed that the overall movement complexity would increase with active (especially happy) music and decrease with inactive music (especially sad) music, since active music is more movement stimulating than inactive music. To this end, we developed a feature that operationalizes the overall complexity of full-body movement.

Method

The data used in this study were obtained in two consecutive experiments. In Experiment 1, music-induced movement data were collected. In Experiment 2, the (perceived) emotional content of the stimuli used in Experiment 1 was assessed.

EXPERIMENT 1 – MOVEMENT TASK

Participants. A total of 60 participants took part in Experiment 1 (43 females; age range = 19-32, M = 24.0,

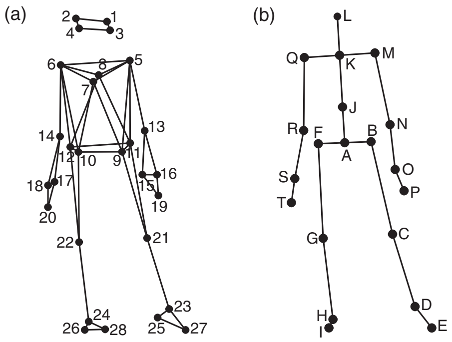


FIGURE 1. (a) Anterior view of the location of the markers attached to the participants' bodies; (b) Anterior view of the locations of the secondary markers/joints used in the analysis.

$SD = 3.3$). Six participants had received formal music education, and three participants were dance pedagogues. Participants were rewarded with a movie ticket.

Stimuli. Participants were presented with 30 randomly ordered musical stimuli representing the following popular music genres: Techno, Pop, Rock, Latin, Funk, and Jazz. All stimuli were naturalistic (existing pop music) to make the experiment as ecologically valid as possible, and had a uniform emotional content throughout. Participants were on average unfamiliar with every stimulus. All stimuli were 30 s long, nonvocal, and in 4/4 time, but differed in their rhythmic complexity, pulse clarity, and tempo. An overview of the stimuli can be found in Appendix A.

Apparatus. Participants' movements were recorded using an eight-camera optical motion capture system (Qualisys ProReflex) tracking, at a frame rate of 120 Hz, the three-dimensional positions of 28 reflective markers attached to each participant. The locations of the markers can be seen in Figure 1a, and can be described as follows (L = left, R = right, F = front, B = back): 1: LF head; 2: RF head; 3: LB head; 4: RB head; 5: L shoulder; 6: R shoulder; 7: breastbone; 8: spine; 9: LF hip; 10: RF hip; 11: LB hip; 12: RB hip; 13: L elbow; 14: R elbow; 15: L wrist/radius; 16: L wrist/ulna; 17: R wrist/radius; 18: R wrist/ulna; 19: L middle finger; 20: R middle finger; 21: L knee; 22: R knee; 23: L ankle; 24: R ankle; 25: L heel; 26: R heel; 27: L big toe; 28: R big toe. The musical stimuli were played back via a pair of Genelec 8030A loudspeakers using a Max/MSP patch running on an Apple computer.

Procedure. Participants were recorded individually and were asked to move to the stimuli in a way that felt natural. Additionally, they were encouraged to dance if they wanted to, but were requested to remain in the center of the capture space indicated by a 115 x 200 cm carpet.

Dancing in a motion capture lab might appear as a very artificial situation, and it is expectable that participants felt strained and uneasy. However, the experimenters made sure that the participants felt as comfortable as possible and did not watch the participants during dancing, as they were concealed by the technical equipment. Most participants reported after the experiment that they enjoyed the experience and quickly forgot about the equipment and the experimenters. Only one participant felt so uncomfortable that she decided to discontinue the experiment. Therefore, we think that participants were able to fulfill the task without being much biased by the situation and environment.

Movement feature extraction. In order to extract various kinematic features, the MATLAB Motion Capture (MoCap) Toolbox (Toiviainen & Burger, 2011) was used to first trim the data to the duration of each stimulus. To remove redundant data while retaining the essential body parts and joints, a set of 20 secondary markers–subsequently referred to as joints–was derived from the original 28 markers. The locations of these 20 joints are depicted in Figure 1b. The locations of joints C, D, E, G, H, I, M, N, P, Q, R, and T are identical to the locations of one of the original markers, while the locations of the remaining joints were obtained by averaging the locations of two or more markers; joint A: midpoint of the four hip markers; B: midpoint of markers 9 and 11 (left hip); F: midpoint of markers 10 and 12 (right hip); J: midpoint of breastbone, spine, and the hip markers (midtorso); K: midpoint of shoulder markers (manubrium), L: midpoint of the four head markers (head); O: midpoint of the two left wrist markers (left wrist); S: midpoint of the two right wrist markers (right wrist). From the three-dimensional joint position data, the instantaneous velocity and acceleration were estimated using numerical differentiation based on the Savitzky-Golay smoothing FIR filter (Savitzky & Golay, 1964) with a window length of seven samples and a polynomial order of two. These values were found to provide an optimal combination of precision and smoothness in the time derivatives. Thereafter, ten movement features were extracted from these data:

- Postural features:
  - Torso Tilt: vertical tilt of the torso (Joints A-K), positive tilt being related to bending forward.

  - Hand Distance: distance between hands (Joints P and T).
  - Foot Distance: distance between feet (Joints (E and I).
- Local kinematics:
  - Magnitude of Head Acceleration (Joint L).
  - Magnitude of Hand Acceleration (Joints P and T).
  - Magnitude of Foot Acceleration (Joints E and I).
- Global kinematics:
  - Area of Movement: the smallest rectangle that contains the projection of the trajectory of the Center of Mass (Joint A) on the horizontal plane (i.e., floor), averaged across four-second analysis windows with a two-second overlap.
  - Fluidity: overall movement fluidity/smoothness measure based on the ratio of velocity to acceleration. The combination of high velocity and low acceleration reflects fluid movement, whereas the combination of low velocity and high acceleration reflects non-fluid movement.
  - Rotation Range: amount of rotation of the body (joints M and Q) around the vertical axis.
  - Movement Complexity: based on Principal Components Analysis (PCA) of the position data (all joints). PCA, in this case, can be understood as a decomposition into different (independent) movement components. The PCA was performed on individual subject level, i.e., one PCA per participant and stimulus. Subsequently, the cumulative sum of the proportion of (explained) variance contained in the first five PCs was determined. We found that for five PCs the cumulative variance had the highest range of variation across movement recordings, so using five PCs discriminated best between complex and non-complex movement. The proportion of variance explained with the first five PCs varied between 67.0% and 99.8%, with an average of 85.5% (subject level). For an excerpt with 99.8% explained variance, almost all movement can be explained with the first five components, whereas for an excerpt with 67.0% explained variance, 33% cannot be explained by the first five components (and more components are needed). Thus, a high proportion of unexplained variance (i.e., a low cumulative sum) would mean that the underlying movement is complex, as a high number of PCs is needed to explain the movement sufficiently. A low proportion of unexplained variance (i.e., a high cumulative sum), on the other hand, implies a simpler movement, as it can be sufficiently explained with a low number of PCs. Figure 2 illustrates this feature by displaying projections of the first five principal components of complex and non-complex movements.

Apart from the Area of Movement feature, all movement features were calculated based on a local coordinate system, in which joint A was located at the origin, and segment BF was parallel to the x-axis of the coordinate system.

Subsequently, the instantaneous values of each variable were averaged across the participants for each stimulus presentation. This yielded a total of ten statistical movement features for each of the 30 stimuli.

EXPERIMENT 2 – PERCEPTUAL TASK

Participants. Thirty-four Finnish musicology students participated in experiment 2 (17 females; age range = 21-47, $M = 25.7$, $SD = 5.9$). All participants were familiar with concepts of music and emotion research, such as different emotion models and research on perceived and felt emotions. None of the participants of this experiment participated in Experiment 1.

Stimuli. The same 30 musical excerpts as in Experiment 1 were used. To conduct the experiment in a reasonable time, we shortened the stimuli to 15 s by removing the first and last 7.5 s of the original excerpt. As both the emotional content of the stimuli and the characteristics of the movement were stable throughout each stimulus (correlations between movement features calculated for 15- and 30-s excerpts and the perceptual ratings resulted in a very similar pattern), shortening the stimuli did not have a notable effect on the correlations between movement and perceptual features.

Procedure. The participants were randomly divided into two groups that were presented with the stimuli at the same time, but in two different random orders. Each stimulus was presented once, followed by a break of 40 s. The participants' task was to rate each stimulus according to its emotional content on six different seven-point scales:

- Arousal: inactive – active
- Valence: unpleasant – pleasant
- Happiness: not happy – happy
- Sadness: not sad – sad
- Tenderness: not tender – tender
- Anger: not angry – angry

Participants were explicitly told in the instructions to rate the excerpts according to what they thought the music expresses/conveys and not what they personally felt when listening to the music.

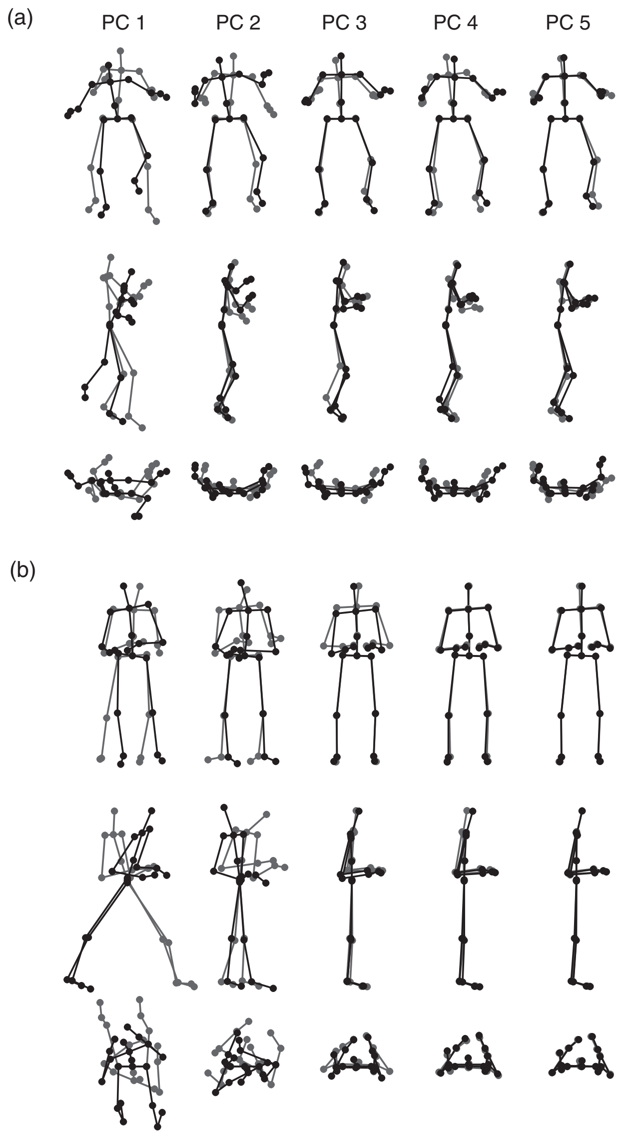


FIGURE 2. Movement complexity illustrated by displaying the projections of the first five principal components of complex movement (a) and non-complex movement (b). The extreme deflections of the projections/dimensions are plotted from the front, from the side, and from the top. (a) High movement complexity, as movement is visible in all five PC projections, thus a high number of principal components is needed to explain the movement; (b) Low movement complexity, as most movement is found in the projection of PC 1—a low number of PCs can sufficiently explain the movement.

TABLE 1. *Results of the Intraclass Correlations for Both the Movement Features (a) and the Ratings of the Emotional Contents (b).*

(a)		(b)	
Movement Feature	r	Emotion	r
Torso Tilt	.86**	Arousal	.98**
Hand Distance	.62**	Valence	.94**
Foot Distance	.73**	Happiness	.95**
Head Acceleration	.95**	Anger	.97**
Hand Acceleration	.94**	Sadness	.92**
Foot Acceleration	.94**	Tenderness	.95**
Area of Movement	.38*		
Fluidity	.96**		
Rotation Range	.60**		
Movement Complexity	.77**		

* $p < .05$, ** $p < .0001$

Results

In order to perform further analysis, we first assessed interparticipant consistency by calculating intraclass correlations (cf. Shrout & Fleiss, 1979) for both the movement features and the ratings of the emotional contents. The values are displayed in Table 1. These values show that all movement features and emotion ratings received significant correlation values. Thus, the interparticipant consistency was sufficiently high for subsequent averaging of movement features and emotion ratings across participants to receive one value per stimulus.

In the first part of the analysis, the six rating items were correlated with the ten movement features to investigate the relations between the emotion ratings and the music-induced movement features. Due to the large number of correlations (60), we utilized false discovery rate control by applying the Benjamini-Hochberg procedure (see Benjamini & Hochberg, 1995). At level $p = .05$, 17 correlations were retained, with a largest p-value of .01. The results are shown in Table 2.

For the group of Postural Features, Torso Tilt correlated significantly with characteristics of the emotional content: negatively with Arousal, $r(30) = -.70$, $p < .001$, and positively with Tenderness, $r(30) = .58$, $p < .001$. Additionally, Foot Distance correlated negatively with Happiness, $r(30) = -.45$, $p < .05$. These results suggest that the participants had a more upright torso when the music was considered active, and were more bent forward during tender music, and that they had a smaller foot distance when the music was happy.

All three Acceleration Features (Local Features group) showed high positive correlations with Arousal: Head Acceleration: $r(30) = .75$, $p < .001$; Hand Acceleration: $r(30) = .76$, $p < .001$; Foot Acceleration: $r(30) = .68$, $p < .001$. Additionally, Head and Hand Acceleration showed significant negative correlations with Tenderness: Head Acceleration: $r(30) = -.63$, $p < .001$; Hand Acceleration: $r(30) = -.47$, $p < .01$. Thus, participants tended to use more accelerated movements of the whole body for active music, and less acceleration of the upper body when the music was considered tender.

Three of the Global Features–Fluidity, Rotation Range, and Movement Complexity–correlated significantly with the emotional content: Fluidity exhibited negative correlations with Arousal, $r(30) = -.84$, $p < .001$, and Anger, $r(30) = -.47$, $p < .01$, and positive correlations with Valence, $r(30) = .45$, $p < .05$, and Tenderness, $r(30) = .66$, $p < .001$, suggesting that participants moved in a fluid way with pleasant and tender music, but more irregularly when the music had an active or angry character. Rotation Range, in addition, correlated positively with Valence, $r(30) = .55$, $p < .01$, and Happiness, $r(30) = .55$, $p < .01$, and negatively with Anger, $r(30) = -.47$, $p < .01$, suggesting that participants rotated their body to pleasant and happy music, whereas they performed less body rotation with music that expressed

TABLE 2. *Results of the Correlation Between the Movement Features and the Rating Items.*

	Arousal	Valence	Happiness	Anger	Sadness	Tenderness
TorsoTilt	**−.70*** **	.29	−.13	−.32	.39	**.58** **
HandDist	.07	.22	.15	−.32	.08	.11
FootDist	−.31	−.38	**−.45***	.34	.38	−.23
HeadAcc	**.75*** **	−.36	.10	.37	−.30	**−.63** **
HandAcc	**.76*** **	−.18	.25	.15	−.40	**−.47***
FootAcc	**.68*** **	.02	.41	−.04	−.43	−.28
MoveArea	.30	.31	.40	−.19	−.33	.11
Fluidity	**−.84*** **	**.45***	−.07	**−.47***	.39	**.66*** **
RotRange	.14	**.55** **	**.55** **	**−.47***	−.27	.38
MoveComp	.08	.42	**.54***	−.42	**−.58** **	.19

* $p < .05$, ** $p < .01$, *** $p < .001$

anger. Movement Complexity showed a significant positive correlation with Happiness, $r(30) = .54$, $p < .01$, and a significant negative correlation with Sadness, $r(30) = -.58$, $p < .001$. Thus, participants tended to use complex movements for happy music, and simpler movements for music that expressed sadness.

Besides the (discrete) emotion ratings, we were interested in the embodiments and relationships between movement features and basic emotion ratings in a two-dimensional space based on the valence and arousal ratings. Thus, the second part of the analysis was motivated by Russell's (1980) circumplex models of affect and the positioning of discrete emotions as polar coordinates in a space consisting of valence and arousal dimensions. Russell asked participants to sort 28 emotional terms (such as happy, excited, angry, or bored) in different ways and created five circular models by computing polar coordinates for each of the 28 terms based on different sorting approaches and scaling methods. These polar coordinates started from the origin (of the circular space) and were specified by the angle and the length based on the sorter agreement.

As an initial step, we developed an angular variable that we called *Circular Affect*, derived from the Valence and the Arousal ratings of each stimulus. This was carried out by first representing the two ratings for Valence and Arousal for each stimulus as a vector in a two-dimensional space (the x-coordinate represents the Valence rating, the y-coordinate represents the Arousal rating). A vector projection at any possible polar angle was then applied to this vector, to receive a value for every angle. The outcome of this approach, the Circular Affect variable, is displayed in Figure 3a. A mathematical description of the derivation is given in Appendix B.

Subsequently, Circular Affect was correlated with the ten movement features and with the four basic emotion ratings. As the last step, the maximal correlation values and the corresponding angles were obtained. This procedure is exemplified in Figure 3b, showing the correlation results of one movement feature (Foot Acceleration) with Circular Affect.

Figure 4 displays the maximal correlation values and the respective polar angles for the ten movement features in a two-dimensional space in relation to Arousal and Valence. Table 3 shows the same data numerically. Due to the definition of Circular Affect (see Appendix 2), the minimal correlations are obtained by reflecting the respective projection vectors about the origin (i.e., rotating by 180°). The respective angles of the minimal correlations are given in brackets, as they will offer further insights into relationships between our results and Russell's (1980) circumplex models.

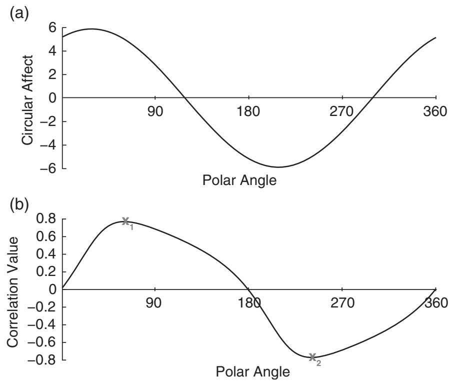


FIGURE 3. (a) Display of Circular Affect, i.e., the polar projection of the (averaged) rating for Valence and Arousal of one stimulus plotted as the function of the polar angle. (b) Correlation values between Circular Affect and Foot Acceleration plotted as the function of the polar angle. X_1 depicts the location of the maximal correlation, whereas X_2 depicts the location of the minimal correlation.

As a first step into analyzing the results, we would like to point out relationships between these results and the results of the previous analysis (discrete ratings, Table 2). Movement features for which the direction of the maximal correlation is close to the Arousal axis, such as Torso Tilt and Head Acceleration (cf. Figure 4), showed, in the discrete rating analysis (cf. Table 2), high correlations for Arousal, but low ones for Valence, whereas features located between the axes (e.g., around 45° or 225°), such as Area of Movement or Foot Distance (while being significant in this analysis), exhibited low correlations for both discrete Valence and Arousal ratings (cf. Table 3).

It can be seen from Table 3/Figure 4 that there are a large number of high correlations between the movement features and Circular Affect. For the postural group, the correlations of Circular Affect with both Torso Tilt and Foot Distance were significant, with the maximal correlation value at an angle of 268° (minimal correlation angle: 88°) ($|r(30)| = .71$, $p < .001$), respectively, for Torso Tilt, and at an angle of 217° (min. corr.: 37°) ($|r(30)| = .65$, $p < .001$) for Foot Distance. These results suggest that participants had the most forward directed torso when the music was considered inactive, and the most upward directed torso when the music was considered active. Foot distance was biggest when the music was inactive and unpleasant, and smallest when the music was active and pleasant.

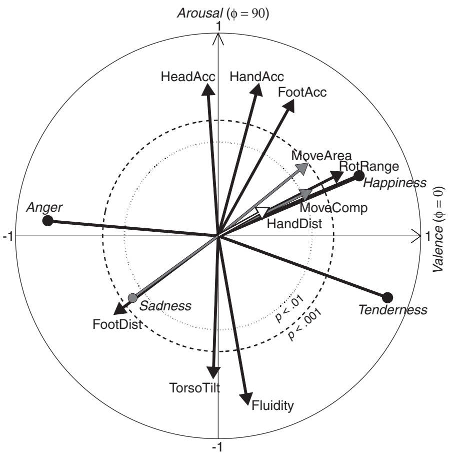


FIGURE 4. Maximal values of the correlations between Circular Affect and the ten movement features and the four basic emotion ratings respectively arranged in a circular model that we called Circular Affect Space (see Table 3 for values and angles in numerical form). The length of the arrows reflects the maximal value of the correlation, and the polar angle indicates the direction of the maximal value. Black arrowheads indicate a significance level of $p < .001$, grey arrowheads of $p < .01$, and white arrowheads are not significant. Movement features are indicated by sharp arrowheads, while the basic emotion ratings are indicated by rounded arrowheads.

All three Acceleration features correlated significantly with Circular Affect, with maximal values at 94° (min. corr.: 274°) ($|r(30)| = .75$, $p < .001$) for Head Acceleration, at 75° (min. corr.: 255°) ($|r(30)| = .78$, $p < .001$) for Hand Acceleration, and at 61° (min. corr.: 241°) ($|r(30)| = .77$, $p < .001$) for Foot Acceleration. Thus, the feet displayed highest acceleration when the music was active and pleasant, and lowest during music of the opposite character, whereas high hand and head acceleration was mostly observed when the music was active with neutral valence, and low hand and head acceleration with inactive music.

For the correlations of the Global Features with Circular Affect, the maximal values were also significant. For Area of Movement, it fell at an angle of 42° (min. corr.: 222°) ($|r(30)| = .64$, $p < .001$), for Fluidity at an angle of 280° (min. corr.: 100°) ($|r(30)| = .85$, $p < .001$), and for Rotation Range at an angle of 27° (min. corr.: 207°) ($|r(30)| = .70$, $p < .001$). This outcome suggests that participants covered the largest area with pleasant and active music and the smallest area when the music was unpleasant and inactive. Participants were most fluid in their movement when the music was inactive and slightly pleasant, and least fluid with active and slightly unpleasant music. Furthermore, participants tended to rotate their body most for pleasant and slightly active music, and least for unpleasant and inactive music. Furthermore, Movement Complexity was found to have the maximal value of the correlation with Circular Affect at 26° (min. corr.: 206°) ($|r(30)| = .52$, $p < .01$). Thus, participants tended to use complex movement for pleasant and slightly active music, and simple movement for unpleasant and slightly inactive music.

TABLE 3. *Maximal Values and Corresponding Polar Angles in the Circular Affect Space of the Correlations Between Circular Affect and the Ten Movement Features as well as Between Circular Affect and the Basic Emotion Ratings.*

Movement Feature/Basic Emotion Rating	Correlation	Max. Corr. Angle
Torso Tilt	**.71*****	268°
Hand Distance	.29	29°
Foot Distance	**.65*****	217°
Head Acceleration	**.75*****	94°
Hand Acceleration	**.78*****	75°
Foot Acceleration	**.77*****	61°
Area of Movement	**.57****	39°
Fluidity	**.85*****	280°
Rotation Range	**.70*****	27°
Movement Complexity	**.52****	26°
Happiness	**.80*****	23°
Anger	**.88*****	175°
Sadness	**.56****	216°
Tenderness	**.93*****	340°

** $p < .01$, *** $p < .001$
Note: We controlled the false discovery rate by using the Benjamini-Hochberg procedure.

Additionally, the basic emotion ratings were correlated with Circular Affect to investigate the relationship between the different ratings. The results are depicted in Figure 4 as well. All four maximal correlation values were significant. For Happiness, the maximal value was at 23°, $r(30) = .80$, $p < .001$, for Anger at 175°, $r(30) = .88$, $p < .001$, for Sadness at 216°, $r(30) = .56$, $p < .01$, and for Tenderness at 340°, $r(30) = .93$, $p < .001$.

Discussion

The results of this study suggest relationships between the emotional content of music and characteristics of music-induced movement. Each emotion rating correlated significantly with a different set of movement features that could therefore be assumed to be emotion-specific.

Furthermore, the results provide support for Leman's (2007) concepts Embodied Attuning and Empathy. Movement features and perceived emotional content of music were found to correlate, so participants might have (unconsciously) used their body to express and reflect the affective content. However, the emotional content is created and shaped by features of the music, such as rhythm, timbre, or tonality–characteristics that could also be linked to the concept of Embodied Attuning. Previous research showed that participants rated the perceived emotional content of music consistently (Eerola & Vuoskoski, 2011), and in order to change the emotional characteristics, musical features have to be changed. Thus, it may be difficult to distinguish between Embodied Attuning and Empathy, suggesting that they are overlapping and coexisting. However, since music can express and convey emotional characteristics, one could assume that the affective content of music has an unconscious effect on music-induced movement and suggests understanding Empathy as an abstraction of Embodied Attuning.

Since previous studies followed different purposes and used different techniques and a wide variety of movement descriptors that partly differ from ours, the comparison of the results is difficult. However, the results of our study still follow general directions proposed in previous studies. Our findings are quite in line with those by Camurri et al. (2003), Boone and Cunningham (1998), and Dahl and Friberg (2007) in that the quasi-spontaneous music-induced movements exhibited qualities similar to movements of professional dancers who embedded different emotional expressions in their dancing, and also to movements of instrumentalists who conveyed certain emotions while playing. Furthermore, our results confirm, at least partly, the hypothesized similarities to studies on emotion-specific nonverbal behavior, in particular to findings by De Meijer (1989), Darwin (1872), or Wallbott (1998). Finally, our results were consistent with the locations of emotional terms in Russell's (1980) circumplex models.

Basic emotions. Happiness correlated positively with Rotation Range and Movement Complexity, suggesting that happy music was expressed by using body rotation and high-dimensional movements. Although these particular movement features were not part of previous research, they might be related to features that were used in other studies. For example, the descriptions by Camurri et al. (2003) (dynamic movement, changes between low and high tension, and frequent tempo changes with long stops) or by De Meijer (1989) (fast, active, open, light, upward directed) could result in complex movement. In previous research, movements expressing happiness were found to be related to movements of the hands and arms (e.g., Boone & Cunningham, 1998; Dael et al., 2012). Although, in the present analysis, hand acceleration did not significantly correlate with the happiness ratings, hand and especially arm movements can be expected to increase the complexity of movements, as the arms offer the best possibilities to move freely in different directions and independently from other body parts. Moreover, the relationship between Rotation Range and Happiness is supported by Burger and Bresin (2010), whose robot conveyed happiness by performing circular movements. Happiness also correlated negatively with Foot Distance. A smaller distance between the feet could be a result of increased movement complexity and rotation for happy music, as the feet are more active and potentially closer together (and even crossing in case of rotating around the vertical axis).

Anger was highly negatively correlated with Fluidity and Rotation Range, indicating that participants, when being exposed to music that is considered angry, were moving in an irregular and jagged way while not rotating the body. The co-occurrence of non-fluid movements and anger is well supported by the literature: Dahl and Friberg (2007) mentioned low fluency, Camurri et al. (2003) frequent tempo changes and short movements, Boone and Cunningham (1998) directional and tempo changes, and Darwin (1872) trembling. Jerkiness and non-fluency of movement might explain the negative correlation between Rotation Range and Anger, as smooth rather than jerky movements are associated with full rotations of the body.

The significant negative correlation between Sadness and Movement Complexity suggests that sad music is reflected in rather simple movements of low dimensionality. This result seems straightforward and is also supported by previous research. For example, Dahl and Friberg (2007) referred to low amount and speed of movement, Camurri et al. (2003) to low tension, few changes, long and smooth movements, and Wallbott (1998) to inactive movements covering little space–all characteristics that indicate movement of low complexity.

Tenderness was found to exhibit significant positive correlations with Torso Tilt and Fluidity and significant negative correlations with Acceleration of Head and Hands, indicating that tender music was embodied through a forward tilted torso and fluid movements with low acceleration. These movement characteristics would certainly fit to tender music, as tender music might not consist of strong and easily perceivable beats. Additionally, a forward tilted body, opposite to an upright torso showing alertness, could emphasize the relaxed character of tender music.

Dimensional model. Variables of the dimensional model have rarely been used in this field of research so far. However, active music was characterized by a high amount of movement and other movements that show the activity in the music, such as high acceleration of head, hands and feet, non-fluent movement, and an upright torso. In previous research (Boone & Cunningham, 1998; Camurri et al., 2003; Dahl & Friberg, 2007; Wallbott, 1998), acceleration-related movement features were associated with the expression of anger and happiness, both considered active emotions. Thus, it would be straightforward to assume that active music is reflected by similar movement characteristics. Moreover, an upright torso was found to be related to happiness and anger (Darwin, 1872; Wallbott, 1998), though De Meijer (1989) and Coulson (2004) reported an upward directed torso for happiness, but a bowed torso for angry. However, this discrepancy might suggest torso tilt to be a feature that discriminates pleasant and unpleasant active emotions.

The high negative correlation of Arousal with Fluidity is somewhat surprising, especially because the correlation value for Anger is lower than for Arousal. Based on related research, one could assume angry music to cause the highest amount of non-fluent movement (Boone & Cunningham, 1998; Dahl & Friberg, 2007), though our results might suggest non-fluid movements to be more characteristic for active than for angry music, at least in our set of stimuli.

Interesting to note is that Tenderness and Arousal exhibit a reverse correlation pattern. This suggests that Arousal and Tenderness might share the same movement features, but with opposite characteristics. The relation of both emotions will be discussed again in the following section.

Fluidity and Rotation Range were the only two significant movement characteristics for pleasant music, thus participants moved smoother and were rotating their body remarkably more when being exposed to pleasant music. Rotation Range was also found to be significant for happy music, thus a pleasant emotional character in music in general could be associated with body rotation.

Circular affect and the circular affect space. The Circular Affect Space indicates the locations of the maximal and minimal correlations between Circular Affect and the movement features, and subsequently reveals interesting insights into the embodiments of the emotion ratings related to the dimensional model. The polar angles corresponding to the maximal correlations are dispersed within the Circular Affect Space with an accumulation in the pleasant active quadrant (cf. Figure 4). Movement features related to amount and complexity of movement, such as Movement Complexity, Area of Movement, or Acceleration, can be found there, thus music perceived as pleasant and active might be encouraging to dance or move to. Likewise, the minimal correlations would cluster in the unpleasant inactive quadrant, suggesting that unpleasant and inactive music decreases the complexity and amount of movement due to the fact that such music is not as inviting to move along with.

Movement Complexity and Rotation Range resulted in similar angles (26° and 27°) for the maximal correlation, which could suggest that they form a characteristic pattern for pleasant, slightly active music. Additionally, these two movement features were found to correlate significantly with the ratings for Happiness, which would indicate Happiness to be located at around 26/27° in the Circular Affect Space in terms of the embodiment of musical emotions. This suggestion is supported by the correlation results between the happiness ratings and Circular Affect, which positions the maximal correlation at an angle of 23°.

Furthermore, Rotation Range was found to be related to Valence in the dimensional model analysis. Thus, body rotation seems to be especially characteristic for pleasant music.

In the basic emotions analysis, anger ratings correlated negatively with Fluidity and Rotation Range. The minimal correlation of Fluidity and Circular Affect falls into the (upper) unpleasant active quadrant, whereas the minimal correlation of Rotation Range and Circular Affect falls into the (upper) unpleasant inactive quadrant. The relatively low correlation values of both features with the anger ratings ($r = -.47$) could motivate the conclusion that the middle of both minimal correlations (around 150°) is the estimated location for Anger in the Circular Affect Space regarding the embodiment of musical emotions. This estimated location is relatively close to the position of the maximal correlation value of the anger ratings with Circular Affect at 175°, which would support the consistency between the discrete ratings to some extent. However, a more active position could have been expected, since anger is usually related to both (positive) arousal and (negative) valence, so this result is somewhat surprising.

The minimal correlations that lie in the unpleasant inactive quadrant–especially Movement Complexity–could be linked to Sadness. Low Movement Complexity was found to be characteristic for sad music in the basic emotions analysis, suggesting that Sadness might be located at around 205° in the Circular Affect Space. This angle might appear as "too active," though it is rather difficult to find genuinely sad stimuli that encourage movement. The correlation between the sadness ratings and Circular Affect resulted in a maximal value located at 216°, which would be consistent with the suggested location mentioned above. Interesting to note is that the location of Sadness is very similar to the location of Foot Distance, although Foot Distance failed to significantly correlate with Sadness in the discrete emotion analysis. Since Foot Distance was not part of previous studies, it is rather difficult to give suitable explanations here. A possible explanation could nonetheless be that participants adopted more stable footing with feet spaced wider apart, since sad music also led participants to move less (cf. the basic emotion rating results discussed above).

By observing the minimal correlations, the characteristic movement pattern for tender music could be linked to this analysis as well. Head and Hand Acceleration received significant negative correlations with the ratings for Tenderness, thus it could be expected that the minimal correlation with these two movement features would be located in the Tenderness (or pleasant inactive) quadrant. However, they are close to the arousal axis, and thus more related to inactive neutral stimuli. It could therefore be suggested that Tenderness might be located quite close to the Arousal axis, which would explain the opposite correlation pattern of Tenderness and Arousal. The lower correlation value for Hand Acceleration (r = -.47) and the higher value for Head Acceleration (r = -.63) could lead to the conclusion that Tenderness might be located in the pleasant inactive quadrant somewhat close to the negative Arousal axis. The maximal correlations of (forward) Torso Tilt and Fluidity would also support the suggested location of Tenderness. The location of the maximal correlation value of the tenderness ratings and Circular Affect (340°), however, is closer to the Valence axis than to the Arousal axis. Thus, this result is somewhat surprising, but maybe just reflects the nature of the stimuli being rather active in order to stimulate movement.

Maximal Fluidity, Torso Tilt, and minimal Hand, Head, and Feet Acceleration fall into the same region of the Circular Affect Space between Tenderness and Inactivity, thus participants were probably not able to distinguish between the two emotional states, which could furthermore explain the opposite correlation pattern of Tenderness and Arousal.

One movement feature, Area of Movement, that failed to be significant in the basic emotions analysis, was found to be significant in the Circular Affect analysis. The maximal correlation of Area of Movement with Circular Affect falls into the pleasant active quadrant; thus, listeners tend to cover a large area when moving to pleasant and active music. This finding could be linked to research by Wallbott (1998), who noted that movements displaying active emotions, such as happiness and anger, covered more space than movements displaying sadness, probably because sad music rather inhibits than encourages movement.

The circular affect space in relation to Russell's (1980) circumplex models. In what follows, we will compare and link the Circular Affect Space to Russell's (1980) circumplex models. We will refer here to the two models that were derived by direct circular scaling and by multidimensional scaling. Table 4 displays the emotions of Russell's circumplex models that are most proximate to the locations of the maximal and minimal correlations between movement features and Circular Affect.

In case of the direct circular scaling, Russell's 'delighted' (25°) is most closely located to Movement Complexity, Rotation Range, and Hand Distance, whereas 'happy' (8°) is located too low on the arousal axis to match any of our movement feature correlations. However, both emotions are positive, thus they could be related to similar movement features. The location of the maximal correlation between Circular Affect and Area of Movement is closest to 'excited' (49°), indicating that participants cover a bigger area when moving to exciting and activating music. The location of 'aroused' (74°) in the upper pleasant active quadrant fits with Acceleration of Feet and Hands, which supports our basic emotion/dimensional model analysis. However, Head Acceleration is closest to 'alarmed' (96°), and also close to 'angry' (99°), which did not correlate in our basic emotions analysis. Foot Distance is located close to 'depressed' (210°) and 'sad' (208°), suggesting that a large distance between feet could be related to unpleasant inactive music. Torso Tilt can be found close to 'tired' (268°), which is straightforward, as one bends forward when being tired. Fluidity is closest located to 'sleepy' (272°), which suggests inactive neutrally pleasant stimuli to be embodied with fluid movements.

The angles corresponding to the minimal correlations for Circular Affect with Rotation Range, Movement Complexity, and Hand Distance, are close to Russell's location of 'sad,' which supports the above suggested location for sadness in the Circular Affect Space and furthermore the finding that sad music can be related to low complexity of movement. Area of Movement is close to 'depressed' and ', suggesting that unpleasant

TABLE 4. *Emotions of Russell's (1980) Circumplex Models that are Closest Located to the Maximal and Minimal Correlations Between Circular Affect and the Movement Features.*

Movement feature	Russell Emotion Corresponding to Maximal Correlation Direction		Russell Emotion Corresponding to Minimal Correlation Direction	
	Direct Scaling	Multidimensional Scaling	Direct Scaling	Multidimensional Scaling
TorsoTilt	Tired	sleepy	tense	aroused
HandDist	Delighted	happy	sad	depressed
FootDist	Depressed	sad	delighted	delighted
HeadAcc	Alarmed	aroused	sleepy	sleepy
HandAcc	Aroused	astonished	droopy	sleepy
FootAcc	Astonished	excited	bored	tired
MoveArea	Exited	delighted	depressed	bored
Fluidity	Sleepy	sleepy	angry	aroused
RotRange	Delighted	happy	sad	depressed
MoveComp	Delighted	happy	sad	depressed

Note: The second and fourth columns of the table correspond to the circumplex model derived by direct circular scaling and the third and fifth columns to the circumplex model derived by multidimensional scaling.

inactive music is embodied with movements within a small area. The minimal correlations of Head, Hand, and Feet Acceleration are close to 'sleepy,' 'droopy' (256°), and 'bored' (241°), suggesting that inactive music is related to low acceleration of the body. Foot Distance is located between Russell's 'delighted' and 'excited,' indicating that small distance between feet might be characteristic for music with such emotional qualities. Torso Tilt is located close to 'tense' (92°), suggesting that tense music is embodied by an upward torso. The minimal correlation of Fluidity is close to the location of 'angry,' which would support both previous research and our basic emotions analysis.

In the case of the multidimensional scaling solution, 'happy' (19°) is located closest to the maximal correlations between Circular Affect and Movement Complexity and Rotation Range (and Hand Distance), and would therefore support our basic emotion analysis and reflect the location of Happiness in the Circular Affect Space. Area of Movement is close to 'delighted' (44°). Acceleration of Feet and Hands are near 'exited' (72°) and 'astonished' (77°), whereas Head Acceleration is closest to 'aroused' (93°), supporting our previous analyses. None of our maximal correlations match Russell's 'angry' (140°). Foot Distance seems again to be related to 'sad' (214°), and 'sleepy' (271°) is close to Torso Tilt and Fluidity.

The minimal correlations of Circular Affect and Rotation Range, Movement Conplexity, and Hand Distance are close to Russell's 'depressed' (204°), and Area of Movement is close to 'bored' (225°), while 'sad' (214°) is located between both emotions, suggesting that such unpleasant, inactive music is embodied with low complexity and simple movements using a small area. Acceleration of Feet, Hands, and Head are close to 'tired' (242°) and 'sleepy' (271°). Foot Distance is close to Russell's 'delighted,' whereas Torso Tilt and Fluidity are closest to 'aroused,' supporting that active music with neutral valence is embodied with an upright torso and non-fluid movements.

The locations of the maximal values of the correlations between the basic emotion ratings and Circular Affect are closer to the locations of Russell's multidimensional scaling solution than to his direct scaling solution. Both Happiness and Sadness obtained similar angles as in Russell's solution. Tenderness, not having a direct counterpart in Russell's analysis, is located close to Russell's 'serene' and 'calm,' suggesting their similarity. Russell's 'angry' can be found at a more active position than our correlational analysis has placed it, close to suggested location based on the correlation pattern between the anger ratings and the movement features.

Conclusion

This study offers insights into body-related nonverbal behavior, and indicates that there are relationships between the emotional content of music and music-induced, quasi-spontaneous movement. Characteristic movements and movement combinations were associated with different emotional characteristics of music. It is thus argued that the body reflects emotional qualities of music, which can be seen as support for the notion of embodied music cognition, especially for the concept of Empathy and its potential boundaries to Embodied Attuning. Furthermore, a circular variable, Circular

Affect, was derived from the arousal and valence ratings, and the locations of the maximal correlations between movement features and these compound ratings were displayed in a two-dimensional space. Both the results of the emotion ratings assessed in this study and Russell's circumplex model of affect show remarkable consistency with the Circular Affect model.

Previous research in the field tended to use professional actors, musicians, or dancers performing certain emotional states, with their attention focused on each emotion. It is likely that this approach results in exaggerated, artificial, and stylized movements. This study, on the other hand, focused on laypersons and investigated how the emotional content of music affects movement when participants are not actively and consciously paying attention to it. As a consequence, the resulting movements might be more diverse and difficult to analyze than movements performed by professional dancers. However, we believe that our approach is ecologically more valid than the acted emotions approaches for investigating the impact of emotional content of music on everyday behavior.

It could be argued that our participants moved rather according to musical features than according to the emotional content. However, in the present analysis approach, the role of musical features was not considered. Future work could therefore include extracting musical features and subsequently using them, for instance, in multiple regression models together with the movement features and the emotion ratings. This would give further insights into the relation between movement features, emotional content, and musical features, and also in possible distinction between Leman's (2007) concepts of Empathy and Embodied Attuning.

Finally, an emotion recognition experiment could be conducted in which participants are asked to rate the movements presented as point-light videos regarding the emotion conveyed or reflected. This would give insight into the perception of emotions in music-induced, quasi-spontaneous movements.

Author Note

We thank Mikko Leimu for help and technical assistance during the motion capture data collection. We also wish to thank the three anonymous reviewers for their thorough and helpful comments on the manuscript. This study was funded by the Academy of Finland (projects 118616, 125710, and 136358).

Correspondence concerning this article should be addressed to Birgitta Burger, Department of Music, P.O. Box 35, FI-40014 University of Jyväskylä, Finland. E-mail: birgitta.burger@jyu.fi

References

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*, 43-64.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*, 289-300.

Boone, R. T., & Cunningham, J. G. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology, 34*, 1007-1016.

Burger, B., & Bresin, R. (2010). Communication of musical expression by means of mobile robot gestures. *Journal on Multimodal User Interfaces, 3*, 109-118.

Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2010). Influence of musical features on characteristics of music-induced movements. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC)* (pp. 425-428). Seattle, WA: University of Washington.

Camurri, A., Lagerlöf, I., & Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies, 59*, 213-225.

Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., & Volpe, G. (2004). Multimodal analysis of expressive gesture in music and dance performances. In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in human-computer interaction. Lecture notes in computer science, 2915* (pp. 20-39). Berlin/Heidelberg, Germany: Springer.

Castellano, G., Villalba, S. D., & Camurri, A. (2007). Recognising human emotions from body movement and gesture dynamics. In A. Paiva, R. Prada, & R. W. Picard (Eds.), *Proceedings of the 2nd International Conference in Affective Computing and Intelligent Interaction (ACII)* (pp. 71-82). Berlin/Heidelberg, Germany: Springer.

Clynes, M. (1980). The communication of emotion: Theory of sentics. In R. Plutchik, & H. Kellerman (Eds.), *Theories of emotions* (Vol. 1, pp. 171-216). New York, NY: Academic Press.

Clynes, M. (1992). Time-forms, nature's generators and communicators of emotion. In *Proceedings of the IEEE International Workshop on Robot and Human Communication* (pp. 18-31), Tokyo, Japan.

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior, 28*, 117-139.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion, 12*, 1085-1101.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception, 24*, 433-454.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London, UK: John Murray.

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior, 13*, 247-268.

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music, 39*, 18-49.

Ekman, P. (Ed.) (1982). *Emotions in the human face.* Cambridge, UK: Cambridge University Press.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29*, 288-298.

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*, 75-99.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, 6*, 123-147.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music, 24*, 68-91.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and Research* (pp. 223-248). New York: Oxford University Press.

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception, 23*, 319-330.

Janata, P., Tomic, S. T., & Haberman, J. M. (2011). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*, 54-75.

Keller, P., & Rieger, M. (2009). Special issue: Musical movement and synchronization. *Music Perception, 26*, 397-400.

Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science, 11*, 45-50.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* New York: Basic Books.

Leman, M. (2007). *Embodied music cognition and mediation technology.* Cambridge, MA, London, UK: MIT Press.

Leman, M., & Godøy, R. I. (2010). Why study musical gesture? In R. I. Godøy & M. Leman (Eds.), *Musical gestures. Sound, movement, and meaning* (pp. 3-11). New York: Routledge.

Lesaffre, M., De Voogdt, L., Leman, M., De Baets, B., De Meyer, H., & Martens, J.-P. (2008). How potential users of music search and retrieval systems describe the semantic quality of music. *Journal of the American Society for Information Science and Technology, 59*, 695-707.

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality, 44*, 714-720.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178.

Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music, 39*, 307-327.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*, 1627-1639.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61*, 81-88.

Schubert, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology, 51*, 154-165.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Toiviainen, P., & Burger, B. (2011). *MoCap toolbox manual.* Jyväskylä, Finland: University of Jyväskylä. Available at http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual.

Van Dyck, E., Moelants, D., Demey, M., Coussement, P., Deweppe, A., & Leman, M. (2010). The impact of the bass drum on body movement in spontaneous dance. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th International Conference in Music Perception and Cognition (ICMPC)* (pp. 429-434), Seattle, WA: University of Washington.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience.* Cambridge, MA: MIT Press.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28*, 879-896.

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research, 34*, 97-113.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*, 494-521.

Appendix A: List of Stimuli

Alice Deejay: Better Off Alone (Who Needs Guitars Anyway?) 2:40-2:54 (loop)

Andre Visior: Speed Up 1:15-1:45

Antibalas: Who is this America Dem Speak of Today? (Who Is This America?) 1:00-1:29

Arturo Sandoval: A Mis Abuelos (Danzon) 1:53-2:21

Baden Powell: Deixa (Personalidade) 1:11-1:41

Brad Mehldau: Wave/Mother Nature's Son (Largo) 0:00-0:29

Clifford Brown & Max Roach: The Blues walk (Verve Jazz Masters, Vol. 44: Clifford Brown & Max Roach) 2:01-2:31

Conjunto Imagen: Medley-Esencia de Guaguanco/ Sonero (Ayer, Hoy y Manana) 2:18-2:48

Dave Hillyard & The Rocksteady 7: Hillyard Street (Playtime) 0:15-0:45

Dave Weckl: Mercy, Mercy, Mercy (Burning for Buddy) 0:10-0:40

Dave Weckl: Tower of Inspiration (Master Plan) 0:00-0:30

DJ Shadow: Napalm Brain/Scatter Brain (Endtroducing . . .) 3:29-3:58

Gangster Politics: Gangster Politics (Guns & Chicks) 1:00-1:29

Gigi D'Agostino: Blablabla (L'Amour Toujours) 0:00-0:30

Herbie Hancock: Watermelon man (Cantaloupe Island) 0:00-0.30

Horace Silver: The Natives Are Restless 0:00-0:29

In Flames: Scream (Come Clarity) 0:00-0:30

Jean Roch: Can You Feel it (Club Sounds Vol. 35) 0:33-1:01

Johanna Kurkela: Hetki hiljaa (Hetki hiljaa) 3:22-3:52

Juana Molina: Tres cosas (Tres Cosas) 0:00-0:30

Kings of Leon: Closer (Only by the Night) 3:17-3:47

Lenny Kravitz: Live (5) 3:02-3:30

Martha & The Vandellas: Heat Wave (Heat Wave) 1:40-2:10

Maynard Ferguson: Fireshaker (Live From San Francisco) 0:00-0:28

MIA: 20 Dollar (Kala) 0:17-0:45

Nick Beat: Techno Disco 2:26-2:56

Panjabi MC: Mundian To Bach Ke (Legalised) 0:47-1:06 (loop)

Patrick Watson: Beijing (Wooden Arms) 2:30-2:59

The Rippingtons: Weekend in Monaco (Weekend in Monaco) 1:13-1:42

Yuri Buenaventura: Salsa (Salsa Movie Soundtrack) 2:17-2:45

Appendix B: Derivation of Circular Affect

In what follows, the derivation of the angular variable Circular Affect is explained in mathematical terms.

The combination of the two ratings for Valence and Arousal for any stimulus can be represented as an emotion vector in a two-dimensional space, e.g., $\mathbf{r} = (v, a)$. In order to obtain a value at any possible polar angle θ, this vector $\mathbf{r}$ was orthogonally projected on the unit vector $\mathbf{e}$ with the respective polar angle,

$$\mathbf{e}_\theta = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta, \theta \in [0, 2\pi],$$

where $\mathbf{i}$ and $\mathbf{j}$ denote the unit vectors parallel to the horizontal and vertical axes, respectively. The vector projection itself was then accomplished by the dot product:

$$C_\theta = \mathbf{r} \cdot \mathbf{e}_\theta.$$

The general principle for this projection is displayed in Figure 5.

Subsequently, Circular Affect (C_θ) was correlated with each of the movement features M_i (respectively each basic emotion rating), resulting in a series of correlation values. In the last step, the maximal correlation value and the corresponding angle θ were obtained from this series.

$$\arg\max_\theta = \mathrm{corr}(C_\theta, M_i).$$

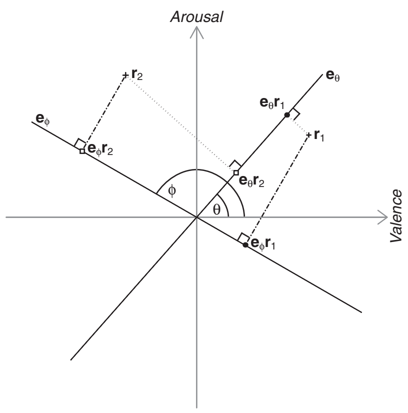


FIGURE 5. Derivation of the dimensional variable Circular Affect. r_1 is an arbitrary rating of Valence and Arousal for one stimulus, whereas r_2 is an arbitrary rating of Valence and Arousal for another stimulus. e_θ and e_ϕ are two arbitrary unit vectors at the polar angles θ and ϕ. The vectors r_1 and r_2 are projected onto e_θ and e_ϕ and result in the values $e_\theta \cdot r_1$ and $e_\phi \cdot r_1$ of the Circular Affect variable for stimulus 1, and respectively in the values $e_\theta \cdot r_2$ and $e_\phi \cdot r_2$ for stimulus 2.

IV

OH HAPPY DANCE: EMOTION RECOGNITION IN DANCE MOVEMENTS

by

Birgitta Burger, Marc R. Thompson, Suvi Saarikallio, Geoff Luck & Petri Toiviainen, 2013

In G. Luck & O. Brabant (Eds.), Proceedings of the 3rd International Conference on Music & Emotion, University of Jyväskylä, Jyväskylä, Finland

Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland, 11th - 15th June 2013. Geoff Luck & Olivier Brabant (Eds.)

OH HAPPY DANCE: EMOTION RECOGNITION IN DANCE MOVEMENTS

Birgitta Burger, Marc R. Thompson, Suvi Saarikallio, Geoff Luck, Petri Toiviainen

Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Finland
birgitta.burger@jyu.fi

Abstract

Movements are capable of conveying emotions, as shown for instance in studies on both non-verbal gestures and music-specific movements performed by instrumentalists or professional dancers. Since dancing/moving to music is a common human activity, this study aims at investigating whether quasi-spontaneous music-induced movements of non-professional dancers can convey emotional qualities as well. From a movement data pool of 60 individuals dancing to 30 musical stimuli, the performances of four dancers that moved most notably, and four stimuli representing happiness, anger, sadness, and tenderness were chosen to create a set of stimuli containing the four audio excerpts, 16 video excerpts (without audio), and 64 audio-video excerpts (16 congruent music-movement combination and 48 incongruent combinations). Subsequently, 80 participants were asked to rate the emotional content perceived in the excerpts according to happiness, anger, sadness, and tenderness. The results showed that target emotions could be perceived in all conditions, although systematic mismatches occurred, especially with examples related to tenderness. The audio-only condition was most effective in conveying emotions, followed by the audio-video condition. Furthermore in the audio-video condition, the auditory modality dominated the visual modality, though the two modalities appeared additive and self-similar.

Keywords: music-induced movement, emotion, perception

1. Introduction

On a daily basis, humans use body movements as an important means of nonverbal communication. Body postures and movements can convey different kinds of information, for instance related to the mental or physical state or personality traits, or to accompany and emphasize speech. It has been argued that speech and movement/gestures are tightly connected and co-occur, as they underlie the same cognitive processes (Iverson & Thelen, 1999; McNeill, 1992). Furthermore, body movements can convey emotions. Various studies have investigated the capability of movement to express emotions and have shown that distinct features of human movement are related to emotion categories. De Meijer (1989), for example, asked observers to attribute emotional characteristics to movements of actors who performed several movement patterns differing in general features, such as trunk or arm movement, velocity, and spatial direction and found that different emotion categories were associated with different movement characteristics. Wallbott (1998) conducted a study in which he used a scenario-based approach with professional actors performing certain emotions. He found movement features characteristic for different emotion categories and computed a discriminant analysis that could, significantly above chance level, classify the emotions correctly. Atkinson, Dittrich, Gemmell, and Young (2004) compared static vs. dynamic whole body expressions and full-light vs. point-light displays and found that all of them could communicate emotions, though the recognition rate and misclassification patterns differed for individual emotions. Pollick, Paterson, Bruderlin, and Sanford (2001) investigated visual perception of emotions in simple arm movements shown as point-light displays, such as drinking and knocking. They found that arm movements could communicate emotions, though observers tended to confuse similar emotions. Gross, Crane, and Fredrickson (2010) studied the perception of knocking movements performed by actors who were subjected to an emotion induction task. The results indicated a limited recognition rate, especially for positive emotions, and systematic confusions between several emotions. Besides in acted emotion approaches, emotion recognition has been studied in other contexts as well, such as in gait. Montepare, Goldstein, and Clausen (1987), for instance, showed that happiness, anger, and sadness could be successfully recognized in walking patterns. Research in linguistics has investigated integration of auditory and visual information in emotion perception using face-voice stimuli. Such studies showed that usually the visual information dominates (Collignon et al., 2008), and that bimodal stimuli can be integrated even if they display incongruent combinations of emotions (De Gelder & Vroomen, 2000; Massaro & Egan, 1996).

Emotions are an essential component of musical expression (e.g., Gabrielsson & Lindström, 2010) and have been investigated in a large number of music-related studies. According to Krumhansl (2002), people report that their primary motivation for listening to music is its emotional impact. Various rating experiments have shown that listeners are able to perceive emotional content in music in a consistent fashion (e.g., Balkwill & Thompson, 1999; Eerola & Vuoskoski, 2011; Gabrielsson & Juslin, 1996; Schubert, 1999; Zentner, Grandjean, & Scherer, 2008).

Musical emotions are conveyed not only by the music itself, but also through movement. While movements are required, for example, to produce sounds when playing a musical instrument, studies have shown that there are certain additional movements that are not used for the actual sound production, but for conveying emotions and expressivity (e.g., Wanderley, Vines, Middleton, McKay, & Hatch, 2005). Davidson (1993) conducted a study in which observers rated expressive movements of violinists and pianists. Results indicate that visual information more clearly communicated the expressive manner of the musician than sound alone and sound and vision presented together. Vines, Krumhansl, Wanderley, and Levitin (2006) examined clarinetists' abilities of communicating tension to observers and found that auditory and visual signals evoked different perceptions of tension, with the sound dominating the judgments. Additionally, their results indicated that the audiovisual presentation increased the perceived tension compared to the audio-only and video-only conditions, suggesting that participants integrated both signals. Dahl and Friberg (2007) studied marimba, saxophone, and bassoon players performing with different emotional intentions and presented observers with only the visual elements of the performances. Observers could detect the happy, angry, and sad performances successfully, but failed with the fearful ones. Sörgjerd (2000) investigated a clarinetist and a violinist who performed a piece of music with different emotional intentions, and reported that happiness, anger, sadness, and fear were better identified than tenderness and solemnity. However, no significant differences for the presentation condition (audio-only, movement-only, audio+movement) were found. Petrini, McAleer, and Pollick (2010) found that in audiovisual presentations of musicians playing with different emotional characteristics the sound dominated the visual signal, both in emotionally matching and mismatching stimuli. Furthermore, when participants were asked to focus on the visual information of the audiovisual stimuli, their ability to correctly identify the emotion decreased in case of emotionally incongruent stimuli, whereas it was unaffected by the video information, when participants were asked to focus on the audio information.

More direct links between music and emotion-specific movement have been investigated in research on dance, in which movement is the only way to convey expressivity and emotion. Several studies showed that dance movement could successfully communicate emotions to observers, both in regular video and in stick-figure animations (Boone & Cunningham, 1998; Dittrich, Troscianko, Lea, & Morgan, 1996; Lagerlöf & Djerf, 2009; Walk & Homan, 1984). Common to these studies was that they used professional dancers (or actors) who were explicitly asked to express the emotions while dancing.

Listening to music make people move spontaneously, for example by rhythmically synchronizing with the pulse of the music by tapping the foot, nodding the head, moving the whole body in various manners, or mimicking instrumentalists' gestures (Leman & Godøy, 2010; Leman, 2007). Studies investigating music-induced movement have suggested such movements to be related to personality (Luck, Saarikallio, Burger, Thompson, & Toiviainen, 2010), mood (Saarikallio, Luck, Burger, Thompson, & Toiviainen, 2013), or musical features (Burger, Thompson, Saarikallio, Luck, & Toiviainen, 2013; van Dyck et al., 2013). In a recent study (Burger, Saarikallio, Luck, Thompson, & Toiviainen, 2013), we could also establish links between movements and the emotional content of the music to which the participants were dancing. In that experiment, we asked 60 participants to move to different pop music stimuli and recorded their movements with an optical motion capture system. The computational analysis of the movement data coupled with a perceptual evaluation of the emotions expressed in the music revealed characteristic movement features for the emotions happiness, anger, sadness, and tenderness.

While there appears to be correlations between perceived emotions in music and movement characteristics, it is neither clear whether music-induced movement can convey emotional content to observers and which emotions can be communicated, nor how auditory and visual information would interact in this process. Therefore, we designed a perceptual experiment using a subset of the movement data collected in the previous experiment and asked observers to rate various stick figure clips regarding the emotions conveyed by these clips. We restricted this study to the use of perceived emotions as opposed to felt emotions, since previous literature has mostly focused on the former (about the importance to distinguish between perceived and felt emotions, see Evans & Schubert, 2008; Gabrielsson, 2002), and we used perceived emotions in the previous study. We decided to include three presentation conditions: audio-only, video-only, and audio and video combined, as was done in studies about musician's movements (Davidson, 1993; Petrini et al., 2010; Sörgjerd, 2000; Vines et al., 2006) and in research on face-voice stimuli (Collignon et al., 2008; De Gelder & Vroomen, 2000; Massaro & Egan, 1996). Besides including the correct combinations of music and movement (i.e., the movements actually performed to that music during the previous experiment), we wanted to further examine the audiovisual integration and generated a set of incongruent stimuli (i.e., combining the movements with another song used in the experiment expressing a different emotion), as was done in Petrini et al. (2010), for instance. Previous studies suggest different scenarios regarding the influence of music and movement on the judgments. However, in line with the results by Petrini et al. (2010) and Vines et al. (2006), we hypothesize that the audio-video condition will receive higher recognition rates than the two unimodal conditions (at least for the congruent stimuli), and audio will dominate the perception of both the congruent and the incongruent audio-video examples.

2. Method

2.1. Participants

Eighty university students, aged 19-36, participated in this study (53 females, average age: 24.7, SD of age: 3.4). Fifty-two participants took some kind of dance lessons and 75 participants reported to like movement-related activities, such as dance and sports. Thirty-seven participants reported to go out dancing more than once a month.

2.2. Stimuli

The stimuli used in this experiment were selected from a motion capture data pool of 60 participants dancing to 30 different musical excerpts that was collected in a previous study (Burger, Saarikallio, et al., 2013). Based on results of a rating experiment conducted within this previous study, four musical excerpts were determined that most clearly conveyed one of the following (target) emotions happiness, anger, sadness, and tenderness. Subsequently, two female and two male participants (from now referred to as “dancers”), were chosen based on their movement characteristics (i.e., participants whose performances received highest values regarding several movement features, such as speed, acceleration, area covered, or complexity). This yielded 16 (4x4) combinations of stick figure dance performances.

The four musical stimuli were of different tempi (100 bpm, 105 bpm, 113 bpm, 121 bpm), so we adjusted their tempi to 113 bpm by time-stretching three of them using the Audacity software[1], to eliminate the effect of tempo on the perceived emotions. Likewise, the movement data was resampled by the appropriate ratio using the Matlab Mocap Toolbox (Toiviainen & Burger, 2013). Besides this, the movement data were rotated to be visible from the front with respect to the average locations of the hip markers (for more information on the marker locations see Burger et al., 2013). In all videos, the stick figures were plotted in black on white background. QuickTime Player 7 (Pro version) was used to combine the time-shifted audio and video material in all possible combinations yielding 16 congruent stimuli and 48 incongruent stimuli. It was checked that the combination appeared synchronized. All stimuli used in the experiment were trimmed to 20 seconds.

2.3. Apparatus

To gather perceptual ratings, a special patch was created in Max/MSP 5, a graphical programming environment, running on Max OS X. The patch used QuickTime Player 7 to play back the video material. The setup enabled the participants to repeat excerpts as often as they wished, to move forward at their own speed, and to take breaks at any moment of the experiment. The stimuli were played back through studio quality headphones (AKG K141 Studio). The participants could themselves adjust the volume to a preferred level.

2.4. Procedure

In the beginning of the experiment, a short questionnaire was filled in to gather information about participants’ gender, age, dance training, and movement and dance activities. The experiment was divided into three sections: one section containing the four (time-stretched) audio clips, a second section containing the 16 silent (and time-shifted) video clips (four dancers moving to the four different musical stimuli), and a third section containing the 64 (time-shifted) audio-video clips (16 congruent and 48 incongruent combinations). To avoid any effect of order, the three sections were presented in random order to the participants. Within each section, the clips were randomized as well. Participants were accomplishing the experiment individually. They were instructed to rate the emotions expressed in the clips (perceived emotions) on seven-step scales for Happiness, Anger, Sadness, and Tenderness. Preceding each section, there was a practice part of one example to get familiar with the interface, the type of stimuli, and the rating scales used. In the beginning of the experiment, participants were explicitly told to rate according to the emotions expressed in the clips (as opposed to felt emotions). The participants were also advised to take breaks in between if they felt like it. The total duration of the experiment was between 45 and 90 minutes. After completing the experiment participants were rewarded with a movie ticket.

3. Results

Outlier detection was performed as the first step of the analysis by calculating the mean inter-subject correlation for each participant (taking all ratings of the three conditions) and each rating scale (happiness, anger, sadness, and tenderness) separately. This yielded a measure of how similarly the participants were rating in relation to each other on each scale.

[1] http://www.audacity.sourceforge.net

We obtained overall positive correlations, apart from one participant who correlated negatively on two scales (happiness and tenderness) with the other participants. We therefore decided to eliminate this participant from further analysis.

Next, we checked for rating consistency between participants by calculating intraclass correlations (cf., Shrout & Fleiss, 1979) for each rating scale separately. All six correlations coefficients were highly significant (between $r = .95$ and $r = .97$, $p < .001$), suggesting that participants' ratings for each scale were similar enough to average across them for any further analysis.

Subsequently, this section will present the analysis regarding which emotions were perceived in the different conditions, which condition was most effective in conveying the target emotion, and whether the perception of the audio-video clips was higher influenced by the auditory or by the visual modality of the clips.

3.1. Emotions perceived

In order to investigate which emotions were perceived in the different conditions and if the target emotions could be perceived, we displayed the ratings of the participants for the four musical stimuli (target emotion) as error bar plots (see Figure 1) and additionally conducted repeated measures ANOVAs (including subsequent post hoc tests) with the four ratings scales per stimulus to assess the significance of the differences between the ratings. The significance level of the post hoc tests (as indicated in Figure 1) was adjusted using Bonferroni correction to account for multiple comparisons.

Figure 1A shows the results for the four stimuli / target emotions of the audio condition. For the stimulus with Happiness as target emotion, the happiness scale received the highest average rating of the four scales/concepts, and the repeated measures ANOVA showed a significant main effect, $F(2.34, 182.67) = 194.04$, $p < .001$ (degrees of freedom adjusted using Greenhouse-Geissner correction as the sphericity assumption was violated – applied to all repeated measures ANOVA results, if not indicated differently). The subsequent post hoc test revealed a highly significant difference ($p < .001$) between happiness and tenderness, the second highest average, thus the target emotion Happiness

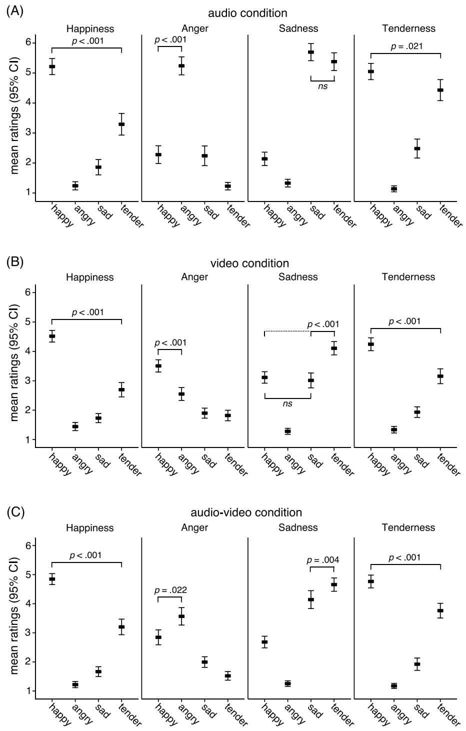


Figure 1. Error bar plots displaying average ratings and 95% confidence intervals for the four target emotion stimuli of the three conditions. (A) Audio condition. (B) Video condition, averaged across dancers. (C) Congruent audio-video condition, averaged across dancers.

could be successfully perceived in the audio condition. For the stimulus with the target emotion Anger, anger received the highest average rating of the four scales/concepts, the repeated measures ANOVA resulted in a significant main effect, $F(2.07, 160.47) = 156.74$, $p < .001$, and the post hoc test showed a significant difference ($p < .001$) between anger and happiness, the second highest average. Thus, the target emotion Anger could be successfully communicated as well. For the stimulus with Sadness as target emotion, sadness obtained highest ratings with a significant main effect revealed by the repeated measures ANOVA, $F(2.35, 183.27) = 310.84$, $p < .001$. The post hoc test, however, indicated that the difference between sadness and the second highest average, tenderness, was non-significant ($p = .68$), so there two emotions were slightly confused in case of the target emotion Sadness. For the stimulus with the target emotion Tenderness, happiness received the highest average rating of the four concepts, followed by tenderness. The repeated measures ANOVA exhibited a significant main effect, $F(2.40, 187.29) = 171.86$, $p < .001$, and the post hoc test resulted in a significant difference between both emotion concepts ($p = .021$), so participants tended to confuse Tenderness with Happiness.

Figure 1B shows the results for the video condition. The ratings for the 16 clips presented in the experiment were averaged across the four dancers to obtain one rating per participant for each target emotion stimulus. For the target emotion Happiness, the happiness scale received the highest average rating of the four concepts, and the repeated measures ANOVA showed a significant main effect, $F(2.36, 183.77) = 254.10$, $p < .001$. The subsequent post hoc test revealed a highly significant difference between happiness and the second highest average, tenderness, ($p < .001$), thus the target emotion Happiness could be communicated successfully in the video condition as well. For the target emotion Anger, happiness received the highest average rating of the four concepts, followed by anger. The repeated measures ANOVA indicated a significant main effect, $F(2.21, 172.36) = 73.19$, $p < .001$, and the post hoc test resulted in a significant difference between both concepts ($p < .001$), so the movements performed to the stimulus rated as angry could not efficiently communicate anger. For the target emotion Sadness, tenderness received the highest average rating, followed by happiness and sadness. The repeated measures ANOVA revealed a significant main effect, $F(2.33, 181.40) = 147.69$, $p < .001$, and the post hoc comparison between sadness and tenderness as well as between tenderness and happiness exhibited significant differences ($p < .001$), whereas the difference between sadness and happiness was shown to be non-significant ($p = 1.00$). From this it can be concluded that the movements failed to communicate the intended emotion. For the target emotion Tenderness, happiness obtained the highest averaged rating, followed by tenderness. The repeated measures ANOVA resulted in a significant main effect, $F(2.33, 181.63) = 198.98$, $p < .001$, while the post hoc test indicated a significant difference between both concepts ($p < .001$), so confusion alike to the audio condition occurred in the video condition as well.

Figure 1C displays the results for the congruent stimuli of audio-video condition. The ratings from the 16 congruent clips presented in the experiment were averaged across the four dancers to obtain one rating per participant for each target emotion stimulus. For the target emotion Happiness, happiness received the highest average rating of the four concepts, and the repeated measures ANOVA exhibited a significant main effect, $F(2.38, 185.44) = 359.41$, $p < .001$. The subsequent post hoc test revealed a significant difference between happiness and tenderness, the second highest average, ($p < .001$), thus the target emotion Happiness could be communicated successfully in the audio-video condition. For the target emotion Anger, anger received the highest average rating of the four concepts. The repeated measures ANOVA resulted in a significant main effect, $F(1.67, 130.41) = 60.76$, $p < .001$, and the post hoc comparison between the two concepts showed a (moderately) significant difference ($p = .022$), so the combination of auditory and visual components could fairly successfully communicate the target emotion Anger. For the target emotion Sadness, tenderness received the highest average rating, followed by sadness. The repeated measures ANOVA showed a significant main effect, $F(1.88, 146.30) = 202.75$, $p < .001$, and the post hoc test exhibited a significant difference between both concepts ($p = .004$), which suggests that there occurred some confusion between the two. For the target emotion Tenderness, happiness obtained the highest average rating, followed by tenderness. The repeated measures ANOVA results revealed a significant main effect, $F(2.18, 169.74) = 317.93$, $p < .001$, and the post hoc comparison indicated a significant difference between the two concepts ($p < .001$), so clips intended to convey tenderness communicated happiness instead. Thus, the same confusion as in the audio and video conditions described above happened in the audio-video condition.

3.2. Conveyance of target emotion

Next, we examined which experiment condition was most effective in conveying the four target emotions. Figure 2 illustrates the differences between the three conditions in conveying the four target emotions. For all target emotions, audio was most effective (highest average ratings), followed by the audio-video condition. The video condition was least effective in conveying the target emotions, especially in the case of anger and sadness. Repeated measures ANOVA results including post hoc tests (adjusted significance level using Bonferroni correction) revealed significant main effects (see Table 1) and differences between the three conditions in all target emotions (significance of the differences indicated in Figure 2).

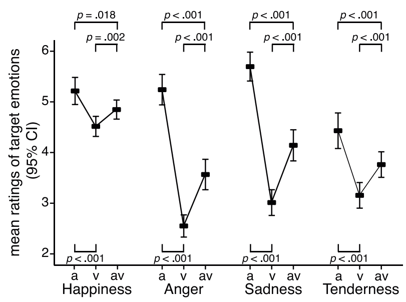


Figure 2. Error bar plot showing the differences in mean ratings for the three experiment conditions with respect to conveying the four target emotions (a: audio condition; v: video condition; av: audio-video condition).

Table 1. Repeated measure ANOVA results for each target emotion.

	F statistics	significance
Happiness	$F(1.49, 116.13) = 14.03$	$p < .001$
Anger	$F(2, 156)^{\#} = 166.63$	$p < .001$
Sadness	$F(2, 156)^{\#} = 164.94$	$p < .001$
Tenderness	$F(1.54, 119.82) = 30.99$	$p < .001$

no Greenhouse-Geissner correction applied (sphericity assumption holds)

3.3. Influence of modalities in audio-video condition

In order to investigate whether the perception of the audio-video clips was more strongly influenced by the auditory or by the visual modality, we conducted a series of linear regression analyses for each of the target emotions: 1) the audio ratings predicting the (congruent and incongruent) audio-video ratings, 2) the video ratings predicting the (congruent and incongruent) audio-video ratings, and 3) both the video and audio ratings predicting the (congruent and incongruent) audio-video ratings. The variances explained by each of the models (R^2 values) are displayed in Table 2.

Table 2. Variances explained by each of the three regression models for each of the target emotions. The asterisks indicate the significance level of the regression models.

	R^2(audio)	R^2(video)	R^2(audio + video)
Happiness	.54 ***	.39 ***	.93 ***
Anger	.82 ***	.14 **	.96 ***
Sadness	.62 ***	.33 ***	.94 ***
Tenderness	.51 ***	.39 ***	.91 ***

*** $p < .001$, ** $p < .01$

For all four target emotions, the ratings for the audio condition could explain a larger amount of variance than the video ratings. We therefore assume that the participants paid more attention to the auditory part of the stimulus than to the visual part when rating the (partly incongruent) audio-video stimuli. Furthermore, audio and video ratings were additive, as adding both R^2 values resulted in a value very close to the R^2 value of the third (audio + video) regression model. Including both audio and video ratings in the (third) regression model, between 91 and 96 % of the total variance could be explained, which means that audio-video ratings could almost perfectly be predicted from both the ratings of the audio and the video conditions.

The final step of the analysis consisted of investigating the distribution and alignment of the ratings of the audio-video condition. To do so, we first reduced the dimensionality of the rating data for the audio-video condition using principal components analysis (PCA). Applying Kaiser's criterion, we retained two components that could account for 77.8% of the total variance (PC 1: 41.25% and PC 2: 36.55%). Subsequently, the components were rotated using varimax rotation and then averaged across dancers and participants yielding one value per audio/video combination resulting in 16 values per PC. Figure 3 displays these results as coordinates in a two-dimensional space. To show the influences of the auditory and visual modality on the audio-video ratings, the figure contains two subfigures: subfigure (A) shows the solution with the same audio stimuli being connected with lines, whereas subfigure (B) shows the same solution, but now with the same movement stimuli connected. The dotted line in each subfigure connects the congruent stimuli. Figure 3 shows that the stimuli having the same audio grouped closer together than the stimuli having the same movements. The happiness and tenderness cluster are overlapping in the same-audio-connected case (Figure 3A), which is in line with the previous results related to the confusion of the two emotion concepts. If we consider the two concepts as one cluster, we can see that there is no overlap between the happiness/tenderness cluster, the anger cluster, and the sadness cluster. In the same-movement case (Figure 3B) however, all clusters overlap, thus it seemed that the auditory domain was indeed dominating the visual domain in the audio-video condition. Furthermore, it is interesting to note that the clusters have self-similar structures with themselves and with the connectors of the congruent stimuli (dotted lines). This suggests that the ratings for the audio-video condition were very consistent and coherent, despite the contradictory information contained in the stimuli. This also accords with the additivity of the audio-only and video-only ratings in the regression models.

4. Discussion

This study investigated whether quasi-spontaneous music-induced movements of non-professional dancers can convey emotional qualities. Participants were asked to rate audio, video, and audio-video clips according to the emotions expressed by music, movement, and combinations of both. The results showed that in general the target emotions could be perceived in the three experiment conditions, although systematic mismatches occurred, especially examples related to ten-

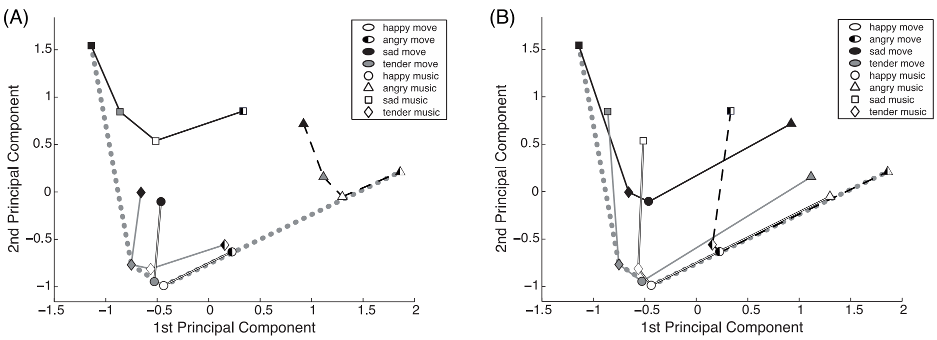


Figure 3. Principal component solution averaged across dancers and participants plotted against each other in two dimensions. (A) Solution with the same audio stimuli being connected with lines. (B) Solution with the same movement stimuli being connected. The dotted line in each subfigure is connecting the congruent stimuli.

derness. Such confusions have been reported in the literature, both in music- and in movement-related research (e.g., Eerola & Vuoskoski, 2011; Gross et al., 2010; Pollick et al., 2001). Eerola and Vuoskoski (2011) reported that for the moderate examples used in their study on emotions perception in music happiness and tenderness as well as tenderness and sadness were confused, while Gross et al. (2010) found in investigating movements of emotional knocking that tenderness was confused with happiness. These findings suggest that it requires very stereotypical examples to communicate emotions, and more specifically to communicate a single emotion with one stimulus. The nature of the stimuli could serve as an explanation for the confusions related to tenderness. The prerequisite for our musical stimuli was to contain some kind of beat to induce and stimulate movement. However, it is difficult (if not impossible) to find musical stimuli that genuinely express tenderness and as well have a perceivable beat structure. Stimuli with a tender character, but possessing a beat are most likely to be rated as happy, as the activity/arousal level would increase due to the beat. Consequently, if the music communicated similar emotional qualities, it is likely that the movement exhibited to such music do so as well.

For the video condition, an interesting finding is that the movements performed to the angry stimulus could not communicate the same emotion. An explanation for this finding could be that dance movements are commonly understood as positive and pleasant (leisure activity, fun), so they might appear and thus be rated (more) positively as well. This explanation would connect the discrete emotions approach with the dimensional model of emotions (Russell, 1980), in which happiness is commonly described as active positive/pleasant, anger as active negative, sadness as inactive negative, and tenderness as inactive positive, so the shift from anger to happiness could be explained in terms of the valence/pleasantness dimension of that model. Additionally, the movements might contain less stereotypical hints to emotional qualities than the music, at least to negative emotions. Thus, it could be that negative emotions can only be communicated in dance movements, if the dancers are asked to move according to the emotional character of the music (and not "just" to the music, as it was in this case). Similar results were also found in studies on dance and on acted emotions (Atkinson et al., 2004; Dittrich et al., 1996), although they would disagree with Walk and Homan's (1984) 'alarm hypothesis': expressions of negative emotions are easier and better recognized than positive emotions. The tendency towards more positive judgments (related to the dimensional model of emotions) can also be found in the case of Sadness as target emotion. The sad examples were rated being high in expressing tenderness, so the observers rated them more positive and pleasant, like they did with the examples expressing anger. A similar shift might have occurred in the ratings for the target emotion Tenderness when being confused with happiness, though it would be rather related to the activity level than to pleasantness in terms of the dimensional model of emotions.

We observed similar shifts in the congruent audio-video condition for the Sadness and Tenderness as target emotions – sad examples being perceived as more positive (cf., dimensional model) and tender examples being perceived as more active – though the results for the target emotion anger were different: anger could be successfully communicated when audio and video was presented together. Thus, it seems that the auditory domain dominates such ratings compared to the visual domain.

A closer investigation of the question which condition was most effective in conveying emotions revealed that the audio condition received the highest average ratings for the four target emotions, followed by the audio-video condition. For all target emotions, the same pattern emerged. Thus, we failed to find support to our initial hypothesis that the audio-video condition would receive highest recognition. We could neither confirm the results obtained in previous research (Collignon et al., 2008; Davidson, 1993; Petrini et al., 2010; Sörgjerd, 2000; Vines et al., 2006). It could be that music is the most powerful carrier of emotions, at least in relation to spontaneous dance movements. As mentioned already, dance movement might rather express pleasure, sensuality, and aesthetics, so they might lack stereotypical and stylized emotion-related qualities.

Earlier studies usually employed professional actors or dancers, whose task was to portray specific emotions. Such an approach leads to rather stylized movements that might successfully express emotions, but may lack naturalism. Our stimuli, on the other hand, were derived from quasi-spontaneous movements to music, so they might lack emotion-specific qualities, as the dancers were not instructed to express emotions with their movements. However, although the dancers were unaware of any implications to emotions, they still moved in a way that observers perceived as expressive of emotions.

The three experiment conditions showed a noteworthy relationship, as the auditory and visual modalities appeared to be additive with regards to the audio-video presentation of the stimuli. The audio-video ratings could be almost perfectly predicted from both the ratings for the audio condition and the video condition, whereas the ratings of either the audio condition or the video condition could only partly predict the audio-video ratings. This would suggest that both domains influence the perception of audio-video stimuli, and that participants tried to integrate both channels, as has already been shown in research on music (Petrini et al., 2010; Vines et al., 2006) and on voice-face integration (De Gelder & Vroomen, 2000; Massaro & Egan, 1996).

Furthermore, we found that in the audio-video condition, the auditory modality dominated the visual modality. Thus, regardless of the combination of the stimuli (i.e., for both congruent and incongruent combinations of music and movements), participants seemed to rather rate according to the music than to the movement. This result supports our initial hypothesis as well as previous studies in music research (Petrini et al., 2010; Vines et al., 2006), but contradicts with results from research on voice-face stimuli (Collignon et al., 2008) – a finding that would suggest that differences exist in musical and linguistic processing with regards to emotion perception. The domination of the auditory domain could mean that the music was more expressive of a certain emotion than was the movement, so it was guiding the ratings of the participants – especially in cases of incongruent combinations. This issue will be investigated further in the future, as data was gathered about the participants' own perception of their rating behavior: the participants were asked after completion of the audio-video part to indicate whether they paid more attention to the music, to the movement, or if they tried to integrate both. These results might give us more insight into the participants' integration of auditory and visual stimulus information.

An interesting finding within the results of the audio-video condition is that the alignment of the ratings showed a self-similar structure. We plotted the principal component solution of the basic emotion ratings in a two-dimensional space and connected the stimuli in two ways: related to the examples containing the same music and related to the examples containing the same movements. The resulting clusters supported the already mentioned finding about the auditory domain having higher influence on the ratings, as the same-music clusters are non-overlapping and smaller than the same-movement clusters. It seems that the music or/and the congruent stimulus combination "attracted" the other movements to cluster close together. Furthermore, the clusters exhibited self-similar structures with each other and with the cluster of the congruent stimuli. This suggests that, despite the contradictory information of the stimuli, the ratings were very consistent and coherent. It could be assumed that both the congruent stimulus combinations and the music were somehow guiding the perception of the incongruent stimulus combinations, so the participants followed a homogeneous rating scheme throughout the experiment.

The result that the movements failed to communicate anger in the video condition would deserve further investigation. In a future experiment, participants could be asked – differently from the present study – to express, while dancing, the emotion expressed in the music. It would be of interest to investigate how the movements (and their perception) change and if participants are able to maintain dance movement characteristics while successfully communicating emotions.

The selection of the musical stimuli in this experiment was somewhat restricted, insofar as they were derived/adopted from a previous study. Therefore, the stimuli might not have as clearly expressed an emotion as stimuli selected purely based on their emotional characteristics. However, a follow-up study could address this issue with a fresh set of stimuli of more clearly expressing an emotion, though being potentially less danceable than the recent selection.

The presented study gave revealing insights into multimodal perception of dance movement stimuli showing that participants could attribute emotions to dance movements. They were furthermore able to meaningfully integrate audio and video signals, even in case of incongruent combinations. This investigation was a follow-up study to a motion capture study (Burger, Saarikallio, et al., 2013), in which we could establish links between movement features and the emotional characteristics of music. The results of both studies can serve as support for each other: we could show in both studies that music-induced movements can express emotional characteristics, with the computational analysis showing in particular that there are relationships between movement features and the emotional content of music, and the perceptual experiment showing that humans can perceive and recognize emotions in music-induced movements.

Acknowledgments

This study was funded by the Academy of Finland (projects 118616, 125710, 136358).

References

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, *33*(6), 717–746.

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, *17*(1), 43–64.

Boone, R. T., & Cunningham, J. G. (1998). Children's Decoding of Emotion in Expressive Body Movement: The Development of Cue Attunement. *Developmental Psychology*, *34*(5), 1007–1016.

Burger, B., Saarikallio, S., Luck, G., Thompson, M. R., & Toiviainen, P. (2013). Relationships between perceived emotions in music and music-induced movement. *Music Perception*, *in press*.

Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, *4*:183.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain research*, *1242*, 126–35.

Dahl, S., & Friberg, A. (2007). Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception*, *24*(5), 433–454.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*, 103–113.

De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*(3), 289–311.

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, *13*(4), 247–268.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*(6), 727–738.

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49.

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*(1), 75–99.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: same or different? *Musicae Scientiae*, *6*, 123–147.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. *Psychology of Music*, *24*, 68–91.

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. Sloboda (Eds.), *Handbook of*

Music and Emotion. Theory, Research, Applications (pp. 367–400). Oxford, UK: Oxford University Press.

Gross, M. M., Crane, E. a., & Fredrickson, B. L. (2010). Methodology for Assessing Bodily Expression of Emotion. *Journal of Nonverbal Behavior*, *34*(4), 223–248.

Iverson, J. M., & Thelen, E. (1999). Hand, Mouth and Brain. *Journal of Consciousness Studies*, *6*(11-12), 19–40.

Krumhansl, C. L. (2002). Music: A Link Between Cognition and Emotion. *Current Directions in Psychological Science*, *11*(2), 45–50.

Lagerlöf, I., & Djerf, M. (2009). Children's understanding of emotion in dance. *European Journal of Developmental Psychology*, *6*(4), 409–431.

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA, London, UK: MIT Press.

Leman, M., & Godøy, R. I. (2010). Why Study Musical Gesture? In R. I. Godøy & M. Leman (Eds.), *Musical Gestures. Sound, Movement, and Meaning* (pp. 3–11). New York, NY: Routledge.

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality*, *44*(6), 714–720.

Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*, *3*(2), 215–221.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, Il: University of Chicago Press.

Montepare, J. M., Goldstein, S. B., & Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behavior*, *11*(1), 33–42.

Petrini, K., McAleer, P., & Pollick, F. (2010). Audiovisual integration of emotional signals from music improvisation does not depend on temporal correspondence. *Brain research*, *1323*, 139–48.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*, B51–B61.

Russell, J. A. (1980). A Circumplex Model Of Affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Saarikallio, S., Luck, G., Burger, B., Thompson, M. R., & Toiviainen, P. (2013). Dance moves reflect current affective state illustrative of approach-avoidance motivation. *Psychology of Aesthetics, Creativity, and the Arts*, *in press*.

Schubert, E. (1999). Measuring Emotion Continuously: Validity and Reliability of the Two-Dimensional Emotion-Space. *Australian Journal of Psychology*, *51*(3), 154–165.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Sörgjerd, M. (2000). *Auditory and visual recognition of emotional expression in performances of music*. Unpublished thesis. Department of Psychology, University of Uppsala, Sweden.

Toiviainen, P., & Burger, B. (2013). *MoCap Toolbox Manual*. University of Jyväskylä: Jyväskylä, Finland. Available at http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The Impact of the Bass Drum on Human Dance Movement. *Music Perception*, *30*(4), 349–359.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*, 80–113.

Walk, R. D., & Homan, C. P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society*, *22*, 437–440.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*(6), 879–896.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, *8*(4), 494–521.

V

EFFECTS OF THE BIG FIVE AND MUSICAL GENRE ON MUSIC-INDUCED MOVEMENT

by

Geoff Luck, Suvi Saarikallio, Birgitta Burger, Marc R. Thompson & Petri Toiviainen, 2013

Journal of Research in Personality vol. 44(6), 714-720

VI

INVESTIGATING RELATIONSHIPS BETWEEN MUSIC, EMOTIONS, PERSONALITY, AND MUSIC-INDUCED MOVEMENT

by

Birgitta Burger, Juho Polet, Geoff Luck, Marc R. Thompson, Suvi Saarikallio & Petri Toiviainen, 2013

In G. Luck & O. Brabant (Eds.), Proceedings of the 3rd International Conference on Music & Emotion, University of Jyväskylä, Jyväskylä, Finland

Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland, 11th - 15th June 2013. Geoff Luck & Olivier Brabant (Eds.)

INVESTIGATING RELATIONSHIPS BETWEEN MUSIC, EMOTIONS, PERSONALITY, AND MUSIC-INDUCED MOVEMENT

Birgitta Burger[*], Juho Polet[+], Geoff Luck[*], Marc R. Thompson[*], Suvi Saarikallio[*], Petri Toiviainen[*]

[*]Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Finland
[+]Niilo Mäki Institute, Jyväskylä, Finland
birgitta.burger@jyu.fi

Abstract

Listening to music makes us to move in various ways. The characteristics of these movements can be affected by several aspects, such as individual factors, musical features, or the emotional content of the music. In a study in which we presented 60 individuals with 30 musical stimuli representing different genres of popular music and recorded their movement with an optical motion capture system, we found significant correlations 1) between musical characteristics and the exhibited movement, 2) between the perceived emotional content of the music and the movement, and 3) between personality traits of the dancers and the movement. However, such separate analyses are incapable of investigating possible relationships between the different aspects. We describe two multivariate analysis approaches – mediation and moderation – that enable the simultaneous analysis of relationships between more than two variables. The results of these analyses suggest mediation effects of the perceived emotional content of music on the relationships between different features of music and movement. It can therefore be assumed that musical emotions can (partly) account for the effect of music on movements. However, using personality as moderator between music and movement failed to show a moderation effect in most cases, suggesting that personality does not generally affect existing relationships between music and movement. Hence it can be assumed that musical characteristics and personality are independent factors in relation to music-induced movement.

Keywords: music-induced movement, multivariate statistics, emotion

1. Introduction

Music makes us move. While listening to music, we often spontaneously move our bodies. Keller and Rieger (2009), for example, stated that simply listening to music can induce movement, and Lesaffre et al. (2008) conducted a self-report study, in which most participants reported moving when listening to music. Janata, Tomic, and Haberman (2012) asked participants to tap to music and found that participants not only moved the finger/hand, but also other body parts, such as feet and head. In general, people tend to move to music in an organized way by, for example, rhythmically synchronizing with the pulse of the music by tapping their foot, nodding their head, moving their whole body, or mimicking instrumentalists' gestures (Godøy, Haga, & Jensenius, 2006; Leman & Godøy, 2010). Moreover, Leman (2007:96) suggests, "Spontaneous movements [to music] may be closely related to predictions of local bursts of energy in the musical audio stream, in particular to the beat and the rhythm patterns". Such utilization of the body is the core concept of embodied cognition, which claims that the body is involved in or even required for cognitive processes (e.g., Lakoff & Johnson, 1980, 1999; Varela, Thompson, & Rosch, 1991). Human cognition is thus highly influenced by the interaction between mind/brain, sensorimotor capabilities, body, and environment. Following this, we can approach musical engagement by linking our perception of it to our body movement (Leman, 2007). One could postulate that our bodily movements reflect, imitate, or help understanding the structure and content of music. Leman suggests that corporeal involvement could be influenced by three (co-existing) components or concepts: "Synchronization", "Embodied Attuning", and "Empathy", which differ in the degree of musical involvement and in the kind of action-perception couplings. "Synchronization" constitutes the fundamental component: synchronizing to a beat, the basic musical element. The second component, "Embodied Attuning", concerns the linkage of body movement to musical features, such as melody, harmony, rhythm, or timbre. Following this idea, movement could be used to reflect, parse, and navigate within the musical structure. Finally, "Empathy" links musical features to expressivity and emotions, so the listener would feel the emotions expressed in the music and reflect them by using body movement.

A few studies have investigated music-induced movement and suggested several factors to affect the characteristics of such movements: Relationships have been found with individual factors, such as personality (Luck, Saarikallio, Burger, Thompson, & Toiviainen, 2010) or mood (Saarikallio, Luck, Burger, Thompson, & Toiviainen, 2013), as well as with musical features, such as pulse clarity and rhythmic strength (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2012; Burger, Thompson, Saarikallio, Luck, & Toiviainen, 2013) or the presence of the bass drum (Van Dyck et al., 2013). Besides such factors, the emotional content of the music was found to be related to music-induced movement as well (Burger, Saarikallio, Luck, Thompson, & Toiviainen, 2013).

So far, these different aspects have only been studied separately. However, among them there might be relationships that separate analyses would be incapable of revealing. One variable, the emotional content of the music, for example, could account for the relationship between music and movement, in other words the music would have an indirect relationship with the movement by influencing the emotional expression, which then influences the movement. In this case, the emotional content would act as a so-called mediator variable (Baron & Kenny, 1986; Hayes, 2009; Preacher, Rucker, & Hayes, 2007). Besides or in addition to such an association, another variable, the personality of the dancer for example, could change the effect of the music on the movement. Personality would then act as a so-called moderator variable (Baron & Kenny, 1986; Preacher et al., 2007).

To explore possible relationships between musical features, perceived emotions in music, personality, and music-induced movement, we conducted a motion capture study to connect the different aspects chosen earlier (separate analyses described in Burger, Saarikallio, et al., 2013; Burger, Thompson, et al., 2013; Luck et al., 2010). Therefore, we reviewed recent statistical approaches in the field of multivariate statistics and found mediation and moderation analysis an interesting and potentially useful approach for our inquiry. This paper will first describe the design of the study and then focus on the proposed analysis methods.

2. Method

2.1. Participants

A total of 60 participants took part in this experiment (43 females; average age: 24, SD: 3.3). They were recruited based on a database of 952 individuals containing their scores of the Big Five Inventory (John, Naumann, & Soto, 2008), a 44-item instrument measuring the five primary personality dimensions (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). We aimed at high- and low-scoring individuals for each of the five dimensions to participate in this study. Participants were rewarded with a movie ticket.

2.2. Stimuli

Participants were presented with 30 randomly ordered musical stimuli representing the following popular music genres: Techno, Pop, Rock, Latin, Funk, and Jazz. All stimuli were 30 seconds long, non-vocal, and in 4/4 time, but differed in their rhythmic complexity, pulse clarity, and tempo, and in their perceived emotional content. The emotional content was assessed in a perceptual experiment, in which 34 participants rated the emotions expressed in the music on scales for happiness, anger, sadness, tenderness, arousal, and valence (for a detailed description of this experiment, see Burger, Saarikallio, et al., 2013).

2.3. Apparatus

Participants' movements were recorded using an eight-camera optical motion capture system (Qualisys ProReflex) tracking, at a frame rate of 120 Hz, the three-dimensional positions of 28 reflective markers attached to each participant. The locations of the markers are shown in Figure 1a, and can be described as follows (L = left, R = right, F = front, B = back): 1: LF head; 2: RF head; 3: LB head; 4: RB head; 5: L shoulder; 6: R shoulder; 7: sternum; 8: spine (T5); 9: LF hip; 10: RF hip; 11: LB hip; 12: RB hip; 13: L elbow; 14: R elbow; 15: L wrist/radius; 16: L wrist/ulna; 17: R wrist/radius; 18: R wrist/ulna; 19: L middle finger; 20: R middle finger; 21: L knee; 22: R knee; 23: L ankle; 24: R ankle; 25: L heel; 26: R heel; 27: L big toe; 28: R big toe. The musical stimuli were played back via a pair of Genelec 8030A loudspeakers using a Max/MSP patch running on an Apple computer.

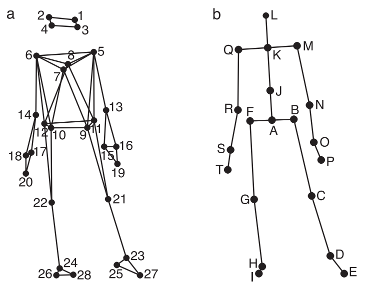


Figure 1. (a) Anterior view of the location of the markers attached to the participants' bodies; (b) Anterior view of the locations of the secondary markers/joints used in the analysis.

2.4. Procedure

Participants were recorded individually and were asked to move to the stimuli in a way that felt natural. Additionally, they were encouraged to dance if they wanted to, but were requested to remain in the center of the capture space indicated by a 115 x 200 cm carpet.

2.5. Movement feature extraction

In order to extract various kinematic features, the MATLAB Motion Capture (MoCap) Toolbox (Toiviainen & Burger, 2013) was used to first trim the data to the duration of each stimulus. Following this, a set of 20 secondary markers was derived – subsequently referred to as joints – from the original 28 markers to reduce marker redundancy. The locations of these 20 joints are depicted in Figure 1b. The locations of joints C, D, E, G, H, I, M, N, P, Q, R, and T are identical to the locations of one of the original markers, while the locations of the remaining joints were obtained by averaging the locations of two or more markers; joint A: midpoint of the four hip markers; B: midpoint of markers 9 and 11 (left hip); F: midpoint of markers 10 and 12 (right hip); J: midpoint of sternum, spine, and the hip markers (midtorso); K: midpoint of shoulder markers (manubrium), L: midpoint of the four head markers (head); O: midpoint of the two left wrist markers (left wrist); S: midpoint of the two right wrist markers (right wrist). From the three-dimensional joint position data, instantaneous velocity and acceleration were estimated using numerical differentiation based on the Savitzky-Golay smoothing FIR filter (Savitzky & Golay, 1964) with a window length of seven samples and a polynomial order of two. These values were found to provide an optimal combination of precision and smoothness in the time derivatives. Subsequently, the data was transformed into a local coordinate system, in which joint A was located at the origin, and segment BF had zero azimuth. From these data, various movement features were extracted, with six being used in the present analysis:

- Magnitude of Head Speed (Joints L).
- Magnitude of Hand Speed (Joints P & T).
- Magnitude of Head Acceleration (Joint L).
- Magnitude of Hand Acceleration (Joints P & T).
- Fluidity: overall movement fluidity (smoothness) measure based on the ratio of velocity to acceleration. The combina-

tion of high velocity and low acceleration reflects fluid movement, whereas the combination of low velocity and high acceleration reflects non-fluid movement.

- Rotation Range: amount of rotation of the body (Joints M & Q) around the vertical axis.

2.6. Musical feature extraction

In order to quantitatively describe the musical content of our stimuli, we performed computational feature extraction analysis of the stimuli used in the experiment. To this end, various musical features were extracted from the stimuli using the MATLAB MIRToolbox (version 1.4) (Lartillot & Toiviainen, 2007), with six being used in the present analysis.

- Spectral Flux of Sub-band 2: indicates the extent to which the spectrum of the frequency band between 50 and 100 Hz changes over time. This feature thus measures the flux in the low frequencies, usually produced by kick drum and bass guitar. For the calculation see Alluri and Toiviainen (2010) and Burger, Saarikallio, et al. (2013).
- Spectral Flux of Sub-band 9: based on the same calculation as spectral flux of sub-band 2, except for the frequency band ranging from 6400 to 12800 Hz. Thus, this feature measures the flux of the high frequencies, usually produced by hihats or cymbals.
- Attack Time: time of the attack at note onsets. The shorter the time, the sharper and more percussive the sound.
- Number of Onsets: sum of note onsets detected in the stimulus. A high number of onsets is related to a high sound density.
- Low Energy: the proportion of time during which the sound (RMS) energy is below the average RMS energy. Vocal music with silences, for instance, will have high values, whereas continuous strings have low values of low energy (Tzanetakis & Cook, 2002).
- Spectral Centroid: geometric center of the amplitude spectrum. It is commonly associated with musical timbre, in particular with brightness (McAdams, Winsberg, Donnadieu, Soete, & Krimphoff, 1995).

All features apart from the number of onsets resulted in time-series, subsequently being averaged to obtain one value per stimulus.

3. Results

3.1. Emotion as mediator

In Figure 2, two models of relationships between variables are depicted: the total-effect model and the mediation model. The total effect is the effect of an independent variable on a dependent variable, whereas a mediator is a variable that accounts for the effect of an independent variable on a dependent variable (Baron & Kenny, 1986; Hayes, 2009; Preacher et al., 2007). Mediation occurs if the effect of the independent variable on the dependent variable is reduced when the mediator is included (i.e., the regression coefficient is smaller for *c′* than for *c*, preferably *c′* being non-significant). Furthermore for mediation to occur, the indirect effect needs to be estimated, which can be assessed by computing confidence intervals for the indirect effect using bootstrap methods: Confidence intervals in which both lower and upper bound are either positive or negative (i.e., the confidence interval does not contain zero) are considered significant (Hayes, 2009). The effect size of the indirect effect can be assessed by calculating kappa-squared (κ^2): a small effect would be around .01, a medium effect around .09, and a large effect around .25 (Preacher & Kelley, 2011).

We hypothesized that the emotional content accounts for the relationship between music and movements, since musical features and emotional content are related with each other and both of them might have influenced participants' movements.

Although it is possible that an indirect effect exists between two variables even if these two variables are not associated when the mediator variable is absent (Hayes, 2009), we will concentrate for the time being on the combinations, in which all the three variables show at least moderately high mutual correlation with each other. Therefore, as the first step of the analysis, correlations between movement, emotions, and musical features were assessed,

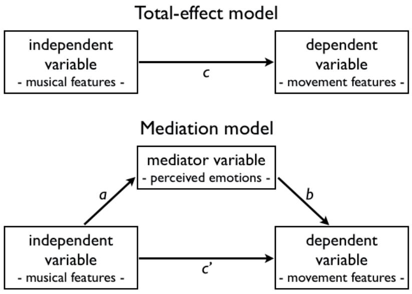


Figure 2. Total-effect model and mediation model. A mediator model decomposes the total effect, *c*, into the indirect effect, *ab* (product of the indirect paths *a* and *b*) and the direct effect, *c*' (with the effect of the mediator removed). The total effect can be describes as $c = c' + ab$, and hence the indirect effect as $ab = c - c'$.

of which the significant ones ($p < .05$) are presented in Table 1.

Table 1. Correlations (n=30) between emotions, movement features, and musical features. Asterisks indicate the significance level of the lowest coefficient of the three pairwise correlations.

		SHe	SHa	AHe	AHa	Fl	Rot
Aro	Ons	*		**	**	**	
	SBF2	**	**	**	**	**	
	SBF9	**	**	**	**		
	AT			*	**	**	
Val	LE						*
	Ons					*	
	SBF2					*	
Hap	SC						*
Ang	LE						*
	Ons					**	
	SBF2					*	
Sad	AT				*	*	
Ten	LE						*
	Ons	*		**	*	**	
	SBF2	**		**	**	*	
	SBF9	*		*	*		

* $p < .05$, ** $p < .01$

Abbreviations: Emotions: *Aro*: Arousal; *Val*: Valence; *Hap*: Happiness, *Ang*: Anger; *Sad*: Sadness; *Ten*: Tenderness – Musical features: *SBF2*: Spectral flux of sub-band no. 2; *SFB9*: Spectral flux of sub-band no.9; *AT*: Attack time; *Ons*: no of onsets; *LE*: Low Energy; *SC*: Spectral Centroid – Movement features: *SHe*: Head speed; *SHa*: Hand speed; *AHe*: Head acceleration; *AHa*: Hand acceleration; *Fl*: Fluid; *Rot*: Rotation Range

Subsequently, we performed mediation analysis using PROCESS, a tool developed for conditional process modeling in SPSS[1] (Hayes, 2013). In all analyses presented below, the significance of the indirect effect was obtained by computing 95% confidence intervals using 10 000 bootstrap samples.

Testing all significant correlations listed in Table 1 would go beyond the scope of this paper. Thus for Arousal and Tenderness, we will only report relationships involving the two Spectral Flux features, since they were found to significantly contribute to shaping music-induced movement (Burger, Thompson, et al., 2013) and neglect the musical features Low Energy, Number of Onsets, Attack Time, and Spectral Centroid. For the other emotions (Valence, Happiness, Anger, and Sadness), all feature combinations as indicated in Table 1 are analyzed.

Arousal was found to mediate the effect of Spectral Flux of Sub-band 2 on Hand Speed, Hand Acceleration, and Fluidity, and the effect of Spectral Flux of Sub-band 9 on Head Acceleration and Hand Acceleration. For all, there were significant total effects, and including the mediator caused the direct effects to become insignificant, while the indirect effects were significant, with the indirect paths being as well significant (cf., Fig. 2) The confidence intervals did not contain zero, and the κ^2-values indicated large effect sizes. Thus, Arousal could successfully mediate the relationship between these variables. The remaining combinations showed mediation effects as well, although the effect of the direct paths remained significant when including the mediator. The statistical results are presented in Table 2.

Valence was found to mediate the relationship between Low Energy and Rotation Range. However, Valence failed to mediate the effect of Spectral Flux of Sub-band 2 on Fluidity and the effect of No of Onsets on Fluidity, as the respective confidence intervals contained zero. The statistical results can be found in Table 2.

Furthermore, Happiness was found to mediate the effect of Spectral Centroid on Rotation Range (results in Table 2).

[1] www.ibm.com/software/analytics/spss/

With Anger we failed to determine any significant mediation effects, as in all three cases the respective confidence intervals contained zero (see Table 2).

Sadness also failed to act as a mediator between Attack Time and both Hand Acceleration and Fluidity, as both indirect effects were shown insignificant (see Table 2).

Table 2. Non-standardized regression coefficients and model significance of total effect *c*, direct effect *c′*, and both indirect paths *a* and *b*, as well as non-standardized and standardized regression coefficients of indirect effect *ab* with 95% bootstrap confidence interval, and effect size κ^2 with 95% bootstrap confidence interval. Insignificant direct effects are indicated in bold. Insignificant models are shaded in grey.

	model	*b(c), p*	*b(c′), p*	*b(a), p*	*b(b), p*	*b(ab)*, [CI]	*β(ab)*, [CI]	κ^2, [CI]
Aro	**SBF2/SHe**	1.05, .00	.80, .0004	.05, .003	5.11, .02	.25, [.03, .60]	.17, [.002, .04]	.22, [.05, .44]
Aro	**SBF2/SHa**	2.84, .01	1.41, **.21**	.05, .003	29.06, .02	1.42, [.28, 3.50]	.23, [.05, .53]	.23, [.05, .45]
Aro	**SBF2/AHe**	15.73, .00	9.71, .001	.05, .003	122.86, .0002	6.02, [2.49, 11.87]	.27, [.12, .47]	.33, [.15, .51]
Aro	**SBF2/AHa**	39.17, .003	13.38, **.22**	.05, .003	526.39, .0001	25.79, [10.03, 50.06]	.35, [.15, .60]	.36, [.15, .57]
Aro	**SBF2/Flu**	-.0003, .02	.00, **.79**	.05, .003	-.01, .00	-.0003, [-.0006, -.0002]	-.45, [-.66, -.22]	.50, [.28, .66]
Aro	**SBF9/SHe**	3.22, .0001	2.25, .01	.15, .01	6.42, .01	.98, [.25, 2.19]	.19, [.05, .44]	.22, [.06, .44]
Aro	**SBF9/SHa**	13.90, .0001	10.51, .003	.15, .01	22.32, .03	3.39, [.64, 8.16]	.16, [.03, .38]	.18, [.04, .38]
Aro	**SBF9/AHe**	43.08, .002	20.66, **.06**	.15, .01	147.59, .0001	22.42, [7.64, 45.28]	.29, [.10, .55]	.32, [.10, .55]
Aro	**SBF9/AHa**	138.91, .003	60.47, **.10**	.15, .01	516.37, .00	78.44, [25.50, 151.89]	.30, [.10, .54]	.33, [.09, .54]
Val	**LE/Rot**	1.56, .03	.63, **.39**	5.30, .006	.17, 0.1	.93, [.25, 1.95]	.23, [.06, .52]	.22, [.05, .48]
Val	**SBF2/Fl**	-.0003, .02	-.0002, **.30**	-.05, .0004	.003, .14	-.0001, [-.0004, 0001]	-.19, [-.46, .07]	.16, [.01, .38]
Val	**Ons/Fl**	-.0001, .003	-.0001, **.06**	-.01, .0003	.002, .34	.00, [-.0001, .00]	-.12, [-.40, 10]	.11, [.005, .33]
Hap	**SC/Rot**	.0001, .046	.0001, **.31**	.0003, .03	.17, .01	.0001, [.0000, .0001]	.19, [.06, .40]	.19, [.05, .38]
Ang	**LE/Rot**	1.56, .03	1.01, **.16**	-4.77, .05	-.11, .04	.55, [-.03, 1.44]	.14, [-.01, .35]	.13, [.01, .35]
Ang	**Ons/Fl**	-.0001, .003	-.0001, .04	.01, .004	-.002, .16	.00, [-.0001, 0.00]	-.14, [-.34, .06]	.14, .01, .31]
Ang	**SBF2/Fl**	-.0003, .02	-.0002, **.22**	.04, .01	-.003, .08	-.0001, [-.0003, .00]	-.17, [-.36, .07]	.17, [.01, 33]
Sad	**AT/AHa**	-39629.34, .01	-31772.85, .03	25.56, .05	-307.33, .15	-7856.47, [-29489.44, 1839.10]	-.10, [-.33, .02]	.10, [.01, .30]
Sad	**AT/Fl**	.40, .01	.32, .04	25.56, .05	.003, .17	.08, [-.06, .39]	.09, [-.08, .43]	.10, [.002, .36]
Ten	**SBF2/SHe**	1.05, .00	.89, .001	-.05, .00	-2.95, .35	.16, [-.24, .57]	.11, [-.17, .39]	.12, [.003, .35]
Ten	**SBF2/AHe**	15.73, .00	11.70, .01	-.05, .00	-73.63, .14	4.03, [-2.28, 10.88]	.18, [-.11, .47]	.19, [.01, .45]
Ten	**SBF2/AHa**	39.17, .003	28.83, .09	-.05, .00	-189.04, .36	10.33, [-12.05, 33.46]	.14, [-.17, .42]	.12, [.003, .32]
Ten	**SBF2/Flu**	-.0003, .02	.0001, **.73**	-.05, .00	.01, .001	-.0004, [-.0007, -.0001]	-.48, [-.76, -.18]	.41, [.12, .64]
Ten	**SBF9/SHe**	3.22, .0001	2.45, .002	-.11, .04	-7.20, .01	.77, [.15, 1.82]	.15, [.03, .34]	.18, [.04, .37]
Ten	**SBF9/AHe**	43.08, .002	28.79, .02	-.11, .04	-133.82, .002	14.29, [3.25, 32.35	.18, [.05, .38]	.21, [.05,.41]
Ten	**SBF9/AHa**	138.91, .003	108.21, .02	-.11, .04	-287.56, .07	30.71, [2.56, 85.76]	.12, [.01, .30]	.13, [.02, .30]

Abbreviations: see Table 1

Tenderness acted as mediator in the relationship between Spectral Flux of Sub-band 2 and Fluidity. Tenderness also mediated the effect of Spectral Flux of Sub-band 9 on Head Acceleration (direct path still significant), as well as on Head Speed (direct path significant) and on Hand Acceleration (direct path significant and *b*-path non-significant). The results for the models for Spectral Flux of Sub-band no 2 with Head Speed, Head Acceleration, and Hand Acceleration did not show any mediation effect of Tenderness. The statistical results are presented in Table 2.

3.2. Personality as moderator

If the strength of a relationship between two variables is dependent on the level of a third variable, this third variable is said to act as a moderator (Baron & Kenny, 1986; Preacher et al., 2007). Moderation effects are usually understood as an interaction between the independent variable and the moderator variable in predicting the dependent variable, where the effect of the independent variable depends on the level of the moderator. Figure 3 depicts the conceptual and the statistical model for moderation analysis.

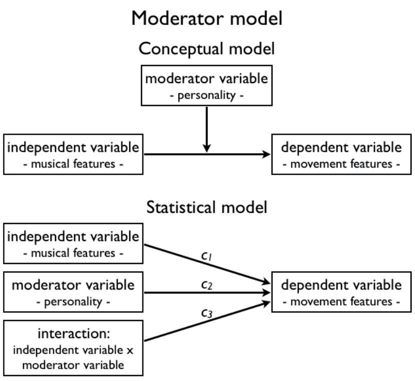


Figure 3. Conceptual and statistical model for moderation.

We hypothesized that personality of the dancers would act as a moderator. Thus, the level of personality trait would affect the relationship between musical features and movement characteristics. For instance, extrovert participants would make use of a certain movement feature to a higher extent than non-extrovert participants given a specific musical characteristic.

For analyzing the effect of personality, we first divided the 60 participants into three groups of 20 high-scorers, 20 mid-scorers, and 20 low-scorers for each personality trait separately and removed the mid-scorers from the analysis. The remaining two groups of participants were coded binary (instead of keeping the original scores). A similar approach was used in Luck et al. (2010).

For moderation to occur, we assume a relationship between the independent and the dependent variables, in our case the musical and movement features respectively. Therefore, we correlated the musical features with the movement features and present the result in Table 3.

Table 3. Correlations (*n*=30) between musical features and movement features.

	SHe	SHa	AHe	AHa	Fl	Rot
LE						*
Ons	*		**	*	**	
SBF2	***	*	***	**	*	
SBF9	***	***	**	**		
AT			*	**	**	
SC						*

* $p < .05$, ** $p < .01$, *** $p < .001$

Abbreviations: see Table 1

Subsequently, we performed moderation analysis employing the SPSS modeling tool PROCESS (Hayes, 2013)[2]. The results indicate that only two combinations showed moderation effect: we found significant interactions between Extroversion and Spectral Flux of Sub-band no 2 on Head Acceleration and between Conscientiousness and Number of Onsets on Fluidity. The results of both regression models are reported in Table 4.

[2] In moderation analysis, it is common to transform the predictors using grand mean centering before analysis, so that we obtain the effect of one predictor when the other predictor is its mean value (and not zero).

Table 4. Regression results of the moderation models with significant interactions.

Extroversion/SBF2/Head Acceleration, $R^2 = .13$

	β	*SE*	*t*	*p*
Extr	.28	25.29	10.25	.00
SBF2	.21	2.69	6.92	.00
Extr x SBF2	.07	2.69	2.32	.02

Conscientiousness /Ons/Fluidity, $R^2 = .04$

	β	*SE*	*t*	*p*
Cons	.03	.0007	1.22	.22
Ons	-.18	.0000	-6.42	.00
Cons x Ons	-.06	.0000	-2.00	.045

4. Discussion

We described two analysis approaches of multivariate nature – mediation and moderation – to enable the simultaneous inclusion of several variables into one analysis to investigate underlying relationships among them.

The results of the analysis suggest that such analyses serve as insightful ways of looking at relationships between different variables. We found the perceived emotional content of the music to mediate the effect of musical features on movement features in case of Arousal, Valence, Happiness, and Tenderness, which suggests to assume that the emotional content could account for the relationship between musical features and participants' movements. Thus, participants might have taken the emotions expressed in the music into consideration when moving to it. This could serve as support for Leman's theory of corporeal articulations (Leman, 2007), in particular for the concepts of "Embodied Attuning" and "Empathy". Participants' movements reflected both musical features and emotions perceived in the music, suggesting that participants have even tried to integrate them into a coherent outcome. This could also indicate that music and emotions are interconnected and co-existing, so "Empathy" could be maybe seen as an abstraction of "Embodied Attuning".

As regards personality as moderator, we found only two instances where moderation occurred, whereas in the remaining 88 combinations, no moderation occurred. This result, especially taken the low R^2 values of both models, would suggest that generally the level of personality did not change the relation between music and movement. In other words, if a relationship between a musical feature and a movement feature exists, it remains unaffected by the personality. One might even propose that an existing relationship between musical features and movement features can suppress the influence of personality. Furthermore, this result could indicate that musical characteristics and personality are independent factors that are related to music-induced movement in different, non-interactive, ways.

Interesting to note is that individual differences, such as personality, do not appear as a concept in the theory of corporeal involvement proposed by Leman (2007). However, if, for instance, movements are seen as being gestural expressions of a certain emotion expressed in the music (as in the concept of "Empathy"), then it seems likely to assume that movements can be seen as gestural expressions related to individual factors of a performer/listener/dancer. Thus, the integration of individual factors into approaches of body-related musical behavior might be a fruitful strategy to formulate a holistic theory of music-related and -induced corporeal articulations.

Further analysis attempts will include additional and more complex model configurations, in particular having several independent, dependent, and intervening variables. Using, for instance, more than one musical and more than one movement feature simultaneously in the analysis would provide a more holistic view on our data and might reveal more general results than with the present approach. Another attempt will combine both emotion and personality into one analysis – for example, in an approach called moderated mediation (Preacher et al., 2007) – to test if the magnitude of the indirect effect (emotions) is dependent on the moderator (personality). The method of path modeling (Schumacker & Lomax, 2010) might provide further opportunities for integrating both aspects into the same analysis. Moreover, structural equation modeling, an approach for testing and estimating causal relations between factors and/or observed variables (Schumacker & Lomax, 2010), will be tested. It offers the possibility to build latent factors from the observed variables, so higher-order relationships – in our case, "music" versus "movement" with intervening factors, such as "emotions" and "personality", might be testable as one comprehensive model. However, one issue regarding the latter approach is the small sample size of our data, as for structural equation modeling, a large sample size is a prerequisite for incorporation of latent variables.

Acknowledgments

We thank Mikko Leimu for help and technical assistance during the motion capture data collection, Jonna Vuoskoski for collecting the personality data, and Asko Tolvanen for support in multivariate statistics. This study was funded by the Academy of Finland (projects 118616, 125710, 136358).

References

Alluri, V., & Toiviainen, P. (2010). Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre. *Music Perception*, *27*(3), 223–242.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173–1182.

Burger, B., Saarikallio, S., Luck, G., Thompson, M. R., & Toiviainen, P. (2013). Relationships between perceived emotions in music and music-induced movement. *Music Perception*, *in press*.

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2012). Music moves us: Beat-related musical features influence regularity of music-induced movement. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference in Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences for Music* (pp. 183–187). Thessaloniki, Greece.

Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, *4:183*.

Godøy, R. I., Haga, E., & Jensenius, A. R. (2006). Playing "air instruments": mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty, & J.-F. Kamp (Eds.), *Gesture in Human-Computer Interaction and Simulation, Lecture Notes in Computer Science, 3881* (Vol. 3881, pp. 256–267). Berlin/Heidelberg, Germany: Springer.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, *76*(4), 408–420.

Hayes, A. F. (2013). The official reference for PROCESS is Hayes , A . F . (2013). An introduction to media- tion , moderation , and conditional process analysis : A regression-based approach . New York : Guilford Press, (2012), 1–39.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of experimental psychology. General*, *141*(1), 54–75.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.

Keller, P. E., & Rieger, M. (2009). Special Issue-Musical Movement and Synchronization. *Music Perception*, *26*(5), 397–400.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books.

Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *Proc. of the 10th Int. Conference on Digital Audio Effects* (pp. 1–8). Bordeaux, France: University of Bordeaux.

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA, London, UK: MIT Press.

Leman, M., & Godøy, R. I. (2010). Why Study Musical Gesture? In Rolf Inge Godøy & M. Leman (Eds.), *Musical Gestures. Sound, Movement, and Meaning* (pp. 3–11). New York, NY: Routledge.

Lesaffre, M., De Voogdt, L., Leman, M., De Baets, B., De Meyer, H., & Martens, J.-P. (2008). How Potential Users of Music Search and Retrieval Systems Describe the Semantic Quality of Music. *Journal of the American Society for Information Science and Technology*, *59*(5), 695–707.

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality*, *44*(6), 714–720.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G. De, & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, *58*(3), 177–192.

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*(2), 93–114.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227.

Saarikallio, S., Luck, G., Burger, B., Thompson, M. R., & Toiviainen, P. (2013). Dance moves reflect current affective state illustrative of approach-avoidance motivation. *Psychology of Aesthetics, Creativity, and the Arts*, *in press*.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, *36*(8), 1627–1639.

Schumacker, R. E., & Lomax, R. G. (2010). *A Beginner's Guide to Structural Equation Modeling*. London, UK: Routledge Academic.

Toiviainen, P., & Burger, B. (2013). *MoCap Toolbox Manual*. University of Jyväskylä: Jyväskylä, Finland. Available at http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual.

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE transactions on Speech and Audio Processing*, *10*(5), 293–302.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The Impact of the Bass Drum on Human Dance Movement. *Music Perception*, *30*(4), 349–359.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

VII

MOCAP TOOLBOX – A MATLAB TOOLBOX FOR COMPUTATIONAL ANALYSIS OF MOVEMENT DATA

by

Birgitta Burger & Petri Toiviainen, 2013

Proceedings of the Sound and Music Computing Conference (pp. 172–178), KTH Royal Institute of Technology, Stockholm, Sweden

MOCAP TOOLBOX – A MATLAB TOOLBOX FOR COMPUTATIONAL ANALYSIS OF MOVEMENT DATA

Birgitta Burger **Petri Toiviainen**
Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music,
University of Jyväskylä, Jyväskylä, Finland
birgitta.burger@jyu.fi petri.toiviainen@jyu.fi

ABSTRACT

The *MoCap Toolbox* is a set of functions written in Matlab for analyzing and visualizing motion capture data. It is aimed at investigating music-related movement, but can be beneficial for other research areas as well. Since the toolbox code is available as open source, users can freely adapt the functions according to their needs. Users can also make use of the additional functionality that Matlab offers, such as other toolboxes, to further analyze the features extracted with the *MoCap Toolbox* within the same environment. This paper describes the structure of the toolbox and its data representations, and gives an introduction to the use of the toolbox for research and analysis purposes. The examples cover basic visualization and analysis approaches, such as general data handling, creating stick-figure images and animations, kinematic and kinetic analysis, and performing Principal Component Analysis (PCA) on movement data, from which a complexity-related movement feature is derived.

1. MOTIVATION AND OVERVIEW

The *MoCap Toolbox* is a Matlab[1] toolbox dedicated to the analysis and visualization of motion capture (MoCap) data. It has been developed for the analysis of music-related movement, but is potentially useful in other areas of studies as well. It is open source, distributed under GPL license, and freely available for download at:

www.jyu.fi/music/coe/materials/mocaptoolbox.

The *MoCap Toolbox* is mainly intended for working with recordings made with an infrared marker-based optical motion capture system. Such motion capture systems are based on an active source emitting pulses of infrared light at a very high frequency, which is reflected by small, usually spherical markers attached to the tracked object (e.g., a participant dancing or playing an instrument). With each camera capturing the position of the reflective markers in two-dimensional, a network of several cameras can be used to obtain position data in three dimensions. Besides optical motion capture, the *MoCap Toolbox* can also be used for analyzing data captured with other tracker technologies, such as inertial or magnetic trackers. However, some features of the toolbox will be limited, since such trackers do not produce position data, but derivative data, (e.g., acceleration). Furthermore, the toolbox is optimized for the use of 3-dimensional position data, so using data with six degrees of freedom (position and rotation) might require customized adjustments of functions.

There are proprietary (closed source) software solutions available for motion capture analysis and visualization, such as Visual3D[2] or MotionBuilder[3], and applications that are primarily used for recording data (such as Qualisys Track Manager[4] or Vicon Nexus[5]). However, such applications are usually either too limited in their functionality, too focused on visualization and/or too restrictive to adapt to the needs of the researcher, such as developing new movement features useful for their individual research questions. To overcome these issues, we implemented this toolbox in Matlab, a generic scientific computing environment, and made it available to other researchers to be used in favor of their needs. The *Mocap Toolbox* is not the only Matlab toolbox available for motion capture analysis; one other toolbox worth mentioning is the toolbox created by Charles Verron [1]. This toolbox is more limited than the *MoCap Toolbox*, but offers a graphical user interface (GUI).

Matlab offers pre-built visualization opportunities and gives access to a large range of other functionality. Some functions included in the *MoCap Toolbox* use, for example, the *Signal Processing Toolbox* provided by MathWorks, or the *FastICA* package[6], a freely available third-party toolbox for Independent Component Analysis. Furthermore, the users themselves can make immediate use of the additional functionality and toolboxes provided by Matlab, for example the *Statistics Toolbox*, to further analyze features extracted with the *MoCap Toolbox* without the need to switch between different applications. *MoCap Toolbox* code is written using the generic Matlab syntax and is openly assessable, so users can add and adapt functions to their own needs.

[1] www.mathworks.com

[2] www.c-motion.com/products/visual3d/
[3] www.autodesk.com/motionbuilder
[4] www.qualisys.com/products/software/qtm/
[5] www.vicon.com/products/nexus.html
[6] www.cis.hut.fi/projects/ica/fastica/

The *MoCap Toolbox* supports various motion capture data formats, in particular the .c3d[7] file format (which, e.g., Vicon[8] or OptiTrack[9] optical motion capture systems can produce), the .tsv format and the .mat format, both produced by the Qualisys motion capture system[10], and the .wii data format produced by the *WiiDataCapture* software[11].

The *MoCap Toolbox* provides 64 functions for analyzing and visualizing motion capture data. The main categories can be summarized as data input and edit functions, coordinate transformation and coordinate system conversion functions, kinematic and kinetic analysis functions, time-series analysis functions, visualization functions, and projection functions. Furthermore, it uses three different data structures, the *MoCap data structure*, the *norm data structure*, and the *segm data structure*. To convert between the different data representations and enable certain visualizations, three different parameter structures are used, the *m2jpar*, the *j2spar*, and the *animpar structures*. Both the data and the parameter structures will be discussed and explained in the next section.

2. DATA REPRESENTATIONS

The *MoCap Toolbox* uses three different data structures, the *MoCap data structure*, the *norm data structure*, and the *segm data structure*. A *MoCap data structure* instance is created when mocap data is read from a file to the Matlab workspace using the function `mcread`. A *MoCap data structure* contains the 3-dimensional locations of the markers (in the `.data` field) as well as basic information, including the type of structure, the file name, number of frames of the recording, the number of cameras used for the recording, the number of markers in the data, the frame rate, the names of the markers, and the order of time differentiation of the data. Additionally, the *MoCap data structure* contains fields for data captured with analog data, such as EMG. Finally, the time stamp of the recording and the data type (e.g., 3D) can be added.

A *MoCap data structure* instance is also created when the function `mcm2j` is used. This function transforms a marker representation to a joint representation. These two representations use the same data structure, although they are conceptually different: the marker representation reflects the actual marker locations, whereas the joint representation is related to locations derived from marker locations. A joint can consist of one marker, but it can also be derived from more than one markers. It can, for example, be used for calculating the location of a body part where it is impossible to attach a marker. The midpoint of a joint, for instance, can be then derived as the centroid of four markers around the joint.

The *norm data structure*, created by the function `mcnorm`, is similar to the *MoCap data structure*, except that its `.data` field has only one column per marker. This column contains the Euclidean norm of the vector data from which it was derived. If, for instance, `mcnorm` is applied to velocity data, the resulting *norm data structure* holds the magnitudes of velocities, or speeds, of each marker.

The third data structure, the *segm data structure*, is not, like the other two, related to points in space (markers or joints), but to segments of the body (see, e.g., [2]). The function `mcj2s` performs a transformation from a joint representation to a segment representation and produces as output a *segm data structure* instance. Most fields of a *segm data structure* are similar to the ones of a *MoCap data structure*, however, the `.data` field is replaced by four other fields. The `.parent` field contains information about the kinematic chains of the body, i.e., how the joints are connected to form segments, and how segments are connected to each other. The fields `.roottrans` and `.rootrot` store the location and orientation of the center of the body, the root. The `.segm` field consists of several subfields that store the orientation of the body segments in several ways. The `.eucl` subfield contains for each segment the Euclidean vector pointing from the proximal to the distal joint of the segment. The length of each segment is stored in the `.r` subfield. The `.quat` subfield includes the rotation of each segment as a quaternion representation (see, e.g., [3] and [4]). Finally, the `.angle` subfield contains the angles between each segment and its proximal segment.

To convert between the different representations and to enable certain visualizations, the *MoCap Toolbox* offers three different parameter structures: *m2jpar, j2spar*, and the *animpar structures*.

The *m2jpar structure* is used by the function `mcm2j` and contains the information needed to perform the transformation from marker to joint representation. Besides fields holding the number of joints and the names of the joints, it includes a field with the numbers of the markers defining the location of each joint.

The *j2spar structure* is used by the function `mcj2s` and contains the information needed to perform the transformation from joint to segment representation. Besides the fields containing the segment names and the number of the root (center of the body) joint, it includes fields with the numbers of the three joints that define the frontal plane of the body and a vector indicating the number of the parent segment (the segment that is proximal in the kinematic chain) for each segment.

The *animpar structure* is used by the functions `mcplotframe` and `mcanimate` and contains the information needed to create frame (stick figure) plots and animations. The structure includes fields for the screen size, limits of the plotted area, viewing angles, marker sizes, plotting colors, connection line configurations and widths, and plotting of marker and frame numbers. Additionally, the structure contains fields related to creating animations, such as the frames per second, a substructure for perspective projection parameters, and settings for plotting marker traces.

[7] www.c3d.org
[8] www.vicon.com
[9] www.naturalpoint.com/optitrack/
[10] www.qualisys.com
[11] www.jyu.fi/music/coe/materials/mocaptoolbox

3. USING THE TOOLBOX

In what follows, we will give an introduction to the use of the toolbox for research and analysis purposes.

The *MoCap Toolbox* manual, provided with the download of the toolbox, offers an example chapter with eleven demos explaining the basic usage of the toolbox. Additionally, a demo data set called `mcdemodata`, including motion capture data and associated parameter structures, is provided with the download. The *MoCap data structures* `dance1` and `dance2` used below are available in the `mcdemodata` data set.

3.1 Reading Data and Filling Gaps

Recorded motion capture files can be imported into Matlab using the function `mcread` storing the content of the file as a *MoCap data structure*, i.e.,

```
d = mcread('file.tsv');
```

An essentially useful first step is usually to check for missing frames in the recording. Taking the *mocap data structure* `d`, we can use

```
mcmissing(d)
```

to detect missing frames in the recording. In case of missing data, we can fill them using linear interpolation with the function `mcfillgaps`:

```
d = mcfillgaps(d);
```

From this point onwards, we will use the two *MoCap data structures* `dance1` and `dance2` from the `mcdemodata`. Since they are already available as *MoCap data structures* and do not contain missing data, both importing and gap filling are not required anymore.

3.2 Visualizing and Animating Data

A good approach to get an overview of the data is to visualize and animate data. Using the *MoCap Toolbox*, mocap data can be plotted in different two ways: as a time series or as single frames. As a function of time, marker location data can be plotted with the function `mcplottimeseries`, e.g.,

```
mcplottimeseries(dance1,[1 20 28],
 'dim',3)
```

which plots the third/vertical dimension of markers 1, 20, and 28 (left front head, right hand, and right foot) (see Fig. 1).

Marker locations as single frames can be plotted using the function `mcplotframe` (using the (x,y) projection of the markers):

```
mcplotframe(dance1,450);
```

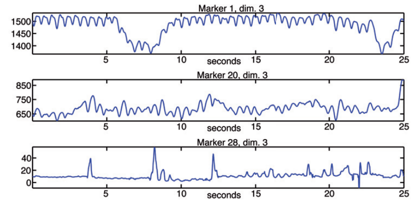


Figure 1. Marker location data plotted as function of time using `mcplottimeseries`.

This call, plotting the 450th frame of the recording (see Fig 2a), uses the default *animation parameter structure*. However, if a customized *animpar* structure is used, we can, for instance, set the connection lines between the markers to obtain a visualization that is easier to understand and that looks more human-like (see Fig. 2b):

```
ap = mcinitanimpar;
ap.conn = [1 2; 2 4; 3 4; 3 1; 5 6; 9
 10; 10 12; 11 12; 11 9; 8 9; 8 10; 8
 5; 8 6; 5 9; 5 11; 6 10; 6 12; 7 11;
 7 12; 7 5; 7 6; 5 13; 13 15; 13 16;
 16 19; 15 19; 6 14; 14 17; 14 18; 17
 20; 18 20; 9 21; 11 21; 10 22; 12 22;
 21 23; 23 25; 23 26; 25 26; 22 24; 24
 27; 24 28; 27 28];
mcplotframe(dance1,450,ap);
```

In case users collected the data with a Qualisys motion capture system and created a bone structure during the labeling process in the Qualisys software, they can export the so-called label list (which contains the marker connections) and use this file to create the connection matrix by employing the function `mccreateconnmatrix`.

We can change the general color scheme and the colors of individual markers, connector lines, traces, and numbers by adjusting the values of the respective fields of the *animpar* structure, for example (see Fig. 2c):

```
ap.colors = 'wrbgy';
ap.markercolors = 'bmgyrrrrrrrrkk';
mcplotframe(dance1,450,ap);
```

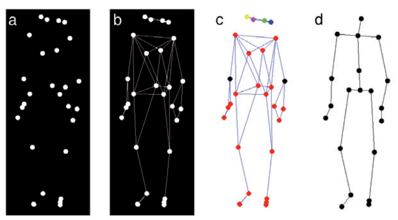


Figure 2. Marker location data plotted as frame using `mcplotframe`: a) using the default parameters; b) using a connection matrix; c) changing colors; d) joint transformation.

The function `mcanimate` is used to create animations:

```
mcanimate(dance1,an);
```

The *MoCap Toolbox* produces the single animation frames as .png files. They have to be compiled into a movie using other software, such as QuickTime Pro on Mac, or MovieMaker on Windows.

Animations can be created as 2D projections in two ways, either orthographic (default) or perspective, the latter one by including the perspective projection parameter:

```
mcanimate(dance1,an,1);
```

3.3 Kinematic Analysis

Kinematic variables, such as velocity and acceleration, are estimated using the time-derivative function `mctimeder`:

```
d1v = mctimeder(dance1,1); %vel.
d2v = mctimeder(dance2,1);
d1a = mctimeder(dance1,2); %acc.
d2a = mctimeder(dance2,2);
```

To analyze such time series, we can calculate their means and standard deviations using `mcmean` and `mcstd` (ignoring eventual missing frames). For this sample analysis, we will take the norm data, that is, the magnitudes of the 3-dimensional data of velocity and acceleration. To simplify the approach, we first combine the data from marker 1 (left front head) of the four *MoCap data structures* using mcconcatenate:

```
dva = mcconcatenate(d1v,1,d1a,1,d2v,1,
 d2a,1);
dva_mean = mcmean(mcnorm(dva));
dva_std = mcstd(mcnorm(dva));
```

The results (see Table 1) show that both mean and standard deviation of velocity and acceleration of the left front head marker are higher for `dance2` than for `dance1`, so the dancer in `dance2` moved faster and at a wider range of speeds and also used more and larger directional changes.

		mean	SD
velocity	`dance1`	235.85	110.79
	`dance2`	520.24	192.27
acceleration	`dance1`	2233.66	1326.55
	`dance2`	3347.21	1423.19

Table 1. Means and standard deviations of velocity and acceleration (magnitudes) of the left front head marker data of `dance1` and `dance2`.

The cumulative distance travelled by a marker can be calculated with the function `mccumdist` (returning a *norm data structure*):

```
d1dist = mccumdist(dance1);
d2dist = mccumdist(dance2);
```

We use the Matlab function `barh` for plotting markers 1 (left front head), 20 (right finger), and 28 (right foot) (see Fig. 3):

```
figure, barh([d1dist.data(1500,[1 20
 28]); d2dist.data(1500,[1 20
 28])],'b');
```

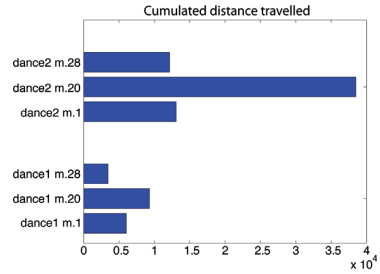


Figure 3. Cumulated distance travelled by markers 1, 20, and 28 of mocap data `dance1` and `dance2` (labels and title were added separately).

We can see in Figure 3 that the three markers, especially the right hand marker, travelled more for `dance2` than for `dance1`, so we can assume that the amount of movement was higher in `dance2`.

A measure related to the amount of movement is the area covered by the movement, which can be calculated using `mcboundrect`. If we want to calculate the bounding rectangle of the four hip markers, we do:

```
br1 = mean(mean(mcboundrect(dance1,[9
 10 11 12])));
br2 = mean(mean(mcboundrect(dance2,[9
 10 11 12])));
```
[12]

The bounding rectangle value for `dance1` equals .1806 and for `dance2`, it equals .9724. Since the value for `dance2` is higher, `dance2` not only had a higher amount of movement, but also used more space than `dance1`. The bounding rectangle measure was found to be a relevant movement feature in [5] and [6].

We can also calculate distances between markers using `mcmarkerdist`. The standard deviation of the distance between left and right finger,

```
md1 = std(mcmarkerdist(dance1,19,20));
md2 = std(mcmarkerdist(dance2,19,20));
```

gives us information about the variability of the marker distance. The standard deviation of the finger marker distance for `dance1` equals 49.0 and for `dance2` 175.65, so the fingers in `dance2` exhibited more variable distances.

Periodicity of movement can be estimated using the function `mcperiod`. It is based on autocorrelation, and

[12] `mcboundrect` uses window decomposition. The function output here is averaged across the windows and the four markers.

either the first or highest peak of the autocorrelation function is taken as periodicity estimation dependent on the parameter input. With

```
d1m1 = mcgetmarker(d1a,20);
d2m1 = mcgetmarker(d2a,20);
[per1 ac1 eac1] = mcperiod(d1m1,2,
 'highest');
[per2 ac2 eac2] = mcperiod(d2m1,2,
 'highest');
```

we calculate the periodicity of the acceleration of the right finger marker. `mcperiod` resulted in a periodicity estimate for each dimension being [1.04, 0.53, 0.52] for `dance1` and [1.04, 1.01, 1.06] for `dance2`. While the first dimension is similar, the second and third dimensions are roughly half for `dance1`, suggesting that in this case the finger moved in double tempo in y and z directions.

A more accurate periodicity analysis can be performed using windowed autocorrelation:

```
[per1 ac1 eac1] = mcwindow(@mcperiod,
 d1m1,2,0.25);
[per2 ac2 eac2] = mcwindow(@mcperiod,
 d2m1,2,0.25);
```

To allow visual inspection of the time development of the periodicity, the enhanced autocorrelation (eac) matrix can be plotted as an image (see Fig. 4). The colors indicate the regularity of periodic movement, with warm colors corresponding to regions of regular periodic movement in the period-time plane:

```
figure, imagesc(eac1(:,:,3)), axis xy
set(gca, 'XTick',0:4:46, 'XTickLabel',
 0.5*(0:4:46), 'YTick',[0 30 60 90
 120], 'YTickLabel',[0 0.5 1 1.5 2.0])
figure, imagesc(eac2(:,:,3)), axis xy
set(gca, 'XTick',0:4:46, 'XTickLabel',
 0.5*(0:4:46), 'YTick',[0 30 60 90
 120], 'YTickLabel',[0 0.5 1 1.5 2.0])
```

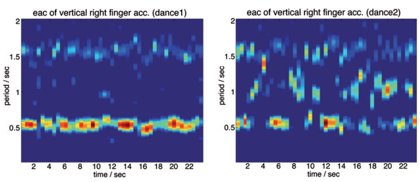


Figure 4. Enhanced autocorrelation function of the vertical components of the right finger acceleration in `dance1` and `dance2`.

We can see in Figure 4 that the vertical component of the right finger acceleration of `dance1` shows quite clear periodic movement with a period of about 500 milliseconds, whereas the periodicity for `dance2` is weaker and more irregular.

3.4 Kinetic Analysis

The *MoCap toolbox* offers the possibility to calculate kinetic variables using Dempster's body-segment model [7]. To make our present data compatible with Dempster's model, we first have to reduce the amount of markers from 28 to 20. We will accomplish this with a marker-to-joint transformation, implemented in the function `mcm2j`. The *m2jpar parameter structure* required for this transformation is created like this:

```
m2j = mcinitm2jpar;
m2j.nMarkers = 20;
m2j.markerNum = {[9 10 11 12],[9 11],
 21,23,26,[10 12],22,24,28,[7 8 7 8 9
 10 11 12],[5 6],[1 2 3 4],5,13,[15
 16],19,6,14,[17 18],20};
m2j.markerName = {'root', 'lhip',
 'lknee','lankle','ltoe','rhip',
 'rknee','rankle','rtoe','midtorso',
 'neck','head','lshoulder','lelbow',
 'lwrist','lfinger','rshoulder',
 'relbow','rwrist','rfinger'};
```

The joint `'root'`, for example, is obtained by calculating the centroid of markers 9, 10, 11, and 12. The marker-to-joint transformation is carried out as follows:

```
d1j = mcm2j(dance1,m2j);
d2j = mcm2j(dance2,m2j);
```

Figure 2d visualizes frame 450 of the joint representation of `dance1`. The next step is to do the joint-to-segment transformation. The *j2spar parameter structure* required for the transformation is created like this:

```
j2s = mcinitj2spar;
j2s.rootMarker = 1;
j2s.frontalPlane = [6 2 10];
j2s.parent = [0 1 2 3 4 1 6 7 8 1 10
 11 11 13 14 15 11 17 18 19];
j2s.segmentName = {'lhip','lthigh',
 'lleg','lfoot','rhip','rthigh',
 'rleg','rfoot','ltorso','utorso',
 'neck','lshoulder','luarm','llarm',
 'lhand','rshoulder','ruarm','rlarm',
 'rhand'};
```

The joint-to-segment transformation is accomplished using the function `mcj2s`:

```
d1s = mcj2s(d1j,j2s);
d2s = mcj2s(d2j,j2s);
```

In order to calculate kinetic variables, such as energy, each body part has to be associated to its parameter (i.e., masses and lengths) specified by the Dempster model. Therefore, a variable is created specifying the types of the segments[13]:

[13] For a list of the segment types, see the *MoCap Toolbox* manual.

```
s_ind = [0 0 8 7 6 0 8 7 6 13 12 10 11
 3 2 1 11 3 2 1];
```

This variable associates each joint with a segment type. Each component indicates the type of body segment for which the respective joint is a distal joint. Joints that are not distal to any segment have zero values.

The parameters for each body segment can be then obtained using the function `mcgetsegmpar`:

```
spar = mcgetsegmpar('Dempster',s_ind);
```

With this body-segment representation we can estimate kinetic variables for each segment individually. The time-average of the kinetic energy of the whole body, for example, can be calculated like this:

```
[trans1 rot1] = mckinenergy(d1j,d1s,
 spar);
[trans2 rot2] = mckinenergy(d2j,d2s,
 spar);
kinEn1 = sum(mcmean(trans1)) +
 sum(mcmean(rot1));
kinEn2 = sum(mcmean(trans2)) +
 sum(mcmean(rot2));
```

The value for the overall kinetic energy of `dance1` equals 2.21, and the value for `dance2` is 11.37, thus more energy was used in `dance2`, which supports our argumentation drawn earlier, that there is more movement in `dance2` than in `dance1`.

3.5 Principal Component Analysis (PCA)

Principal component analysis can be used to decompose motion capture data into components that are orthogonal to each other. By using

```
[pc1 p1] = mcpcaproj(d1j,1:5);
[pc2 p2] = mcpcaproj(d2j,1:5);
```

we calculate the first five principle component projections of the position data (as joint representations) of `d1j` and `d2j`. `p1.l` and `p2.l` contain the amount of variance explained by each component. From these variances, we can derive, for instance, a measure of movement complexity, defined as the cumulative sum of the proportion of explained variance contained in the first five PCs (see, e.g., [5] and [8]):

```
pcapropvar1 = cumsum(p1.l(1:5));
pcapropvar2 = cumsum(p2.l(1:5));
```

The results, presented in Table 2, indicate that, in case of `pcapropvar1`, most movement is already explained with the first component, and the first five components explain almost all movement. In case of `pcapropvar2`, however, only about 50% of the movement is explained with the first component, and the first five components explain less than the first five components of `pcapropvar1`, so more components are needed to fully explain the movements of `dance2`. Such a movement would be characterized as complex, since a high number of PCs is needed to explain the movement sufficiently, whereas a low proportion of unexplained variance (`dance1` case) implies a simpler movement.

	`pcapropvar1`	`pcapropvar2`
`cumsum(1)`	0.79	0.48
`cumsum(1:2)`	0.90	0.78
`cumsum(1:3)`	0.95	0.85
`cumsum(1:4)`	0.97	0.90
`cumsum(1:5)`	0.98	0.93

Table 2. Cumulative variances of the first five principle components for `dance1` (`pcapropvar1`) and `dance2` (`pcapropvar2`).

4. CONCLUSION

The *MoCap Toolbox* is a Matlab toolbox dedicated to the analysis and visualization of motion capture data. It has been developed for the analysis of music-related movement, but is potentially useful in other areas of studies as well. It has attracted researchers' attention working in various fields and has been downloaded for being used in a wide range of different research purposes; music-related, but also, for instance, face recognition, sports, gait, or biomechanics research. It has also gained attraction in artificial intelligence research, such as robotic motion, human-robot interaction, and machine learning.

The *MoCap Toolbox* has continuously been developed further since its first launch in 2008 by both the authors and the users, whose bug reports and suggestions for new functionality has greatly helped to improve and extend it.

In the future error handling will be improved, for instance, when wrong data structures are used. Toolbox functions usually recognize the mistake, but in the present version, some functions do not return sufficiently clear error messages.

Furthermore, some functions will be adapted to standard Matlab conventions, as it is already done in, for instance, `mcplottimeseries` (specifying the plotting parameters as a strings-value combination).

Individual functions will be improved, such as `mcfillgaps`, that would benefit from the implementation of more advanced gap-filling methods than linear filling, for example spline interpolation. Additionally, more body segment models besides Dempster's model will be included, such as models proposed in [9] or [10].

As commercial tools (e.g., Visual3D) commonly provide GUIs instead of operating on a command-line basis, a graphical user interface could also be implemented for the *MoCapToolbox*. It would make the toolbox more user-friendly – for example, connection matrices of stick figures could be drawn in the GUI, or gap filling could be graphically supported.

Acknowledgments

This study was supported by the Academy of Finland (project 118616).

5. REFERENCES

[1] C. Verron, Traitement et Visualisation de Vonnées Gestuelles Captées par Optotrak. IDMIL Report, 2005.

[2] D.G.E. Robertson, G.E. Caldwell, J. Hamill, G. Kamen, and S.N. Whittlesey, Research Methods in Biomechanics. Human Kinetics, 2004.

[3] P. Kelland, Introduction to Quaternions, with Numerous Examples. Rarebooksclub.com, 2012.

[4] A.J. Hanson, Visualizing Quaternions. Morgan Kaufmann Publishers, 2005.

[5] B. Burger, S. Saarikallio, G. Luck, M.R. Thompson, and P. Toiviainen, "Relationships between perceived emotions in music and music-induced movement," in Music Perception 30, 2013, pp. 519-535.

[6] G. Luck, S. Saarikallio, B. Burger, M.R. Thompson, and P. Toiviainen, "Effects of the Big Five and musical genre on music-induced movement," in Research in Personality 44, 2010, pp. 714-720.

[7] W.T. Dempster, Space Requirements of the Seated Operator: Geometrical, Kinematic, and Mechanical Aspects of the Body with Special Reference to the Limbs. WADC Technical Report 55-159, Wright-Patterson Air Force Base, 1955.

[8] S. Saarikallio, G. Luck, B. Burger, M.R. Thompson, and P. Toiviainen, "Dance moves reflect current affective state illustrative of approach-avoidance motivation," in Psychology of Aesthetics, Creativity, and the Arts, in press.

[9] C.E. Clauser, J.T. McConville, and J.W. Young, Weight, volume and center of mass of segments of the human body. AMRL Technical Report 69-70, Wright-Patterson Air Force Base, 1969.

[10] C.L. Vaughan, B.L. Davis, and J.C. O'Connor, Dynamics of Human Gait. Human Kinetics, 1992.