JYVÄSKYLÄ STUDIES IN HUMANITIES
176

Marc Thompson

The Application of Motion Capture to Embodied Music Cognition Research

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN HUMANITIES 176

Marc Thompson

The Application of Motion Capture to Embodied Music Cognition Research

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa
maaliskuun 21. päivänä 2012 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities of the University of Jyväskylä,
in building Seminarium, auditorium S212, on March 21, 2012 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2012

The Application of Motion Capture to Embodied Music Cognition Research

JYVÄSKYLÄ STUDIES IN HUMANITIES 176

Marc Thompson

The Application of Motion Capture to Embodied Music Cognition Research

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2012

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

Cover picture by Veera Aivio

URN:ISBN:978-951-39-4690-6
ISBN 978-951-39-4690-6 (PDF)

ISBN 978-951-39-4689-0 (nid.)
ISSN 1459-4331

Jyväskylä University Printing House, Jyväskylä 2012

ABSTRACT

Thompson, Marc R.
The application of motion capture to embodied music cognition research
Jyväskylä: University of Jyväskylä, 2012, 86 p.
(Jyväskylä Studies in Humanities
ISSN 1459-4331; 176)
ISBN 978-951-39-4689-0 (nid.)
ISBN 978-951-39-4690-6 (PDF)

Music and expressive body movements are inseparable. Musicians use movements to convey emotional expressivity, conductors use gestures to convey musical structure, and co-performers rely on corporeality to maintain group cohesion. Additionally, the spontaneously-produced movements that occur when listening to music appear to naturally synchronize with the musical pulse. Despite this tendency common among most people, movement patterns appear to be highly individualistic and perhaps related to expressive cues inherent in the music. A growing awareness of the significant role that expressive body movement plays in music-related activities has given rise to the application of embodied cognition to music. Embodied cognition is the view that knowledge is attained through sensorimotor representations of the world and that the human body acts as the main mediator between our thoughts and the environment. Music-related activities such as performance and dance are excellent examples embodied cognition, as they require online tracking of perceptual stimuli that is in turn being augmented by our own actions. Furthermore, the embodied approach to music assumes that we can interact with music at different levels of engagement ranging from synchronization with low-level features such as the beat to empathizing with the music's higher-level features such as expressive intentions. In both cases, musical meaning is developed through goal-directed corporeal actions, which both react and contribute to musical experiences. The study of such phenomena has greatly benefited from advances in optical motion capture systems, which can track movements with a high degree of precision. For analysing the movements, semantically significant features are then extracted from the raw motion capture data. These variables can describe the kinematics, kinetic or other behavioural features of musicians and dancers, and be subjected to an infinite variety of analysis methods. The aim of this thesis is to demonstrate, through experimental research, some applications of motion capture within the area of embodied music cognition. Across six independent studies, four music-related activities were analysed: piano performance, conducting gestures, choreographed choral performance and music-induced movements.

Keywords: embodied music cognition, motion capture, gestures, music-related movements

Author's address: Marc R. Thompson
Department of Music
PL 35 (M)
40014, University of Jyväskylä
Finland

Marc.Thompson@jyu.fi

Supervisor: Dr. Geoff Luck
Finnish CoE in Interdisciplinary Music Research
University of Jyväskylä

Reviewers: Professor Rolf Inge Godøy
Department of Musicology
University of Oslo

Dr. Peter Keller
MARCS Auditory Laboratories
University of Western Sydney

Opponent: W/Professor Jane W. Davidson
School of Music
The University of Western Australia

AKNOWLEDGEMENTS

The real voyage of discovery consists not in seeking new landscapes but in having new eyes.
—Marcel Proust (1871-1922)

I move on to another day, to a whole new town with a whole new way.
– The World at Large by Modest Mouse

At a crossroads such as this, it is opportune to reflect on and show gratitude to the talented people who have generously devoted their time, knowledge and tutelage to me throughout these last four years.

First, I wish to thank Dr. Geoff Luck, who has supervised both my Masters and Doctoral studies. I am grateful for his mentorship, partnership and friendship throughout every stage of this project, and for assisting me in converting my unpinned ideas into clear and achievable goals.

To Prof. Petri Toiviainen, I am grateful for his continued support and guidance since September 2005. I am appreciative to Petri for sharing his vast knowledge on a daily basis on subjects as varied as linear algebra, empirical mode decomposition, Wii-controller hacking and Finnish mushrooms, to name a few. His intelligence, kindness and modesty embody the qualities that all great academics ought to have.

At McGill University, I thank Marcelo M. Wanderley, whose guidance on my bachelor's thesis some years ago, which focused on musical gestures and the role they play in music performance, influenced me to study music cognition in the first place. I am very glad to have had the opportunity to learn from him again during my six-month sabbatical at IDMIL in 2010.

From an administrative point of view, working at the Music Department was made easy with the great help of Markku Pöyhönen. Markku was a constant informant of upcoming events, tasks and responsibilities. I believe he is the key reason the Department is as successful as it is.

To Tommi Himberg, I am thankful for his invitation to join several exciting projects of which I am proud to have been a part of, and for the numerous conference proceedings and articles we have written together.

To my fellow doctoral students, Rafael Flores Ferrer, Vinoo Alluri, Jonna Vuoskoski, Birgitta Burger, Anemone Van Zijl, Pasi Saari and Marko Punkanen, thank you all for the enlightening conversations, laughter, endless lunch hours and the vast array of sweet and salty snacks we have shared over the years.

I also wish to thank Suvi Saarikallio, Oliver Lartillot, Prof. Tuomas Eerola, Prof. Jaakko Erkkilä, Esa Ala-Ruona, Jörg Fachner, Jukka Louhivuori, Pekka Toivanen, Hannes Juutilainen, Mikko Myllykoski, and Mikko Leimu who have all in their unique ways played a role in helping me achieve success at this department.

This summary greatly benefitted from having been reviewed by two eminent scholars. I thank Peter Keller and Rolf Inge Godøy for their insightful comments and support of this thesis and its constituent articles. In addition, I thank Alex Reed for help in proofreading this summary.

Though away from my homeland, I never feel without family. I thank Virpi, Mikko and Lauri who have given me so much since my first months in Jyväskylä. I also thank my in-laws, Tarja, Antero, Janne, Veera and Samppa for their love, kindness and acceptance.

In Canada, I thank my parents, Christopher and Nicole, and my siblings, Philippe, Charles and Amanda and their families, for supporting me in whichever strange and unlikely avenues my journey has taken me.

Most of all I am thankful to my wife Noora and our beautiful daughter Elsa, to whom I dedicate this work. Their love, support and encouragement—past, present and future—continue to be my guiding light.

Jyväskylä, November 10th, 2011
Marc R. Thompson

LIST OF ORIGINAL PUBLICATIONS

I. Thompson, M. R., & Luck, G. (2008). Effect of pianists' expressive intention on amount and type of body movement. In Miyazaki, K., Hiraga, Y., Adachi, M., Nakajima, Y. & Tsuzaki, M. (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition, (ICMPC10)* (pp. 540-544). Sapporo, Japan: Hokkaido University.

II. Luck, G., Toiviainen, P. & Thompson, M.R. (2010. Perception of expression in conductors’ gestures: A continuous response study. *Music Perception 28*(1), 47-57.

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

IV. 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.

V. Himberg, T. & Thompson, M. (2011). Learning and synchronizing dance movements in South African songs—cross-cultural motion-capture study. *Dance Research vol. 29(2), 305-328.*

VI. Thompson, M. R., & Luck, G. (2011). Exploring relationships between pianists’ body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, DOI: 10.1177/1029864911423457.

AUTHOR’S CONTRIBUTIONS TO THE ARTICLES

I. This is a proceedings paper on which I am the first author. I collected the motion capture data myself, conducted the analysis and wrote the paper. The second author, Geoff Luck advised me as to how to design the study and suggested how the data could be analysed.

II. I acted as this paper’s third author. My role in the data collection process consisted of developing the computer interface in Max/MSP with which the continuous slider ratings were collected and stored. I also rendered the animation frames viewed by the participants. For the data analysis, I computed the movement variables that were used as predictors in the regression models. In the article itself, I wrote the Methods sections and portions of the Introduction.

III. I acted as this paper’s third author. My main role consisted of scheduling and administrating the data collection phase. I wrote part of the Methods sections and proofread the paper prior to its submission for review.

IV. I acted as this paper’s fourth author. My main role consisted of collecting the data from 64 participants. I also programmed the Max/MSP patch that played back the musical stimuli for the participants. This study is part of an on-going project entitled Music, Movement & Personality (MMP). I have been and continue to be a consultant and analyst for the other aspects of the MMP project and its forthcoming publications.

V. This is a joint first-authorship paper written by Tommi Himberg and I. For this project, I was involved with the planning and technical assisting during the data collection process, which took place during a two-day inter-cultural dance workshop held at the University of Jyväskylä. My contribution to the article was to conduct the pairwise synchronisation analysis. Throughout the writing process, Tommi and I consulted each other regarding our individual sections.

VI. I wrote this paper with consultation and proof reading by the second author Geoff Luck. I also collected the movement data and conducted the analysis.

CONTENTS

1 BACKGROUND

1.1 Introduction

From a multidisciplinary perspective, this thesis aims to demonstrate, through experimental research, some applications of motion capture within the area of embodied music cognition.

Music has an inherent association to movement. We use physical movements to produce music and very often music causes us to move spontaneously, freely and expressively. The embodied approach to cognition suggests that we are first and foremost embodied beings—that we gain knowledge through sensorimotor interactions with the environment. This is a contrasting view to the cognitivist model of cognition, which starts from the premise of mental facilities being akin to a digital computer. The embodied approach reappears in many fields of research. One primary example comes from linguistics, and emphasizes the notion that figurative language (i.e. metaphors) is linked to the physical properties of the human body to shape our understanding of the world (Lakoff & Johnson, 1980; Lakoff & Johnson, 1999; Shapiro, 2007). It is not surprising then that the notion of being *moved by music* is one of the primary metaphors used to describe an emotional encounter with music. This concept juxtaposes a change in our emotional state with a spatial displacement of our physical state, both of which happen frequently when engaging in music-related activities. The metaphorical link between music and movement dates back to the ancient Greeks, and possibly earlier. Aristotle wrote in *Politics* (book VIII, chapter 5) that when we listen to music, our souls imitate the music's emotional character (Schoen-Nazzaro, 1978): depending on the emotional character of the music, the soul can be *moved* towards virtue or vice.

The central idea driving this thesis is that music-related body movements are ubiquitous throughout all forms of music-related activity, and serve a purpose for musicians and non-musicians alike. It is striking that the role of body movement within music-related research has only really been taken into account in the past 20 years. Part of the motivation pushing research in this direction is the newly realized hallmark that in music performance, movement is not secondary to sound but plays an integral part within the larger musical experience. The ways in which musicians move their bodies when performing, indicate expressive intentionality (Davidson, 1993; 1994), communicate specific emotions (Dahl & Friberg, 2007) and can influence an observer's interest in the music (Broughton & Stevens, 2009). Conductors use a wide range of facial gestures and arm movements to indicate, not only musical structure, but also expressive intentionality to an ensemble (Luck & Toiviainen, 2006). Co-performers also use gestures and other visual cues in rehearsals to communicate musical ideas to each other (Williamon & Davidson, 2002) or to synchronize with one another (Goebl & Palmer, 2009). The communicative and expressive roles of body movements extend to other music-related activities. Anyone who has been to a nightclub will recognize that dancers exhibit a wide range of movements to express themselves. Generally, dance moves are inadvertently synchronized to the music's periodic structure, which suggests that movement plays a role in parsing the music's metrical hierarchy (Toiviainen, Luck & Thompson, 2009). However, the precise characteristics of dance moves such as ones found in dance clubs can be driven by a plethora of individual factors such as, for instance, the dancer's personality traits (Luck, Saarikallio & Toiviainen, 2009). The communicative aspect of body movements is also useful for teaching dance choreography. As a social activity, dancing offers an interesting framework in which to study the dynamics of interpersonal communication and joint actions between experts and novices (Keller, 2008; Himberg & Thompson, 2009, 2010).

Moving to music appears to be inherently social. Infants are born with facilities to imitate and engage in rhythmic games with their parents (Trevarthen, Kokkonake & Fiamenghi, 1999). Young people increase their movements and are more synchronized with the music's beat when dancing in groups (de Bruyn et al., 2009). Our innate ability to synchronize to a musical beat anticipates more sophisticated forms of musical communication. For example, it has been shown that even individuals with no musical experience can easily synchronize to the pulse of a conductor's gestures (Luck, 2000). The perceptual system is able to recognize rhythmic movements in others, which enables us to make predictions about future behaviour (Condon, 1982).

Moving in relation to music is shared by all musical cultures (Blacking, 1973). This may have evolved because of music's ancient association to dance, ritual and ceremony. Just as in dance, music-related movements can act as an indicator of expressive and emotional states. The communication between individuals functions as a catalyst of social bonding and allegiances to one's community (Huron, 2001). Levitin (2008) argues that the genesis of musical activity occurred early in our history during periods of challenging manual group labour. With the transition to *Homo sapiens* came enhanced cognitive and social flexibilities (Cross, 2003), and these new abilities helped us to deal with the complexities of social living. Work songs not only consisted of synchronizing movements, but they also augmented the capacity for social interaction and contributed to the forming of larger working communities.

The work produced for this thesis stems from the idea that our understanding of musical experiences is governed by how our sensorimotor capacities, bodies and environment interface with each other. This is the embodied approach to music cognition. In this approach, the human body is thought of as a significant mediator between our sensorimotor capacities and physical environment. Body movements performed during the production and/or perception of music seem to embody one's musical intentions. Some have argued that musical meaning arises from body movements, known as gestures (Cadoz & Wanderley, 2000), when deemed significant by an agent (Godøy & Leman, 2010).

The embodied music framework proposed by researchers such as M. Leman (2008), R. I. Godøy (2010) and A. R. Jensenius (2007) is based upon a rethinking of the nature of the mind within various fields such as phenomenological philosophy (e.g., Merleau-Ponty, 1945), artificial intelligence (Anderson, 2003) and cognitive neuroscience (Lakoff & Johnson, 1999; Gallese & Lakoff, 2005). In the case of neuroscience, theories of embodied cognition are supported by the relatively recent discovery of mirror neurons (Gallese et al., 1996). Mirror neurons in the premotor cortex are activated when performing or viewing goal-directed actions. This implies that we run simulations of others' actions in our own brains in order to understand intentionality (Grafton, 2009). Knowledge of the world is thus linked to our sensorimotor system (Gallese & Lakoff, 2005). These findings have significant implications for music research as music-related activities rely heavily on the brain's facilities dedicated to sensory integration (see Schutz & Lipscomb, 2007), action/perception coupling (Gibson, 1979/1986) and the ability to entrain with others (Clayton, Sager & Will, 2005; Keller, 2008).

1.1.1 Technology-driven research

By now, the value in studying body movement in music-related activities should seem evident. However, for a long time the modern era of music research neglected the integral role that movement plays in music. In 1936, Carl Seashore outlined a framework for studying music psychology thus:

> All that is conveyed from the musician to the listener as music is conveyed on sound waves. Countless other factors–such as dramatic action, gesture, grimaces, smiles and frowns, picture hats and jewellery, personal charm, environment and audience–all contribute to the pleasure or displeasure in the musical situation; but they are not music. Recognition of this fact simplifies our problem. (Seashore, 1936, p. 22)

At first this quote seems like a dismissal of what is now commonly argued—that dramatic actions and gestures do have musical qualities to them, and have such a strong impact on observers that they affect the musical experience. However, in a footnote to the above quote, Seashore stresses the need for *qualifying statements* when making such affirmations:

> The reader will do me a kindness to assume that qualifying phrases could be added for this and other direct and categorical statements…Such phrases as "other things being equal", "as a general principle", "subject to exceptions in minor detail," "in our present state of knowledge," etc., should be understood throughout. (Seashore, 1936, p. 22)

Rather than being a dismissal of movement and other types of stimuli within music performance, it now seems that these are simply factors beyond the scope of his present research question. He is therefore describing a self-imposed restriction as a means to simplify his topic, and to provide a placeholder for the future. In this respect, perhaps the most telling phrase in the last quote is "in our present state of knowledge".

Technology has certainly changed since the days of Seashore's studies. Currently, it seems intuitive that we should use all of the new tools that are available to catch up for lost time by studying musical behaviour such as the gestures, actions and expressive body movements that pervade all music-related activities. However, it is also of great importance to first develop a framework in which to study such phenomena. Some areas, such as the one focusing on musical gestures, have already taken significant strides to developing what R.I. Godøy calls a *conceptual apparatus.* This apparatus consists of having a "preconception of what we are looking for, including both a typology of sound-related gestures and an overview of gesture-pertinent features of musical sound, before making choices regarding the setup of observation studies, experiments, choosing motion capture technologies, and so on" (Godøy, 2010, p. 103).

The conceptualization of a perfect research framework for the investigation of (what I will henceforth refer to as) *music-related movements* is challenging and cannot be fully realized by a single individual. However, the main aim of this thesis is simply to showcase some of the methods I have used over the last four years for investigating music-related movements in terms of *music-related activities.* Between the six articles that make-up this thesis, a focus has been placed on four music-related activities: piano performance, conducting an ensemble, choreographed choral performance and moving freely while listening to music.

From the onset of my doctoral studies, I have been intrigued by the embodied approach to music cognition as well as its philosophical underpinnings. Likewise, optical motion capture captivates me, as it is a powerful technology that can be simultaneously used for scientific research as well as artistic purposes. These are the two main tools I have used throughout my research. Motion capture has allowed me to record with a significant level of accuracy the behaviours executed during the production and perception of musical activities. Additionally, the output of the motion capture system allows for the computation of semantically significant features, which can then be subjected to statistical testing. In brief, motion capture allows for objective descriptions of embodied musical phenomena. The main advantage of adopting embodied cognition in this project has been to go beyond the behaviourist approach by taking into account the participants' first-person experiences. The embodied approach believes "human beings to be creatures who cannot be understood without taking into account how they understand themselves" (Cayley, 2011, quote appears at 28:18 of podcast). In most of the studies presented, verbal descriptions and subjective experiences were taken into account when parsing the data. Through the balancing of objective measurements with subjective experiences, I have sought to conceive my own apparatus that is specifically for the study of music-related movements.

1.1.2 Outline of the summary

The introductory sections of this thesis summary begin with a revision of the themes and goals of the embodied cognition research program (Section 1.2) which is followed by a section that addresses embodied music cognition as conceptualized by M. Leman, R.I. Godøy and A.R. Jensenius (Section 1.3). These sections demonstrate the framework within which I have chosen to conduct my research.

Following this, there is a section focusing on music-related movement research. This encompasses an overview of musical expressivity research (an important aspect of music-related movements) and the various subfields within music psychology and perception research that have explored the role of the body in music (Section 1.4).

Because I have devoted so much time to motion capture, I have then included a section detailing the history of motion capture, its current usage within music research and some examples of how I have used this technology within my research (Section 1.5).

The aims of the thesis and the organization of the studies are outlined in detail in Section 2. Section 3 is the section in which each of the six studies is summarized. Finally, Section 4 is a brief discussion outlining the main findings of the thesis, future directions of the research, and a main conclusion to the summary.

1.2 Embodied cognition: a primer

Embodied cognition is a research approach that views the body as being central to our experience of the world. From this perspective, cognitive processes are governed by our sensorimotor capacities and their interactions with an environment in constant flux. Like other approaches to studying cognition, embodied cognition is an interdisciplinary field of research. It straddles psychology, philosophy of mind, robotics, linguistics, biology and neuroscience. These disparate fields work together to develop "cognitive explanations that capture the manner in which mind, body, and world mutually interact and influence one another to promote an organism's adaptive success" (Cowart, 2005, paragraph 1).

Because embodied cognition is the product of various research fields and schools of thought, it is often referred to it as a *research program* rather than a *theory* (Shapiro, 2007). Indeed, it is a banner term used by many researchers involved in a plethora of disciplines. The vastness of opinions and hypotheses can cause the goals of embodied cognition to seem vague or unfocused. However, one factor that unites each of its sub-fields is its critique of traditional approaches to theories of mind and cognitive processing that have permeated philosophy and traditional cognitive science (Wilson & Foglia, 2011).

Another useful way to explain embodied cognition is to compare it to its related fields. Neighbouring fields include ecological psychology (originally focused on vision), the motor theory of perception (linguistics) and some aspects of neuroscience. Gibson's ecological psychology and in particular its notion of *affordances* is relevant for the embodied approach (Gibson, 1979/1986). A person's treatment of environmental phenomena is influenced by his or her inherent or learned knowledge, past experiences and current mood. Likewise, the motor theory of perception, particularly its treatment of action/perception couplings, is also relevant to the embodied approach (Liberman & Mattingly, 1985; Godøy, 2009). The action/perception mechanism allows for our actions and perception of the world to be mutually influenced by one another. Embodied cognition can thus be studied empirically and ecologically by observing action/perception couplings within real-world *situated* activities such as sports or driving a car. Both these activities require the 'online' tracking of environmental changes (Wilson, 2002). Lastly, advocates of embodied cognition claim that its validity is in part supported by advances in neuroscience. The discovery of mirror neurons has demonstrated that the motor cortex is activated not only by our own actions, but also when viewing the actions of others (Gallese et al., 1996; Gallese & Lakoff, 2005; Grafton, 2009). This has led many to speculate that the understanding of one another's actions is manifested through our own mental simulations. Neuroscience that focuses specifically on music cognition is currently a burgeoning field. Of interest for this thesis are the findings that deal with the cross-modal integration of visual and auditory stimuli (Chapados & Levitin, 2008). Even when lying still, it has been reported that participants who are listening to music show activation in those brain regions involved with motor action planning (Levitin & Tirovolas, 2009).

With this interdisciplinarity in mind, the following section covers the philosophical foundations of embodied cognition, and places it in the context of general cognitive science.

1.2.1 Philosophical foundations

The current state of embodied cognition as an approach to cognitive science draws much inspiration from a subset of continental philosophy known as phenomenology. Phenomenology came to rise in early twentieth century Europe with its main proponents including Edmund Husserl (1859-1938), his student and colleague Martin Heidegger (1889-1976), as well as Maurice Merleau-Ponty (1908-1961) and Jean-Pierre Sartre (1905-1980), among others. The ideas articulated by the phenomenologists stand in stark contrast to other schools such as classic dualism and empiricism. For instance, they reject Descartes' idea that the mind and body are ontologically distinct from one another, and Hume's idea that knowledge of the world is constructed within the mind based on sensory input. The opposition to such views is significant as it goes against ideas deeply rooted within the Western subconscious (such theories have origins in Platonic and Aristotelian writing, not to mention Western theological doctrines). Further, the phenomenologist perspective draws into question the traditional model used in cognitive science and artificial intelligence, in which mental processes are treated like the input and outputs of a digital computer.

To put it in very general terms, phenomenology can be thought of as the study of subjective and conscious experience. The phenomenologists argue that experience is constituted by various structures of consciousness including temporal awareness, spatial awareness, attention, memory, and embodied action. For Husserl, these and other structures of phenomena are experienced as the things in themselves, rather than experienced as representations of the things in themselves (Cooper, 2001). Our experience of phenomena is therefore not a reconstruction of a world outside ourselves, but rather, we are fully integrated in the experienced phenomena. This idea could be contrasted with Hume's notion of *impressions*, which emphasizes that our experience of reality is a psychological construct, built upon sensory input from the outside world (Frame, 2009). Merleau-Ponty extended the ideas of phenomenology by focusing on the human body and its central role in our experience of the world. Given that the body lies at the centre of experience, there is no distinction between the self and corporeality. Rather, the self is constructed through corporeal interactions with the world (Cayley, 2011).

Phenomenology has been influential in various areas of research that concern embodied cognition. From linguistics, Lakoff & Johnson (1980; 1999) have argued that much of our understanding of the world is guided by the use of metaphors. Such metaphors can form relations between physical and abstract concepts. For instance, we make use of physical embodied domains, such as space, to describe more abstract ones, such as time. A more radical application of phenomenology appears in extended cognition, proposed by Andy Clark and David Chalmers (Clark, 1997). This idea proposes that our cognition truly extends to the outside world and that objects such as notebooks, or even iPhones, can be counted as constituents of the mind.

The phenomenological perspective has also garnered attention in neuroscience. Neurophenomenology, championed by Francisco Varela (1946-2001), pragmatically studies how the brain deals with consciousness (Varela, 1996; Varela, Thompson & Rosch, 1991; Thompson, 2007). As an ambitious research program, it seeks to close the explanatory gap between the phenomenal character of experience and the physical nature of the brain (Bayne, 2004).

1.2.2 Cognitivist approach to cognition

As was mentioned in the previous section, it is the cognitivist/classicist approach that is the most established within cognitive science and traditional work on artificial intelligence (AI). Under this framework, mental processes are envisioned using the simplest metaphor of a digital computer—the mind functions as a central information processor of inputs and outputs abstracted from the physical environment (Cowart, 2005). A key aspect within cognitivism is the concept of representation (Anderson, 2003). Knowledge, features and processes within the environment are represented in cognitive models as symbols. Within the model, these symbols are manipulated and compared according to rule-based transformations assigned to the model. The rule-based transformations applied to symbols are meant to simulate how mental processes such as perception, memory and reasoning function (Atkinson et al., 2000, p. 17).

Jensenius (2007) points to some rule-based approaches that have been successfully implemented to explain attributes of musical structure. Such work has resulted in models used to explain the generative theory of tonal music (Lerdahl & Jackendoff, 1983), melodic peaks (Eitan, 1997) and melodic similarity (Hewlett and Selfridge-Field, 1998). One drawback to these models is that although they "reveal many aspects of musical structures and form across large collections of data", they "often fail to locate and explain structures that fall outside of the rules being used in the model" (Jensenius, 2007, p. 14).

Models constructed within the cognitivist approach have therefore been said to model cognition in the *narrow sense* (Wilson & Foglia, 2011). A central concern that advocates of the embodied approach have with traditional AI models is that they refer to a central information processor abstracted from bodily mechanisms of sensory processing in relation to the exterior environment. The embodied approach instead focuses on cognition in the *broad sense*. Our processing is not insular but reliant on the body's features as well as our sensorimotor capacities.

1.2.3 Embodied cognition in comparison

Perhaps the most important contrast between the cognitivist and embodied approaches has to do with problem solving and retrieval of knowledge. In the cognitivist view, computations are made prior to action taking place. Actions are planned and decisions are made based on the passive retrieval of prior knowledge. In the embodied view, cognition and knowledge of the world is constructed on the basis of goal-directed actions taking place in real-time. Learning to perform certain actions through the exploration of the environment leads to the understanding of one's perceptual motor abilities and later more complex cognitive tasks such as language.

The differences between each approach can be illustrated using an example from baseball. Baseball players are particularly good at predicting where a fly ball will land. From the cognitivist perspective, outfielders predict where a ball will land based on a mental model of the ball's trajectory (Fink, Foo & Warren, 2009). However, this involves performing complex calculations involving the arc, acceleration and distance of the ball—and humans are unable to accurately discriminate accelerations (McBeath, Shaffer & Kaiser, 1995). Additionally, when observing baseball players, it is clear that they track the ball's trajectory by continually changing their own position. This strategy enables their perspective of the ball's trajectory to appear linear with respect to their current position. Whenever the ball appears to be curved to the ground, the players must reposition themselves. Clark (1999) notes the differences between both cognitive strategies. The traditional model involves a linear processing cycle: perceive, compute and act. On the other hand, from the embodied perspective the problem of where the ball will land is not a temporal one, but a spatial one. The issue is not solved or predicted ahead of time but worked out in the context of the changing conditions of the environment. Fink et al. (2009) conducted a comparison of the main competing theories that explain the *outfielder problem* using a virtual environment. Participants had to catch virtual fly balls whose trajectories had been synthetically perturbed. Based on the participants rapid responses to mid-flight perturbations, Fink and colleagues concluded that the participants' perception of the ball's trajectory was based on a continuous coupling of visual information and movement, and that no mental representations of the ball's trajectory were involved.

1.2.4 Online processing

The outfielder problem is an example of online *situated cognition* (Pfeifer & Scheier, 1999; Clark, 1999) that "takes place in the context of inputs and outputs" (Wilson, 2002, p. 626). Within these contexts, the action/perception couplings occur in quick succession as the environment is navigated and augmented. Physical activities such as sports and music-making constitute apt examples of such embodied situated cognition, because they are activities requiring a motoric interface with the environment. In this respect, these activities are akin to predator tracking behaviour—we make online decisions based solely on the conditions of the environment and on what is possible through corporeal articulations and perceptual systems. This observation reveals that embodied cognition may account for our own evolutionary history. An organism learns about its environment through motor-based exploration according to the limitations of its sensorimotor and perceptual systems. Anderson argues, "our powers of advanced cognition vitally depend on a substrate of abilities for moving around in and coping with the world which we inherited from our evolutionary forbears" (Anderson, 2003, p. 126). In this regard, our ability of abstract reasoning maintains evolutionary continuity.

Primal goal-directed actions constitute the basis for higher-level cognitive processes such as language (Cowart, 2005). The basis for such an assumption is found in the breadth of work on developmental psychology. For example, Thelen and Smith (1994) found that infants learning to reach out for toys faced unique and individual challenges based on their body mass and pre-reaching behaviour. Strategies for attaining the fine motor control needed to reach for toys were equally individualistic.

1.2.5 Offline processing

It should be stated that much cognition could be categorized as being *offline*—after all, the ability for abstract thought is a cornerstone of human cognition. Offline cognition denotes the types of processing that occur in the planning of the future, the remembering of the past or the imagination of an alternate reality. It has been suggested that this type of processing is also body-based (Wilson, 2002). An example of this might be counting on one's fingers. This can be done with large visible gestures of holding down one finger at a time. However, this action can also be primed without the need of any physical movement. A case could be made that abstract thought actually relies greatly on sensorimotor representations. Through the imagery of real objects within the environment, simulations can be run in accordance with the current task. For example a music student in musicianship class might be asked to sight-sing a melody using solfège. One might speculate that the student's strategy would involve the mental image of piano keys and playing the melody on the piano as each interval is sung—or, if the student is a trumpet player, moving his or her fingers as if pressing down the valves on a trumpet. This could be described as symbolic off-loading—in the sense that we outsource cognitive work to external resources.

1.2.6 Conclusion for this section

The embodied paradigm constitutes an alternative viewpoint from traditional cognition models. Like the older paradigm, it is rooted in a philosophy of the mind, but one that focuses on corporeal interaction with the physical world, even for the purposes of abstraction—as opposed to a mind model that uses arbitrary symbols to represent a world abstracted from itself. Although it has become popularized in the past decades, it remains a relatively young field with many perspectives (Wilson, 2002).

The ideas brought forth by embodied cognition are immediately applicable to many music-related activities. Music is an activity reliant on the acquisition of fine motor skills and constant readjustment of body movements to fit the desired musical expression. When performing in an ensemble, cognitive work and focus must be distributed across one's playing and the playing of others. Even in performances involving a conductor or a group leader, mutual adaptation and the tracking of others' behaviour is required on the part of leader and follower. The following section will explore how embodied cognition has been applied to music, and how its proponents have developed an alternative approach to musical meaning formation based on embodiment.

1.3 Embodied music cognition

Embodied music cognition applies the above approach to music by investigating the role of the human body in music-related activities. In this view, the body acts as a mediator of musical engagement and intention between the mind and physical environment (Leman, 2008). This form of engagement has three levels: synchronization, embodied attuning and empathy. Synchronization can refer to the imitation of lower level features of music, such as the sonic pulsations embedded in the audio stream. An example of this occurs when a dancer parses music's metrical structure through their movements.

Embodied attuning refers to the corporeal imitation of higher-level music features including melody and harmony. This level of engagement draws upon what Leman refers to as an action-oriented ontology, which could be thought of as having to do with an individual's personal history. In this scenario, the dancer may draw attention to musical cues that he or she perceives as significant. Lastly, empathy refers to the participation, identification and understanding of the music's emotional expressiveness. Leman argues that empathic engagement can be studied using an embodied listening framework in which participants express perceived music through corporeal articulations (Leman et al., 2007), which can be investigated in terms of emergent action/perception couplings. In short, the levels of engagement draw from the notion that our perception of music is related to a tendency to imitate its features.

Embodied music cognition is constituted by and related to several areas of research. The concepts of joint action (Keller, 2008; Sebanz, Bekkering & Knoblich, 2006) and shared representations (Tomasello & Carpenter, 2007) lend themselves nicely to embodied music cognition. Music performance relies on the coordination of individuals working towards a common goal. When driving a car, we must navigate through traffic. We use our past driving experiences to create simulations of another car's future actions. Our driving history also informs our own driving behaviour. Action/perception couplings extend to musical activities as we use our perceptual system to navigate through printed scores or rehearsed songs while shifting attention between our own playing and that of our co-performer's. The process of joint action is to "link two minds to the same actualities" (Sebanz et al., 2006, p. 70). In music this has implications for both performing and perceiving music. For conductors, their gaze and/or gestures as they begin a performance initiate coordinated action between the musicians. And such communicative gestures also occur between musicians in small ensembles, jazz combos and rock groups. Once engaged in music production, the conductor or group leader coordinates ongoing joint actions while taking into account a barrage of perceptual information (i.e., flat instruments, late entries, etc.). Such interactions are the key to productive and successful performances.

Many music psychologists and researchers have utilized an embodied approach without ever referencing embodied cognition. For instance, Eric Clarke based his ideas concerning ecological listening on the theory of ecological psychology, developed by J. J. Gibson (Clarke, 2005; Gibson, 1979/1986). The ecological approach holds that structure exists in the environment and is not a construction of the mind (Spiegelberg, 2006). Rather, we understand the environment based on the affordances granted by objects. Our understanding of the environment is formed through a perpetual coupling of actions and perceptions. Applied to music, this means that a sound can contain a number of affordances onto which we project meaning. For instance, our auditory system is able to discriminate between different sounds, and has evolved with the capacity to inform whether a sound has a threatening connotation or not. If a sound does pose a threat, the action taken in light of the perceived threat is equally determined by the affordances of the current environment (Jensenius, 2007; Bregman, 1990).

The following sections highlight three key aspects of embodied music cognition: music mediation technology, descriptions of musical engagement and musical gestures.

1.3.1 Music mediation technology

Embodied music cognition not only highlights the role of the body in communicating musical intentions, but also acts as a framework for developing real-time interactive systems, in which musical intentions are mapped to technology (Leman, 2008). As technology progresses, so do the possibilities for producing and interacting with music. For music information retrieval, it has been suggested that because music-related movements carry information about genre, expression and emotion, body movements could be used as inputs to search, navigate and retrieve from music information libraries (Godøy & Jensenius, 2009).

Another application of embodied music cognition lies in the design of digital musical interfaces. One of the challenges in designing such instruments is to develop control parameters that appear natural and intuitive to the performer. For a novel device to be successful, the mappings between the sensor technology and sensorimotor feedback should take into account human cognition and be sensitive to the expressive information communicated via the performer's gestures (Levitin, McAdams & Adams, 2002).

Whether mapping movements to technology for navigating libraries or actually performing music, a major problem lies in how to recognize musical intentions via corporeal articulations, which are inherently varied and individualistic. Leman (2008) expressed the problem as identifying the semantic gap between subjective and objective descriptions of musical engagement.

1.3.2 Descriptions of musical engagement

In some ways, embodied music cognition represents a novel approach to musicology. Leman argues that traditional musicology is overly reliant on subjective *first-person* descriptions to explain musical meaning. Embodied music cognition applies empirical methods to musicology by emphasizing the scientific method and evidence-based research (Leman, 2008).

In Leman's view, first-person descriptions of music rely on subjective interpretations and the framing of an argument within a historical or cultural context. For example, music theorists use semantic interpretations interwoven with jargon from music theory to form a hypothesis that may be unfalsifiable and therefore unscientific (Leman, 2008, p. 9). The differences between traditional musicology and evidence-based empirical musicology were outlined by Cook and Clarke (2004). In a special issue of Music Perception (2006), the work of three music theorists—each working on Mozart's Sonata in Eb, K 282—was published. Cook and Clarke (2004) describe how the theorists hardly made any effort to mention or describe each other's work. In the rare instances that they did, a difference of terminology necessitated the translation of concepts from one nomenclature to another. In traditional musicology, theories are usually non-falsifiable. This isn't a problem in and of itself because quite often theorists aren't looking for empirical evidence. From this first-person perspective, two conflicting theories can co-exist without the fear of one usurping the other (Huron, 1999; as cited in Cook & Clarke, 2004). As an alternative, the embodied approach seeks to put greater emphasis on evidence-based and falsifiable methods of investigating musical meaning within musicology

While first-person descriptions are subjective and lead to symbolic/linguistic interpretations, third-person descriptions refer to objective observations. This refers to descriptions based on the fundamental attributes of music such as pitch, tempo and timbre. Third-person descriptions are generally repeatable and can be monitored systematically. While humans can monitor the salient features of music, computers can assist in extracting the lower features of music such as spectral energy and spectral flux. Furthermore, subjective involvement with music such as spontaneous movements to music, which are motion-captured and then later analysed, can be thought of as being third-person descriptions.

The goal of Leman's embodied music cognition is not to adopt a behaviourist approach to music research, but to integrate subjective experience with objective measurement. From this union arises the second-person level of description. Second-person descriptions are more than a grammatical play on words, they are based on actual corporeal articulations between two subjects. Leman sees second-person descriptions as the best candidate for filling the semantic gap between first- and third-person descriptions. Through several modes of communication ranging from low-level, non-verbal body movements to vocalizations, private first-person experiences are transmitted from one subject to another, as well as third-person experiences shared in the music, using specifically second-person descriptions. These are descriptions tailor-made for situated activities between two individuals. Just as a significant part of a doctor's diagnosis is based on the descriptions revealed in a patient's words and behaviour in the second-person context of a doctor's appointment, so are situated activities like music-making reliant on second-person descriptions for the coordinated activity that it requires. Second-person descriptions could be said to appear in music when a teacher is showing their student proper technique. Here, the teacher is articulating the desired movement through both gestures and vocal instructions. Another example occurs when conductors use corporeal articulations to communicate tempo and dynamics to an ensemble, or when members of a band communicate through gestures to each other. Gestures are particularly useful in second-person descriptions of musical engagement as they are corporeal articulations to which musical meaning has been attributed by an agent (e.g., an observer). The next section examines the nature of musical gestures and how the studies within this thesis are organized according to the various functional roles of gestures in music-related activities.

1.3.3 Musical gestures

Gesture is a term addressed by many areas of music research. Two volumes, edited by Anthony Gritten and Elaine King, demonstrate the multiple directions that research on musical gestures can take (Gritten & King, 2006; 2011). The core definition of a gesture, across most disciplines, is "a movement or change in state that becomes marked as significant by an agent" (Gritten & King, 2006, p. xx). For a movement to gain significance, an agent is required to interpret its meaning and intention. This definition recalls Gibson's concept of affordance. In Gibsonian psychology, an object can take several meanings and be interpreted in several ways according to the agent's prior experience with said object. Gestures can also be thought of as representations akin to linguistic semiology. Ferdinand de Saussure's *sign model*, for instance, explains the semiotic relationship between a dyadic unit composed of a signifier and a signified. In this pairing, the signifier is the form the sign takes while the signified is a represented concept (Chandler, 2006). The model demonstrates the symbiotic and egalitarian association between a mental image and a physical object. For example, the letters that form the word *tree* immediately spark a mental image of a tree, and vice versa. Within the current context of music-related movements, the role of the signifier is represented by physical gesture and the signified is represented by the gesture's semantic meaning within the context, as interpreted by either the gesture's producer or perceiver.

Gestures are often used when speaking to someone. Together, gestures and speech form a unit of meaning, said to be *co-expressive* and *not redundant* of the speaker's ideas (McNeil, 2005). For instance, gestures can support and even contradict speech (Goldin-Meadow, 2003, p. 3). Although their primary purpose is to convey thoughts and ideas, gestures are performed in speech even when we cannot see the person we are speaking with. It is therefore understood that gestures are used for self-stimulation when we attempt to articulate mental images of our own. In music, it has been observed that expressive gestures are performed whether an audience is present or not. For instance, Glenn Gould continued performing in the studio with a tremendous array of expressive gestures well after he had retired from concert life. Using the notion of gestures to explain embodied music cognition is advantageous as it merges movement and meaning, which would otherwise be thought of as distinct concepts (Jensenius et al., 2010). Gestures transcend "the Cartesian divide between matter and mind. In that sense, the notion of gesture provides a tool that allows a more straightforward crossing of the traditional boundary between the physical and the mental world" (Jensenius et al., 2010, p. 13).

The literature on musical gestures is marked with numerous definitions and categories that complicate this seemingly simple concept. For instance, Kendon's (1982) gesture typology consisted of *gesticulation, emblems, pantomime* and *sign language.* McNeil (1992) later noted that these types of gestures could be organized on a *continuum* whereby one extreme is occupied by gestures that require speech for their meaning to be clear (gesticulations) and the other extreme by gestures that are linguistically self-contained (sign-language) (Jensenius, 2007, p. 37). Gestures specific to music-related activities can also be classified. One basic categorization scheme for music performance is to differentiate between gestures used for controlling an instrument (sound-producing) and gestures used for communicating musical ideas (empty-handed gestures). Some authors have found it helpful to align different kinds of gestures along a continuum ranging from the most physical to the most metaphorical. Such a framework was employed by Delalande (1988) when studying the performance gestures of Glenn Gould. Originally in French but translated into English by Wanderley (Wanderley et al., 2005, p. 97), the three-tiered gesture topology consisted of the following:

- *Effective gesture:* Those that make the sound (fingers on piano keys, bowing of strings)
- *Accompanist gesture:* Body movements that act as subordinate or supporting gestures to the effective gestures. (Swaying of the torso, facial contortions)
- *Figurative gestures:* Sonic gestures perceived by the audience by means of the produced sound that have no direct correspondence to physical movement (Changes in notes articulation, melodic variations)

Similarly to Leman's first-, second- and third-person descriptions, the levels proposed by Delalande are organized from being purely objective (effective) to being purely subjective (figurative). Effective and accompanist gestures represent gestures for controlling a physical object and open-handed gestures for expression mentioned above. Figurative gestures refer to the associative meaning that is applied to sound gestures. In Wanderley's work on the performance gestures of clarinettists, effective and accompanist gestures are rebranded as *instrumental* and *ancillary* respectively (Wanderley, 1999; Cadoz & Wanderley, 2000; Wanderley et al. 2005). The explanation regarding the role of instrumental gestures in music performance is quite simple. They are the bodily movements that trigger musical events such as bowing, plucking, blowing and pressing a key. Ancillary gestures, on the other hand, are not involved in the production of sound but occur in virtually every performance (Wanderley, 1999). Wanderley argued that such gestures constitute the expressive component of musical performance and are akin to paralinguistic gestures. In other words, performing without ancillary gestures resembles a monotone speaker.

The categories outlined by Delalande and Wanderley represent but a small portion of the vast amount of schemes used to classify gestures. Alexander Refsum Jensenius' doctoral thesis (2007, p. 35-58) provides an excellent review outlining the different ways the term musical gesture has been espoused by musicologists, psychologists and computer hardware engineers. Based on his experiences and a careful reading of the literature, Jensenius points out the inconsistencies between fields of research when detailing what gestures are and how they can be grouped. For example, he notes that the human-computer interaction field is imprecise in its definition of gesture compared to other disciplines that are more precise and rigid in their vocabulary. The inaccuracies amongst and incompatibilities between the fields of research are problematic as the study of gestures and their applications are often the subjects of multidisciplinary research projects. To circumvent confusion, Jensenius elected to avoid using the term gesture at all and instead adopted four singular terms to describe music-related movements (Movement: a physical displacement; Action: a self-contained movement unit or chunk that is goal-oriented; Fidgeting: random unintentional movements; Interaction: movements which have a mutual influence on one another).

Music-related movements can be categorized according to their function within a music-related activity. Jensenius suggested a framework based on the work of Cadoz and Wanderley (2000), Delalande (1988) and Wanderley et al. (2005), which I also intend to use. The following are therefore paraphrased definitions of the functions from Jensenius (2007):

- *Sound-producing actions* are movements that produce sound. A further distinction subdivides them by their function into movements of excitation or modification, although these will only be touched upon here.
- *Ancillary movements* support sound-producing actions. They are movements not involved in the production of sound but important for musical expression.
- *Sound-accompanying movements* mimic or follow musical structure. For the purpose of this thesis, I will explicitly include dance into this category because, although it can sometimes be far-removed from musical structure, dance usually follows music's meter and tempo.
- *Communicative movements* are movements used explicitly for communication from one performer to another. In this context, conductors will fall into this category.

These categories were originally applied to only music performers, but I have reinterpreted them slightly to fit the organization of my thesis. This is further explained in Section 2.

1.3.4 Conclusion for this section

Embodied music cognition has provided an excellent theoretical framework in which to conduct my doctoral research on music-related movement. Within this topic, I have particularly focused on four music-related activities: piano performance, conducting an ensemble, choreographed choral performance and moving freely while listening to music. Literature surrounding each of these activities will be discussed in the next section.

1.4 Music-related movement

While the previous two sections on embodied cognition and specifically embodied music cognition presented some background information concerning the framework of my research interests, the current section presents literature that is immediately relevant for the particular context of each of the thesis articles. The music-related activities that have been researched throughout my doctoral studies vary in these terms, as well as in their respective aims. Unsurprisingly therefore, a unique body of research literature surrounds each activity, which it is the aim of this section to present.

A good starting point for such an endeavour is to cite research on musical expression, as the communication of expressive intentions is a common area of interest for each of the articles. The current model of musical expression in music psychology research is that of a multi-dimensional construct in which auditory cues such as loudness, tempo, melodic articulation, etc. can be altered independently according to musical goals and aesthetic sensibility. As a visual cue, body movements have not traditionally been considered a parameter of musical expression. However, as the work of Jane Davidson has shown, body movement is in fact one of the most salient features when communicating expressive intentions. This discussion on musical expression and emotion research (section 1.4.1) applies to all the thesis articles, but in particular, to Studies I, II and VI, which adopted frameworks commonly used in musical emotion research. The subsequent sections then go on to discuss the themes highlighted specifically in the thesis articles: namely, the perception of expressive body movements (section 1.4.2); the production of expressive body movements (section 1.4.3); co-performer communication and synchronization (section 1.4.4); and music-induced movements (section 1.4.5). Near the end of each section, I have indicated to the reader the ways in which the section's topic relates to one or more of the thesis articles.

1.4.1 Communicating musical expression and emotion

Research and scholarship surrounding musical expression has traditionally related to the audible characteristics of music. Fluctuations in amplitude, tempo and timbre constitute the core of the auditory parameters that are manipulated by performers to match their expressive goals. For Westerners, this inclination towards auditory parameters is quite natural as the dynamic markings on scores traditionally instruct the performer how the performance should sound as opposed to how the performer should move. For a listener, musical expression greatly influences their perception of the performance. Patrick Juslin emphasizes the importance of expression in music performance (Juslin, 2003), stating that:

> It is expression that makes people go through all sorts of trouble to hear human performances rather than the "dead-pan" renditions of computers; it is expression that makes possible new and insightful interpretations of familiar works; and it is on the basis of expressive features that we prefer one performer rather than another. (p. 274)

From this quote, it appears that our enjoyment of music depends on how we perceive musical expressivity. Despite emotional expressivity being a central concern of musical aesthetics, definitions of musical expression are varied. Traditionally, musical expression is written about as something shrouded in mystery and historical literature often alludes to its transcendent and ineffable qualities. Examples of this dominate contemporary descriptions of the great Romantic virtuosos such as Liszt, Paganini and Chopin. The poet Moritz Saphir described Franz Liszt's performances as "an inexplicable phenomenon, a compound of such heterogeneous, strangely mixed materials, that an analysis would inevitably destroy what lends the highest charm[:] the individual enchantment…of whimsy and divine nobleness" (*Allgemeine musikalische Zeitung,* May 1838, as cited in Weiss & Taruskin, 1984, p. 365). Another example comes from an unnamed London critic, who described Chopin as a performer who "seems to abandon himself to the impulses of his fancy and feeling, to indulge in a reverie and to pour out unconsciously, as it were, the thoughts and emotions that pass through his mind" (Hedley, 1947, p. 107; as cited in Weiss & Taruskin, 1984, p. 371).

If anything, these passages are wonderful descriptions of cathartic musical experiences. But, they also imply that trying to understand musical expressivity empirically is a fool's errand. Even modern musicians are reluctant to submit their sacred craft to testing (for some perspectives from modern musicians, see Dubal, 1985). These traditional opinions of music expressivity have influenced the way in which some modern research is conducted. For example in music education research, which often uses music student surveys, music expressivity is defined as a "homogenous category, which there is more or less of'" (Juslin, 2010, p. 454; writing in reference of Marchand, 1975). This characterization of musical expression is not entirely unhelpful—people are generally able to understand what playing *expressively* or *playing without expression* means[1]. However, it paints a less than perfect picture of what musical expression actually is.

---

[1] For music students, musical expressivity gets associated with concepts such as 'playing with feeling' (Lindström et al., 2003). This is sometimes achieved by incorporating memories of emotional life experiences in their playing (Woody, 2000).

In the past quarter century, music psychologists have taken up the subject of musical expressivity and have been more systematic in their definitions. Palmer states that musical expression refers to “the large and small variations in timing, intensity or dynamics, timbre, and pitch that form the microstructure of a performance and differentiate it from another performance of the same music” (Palmer, 1997, p. 118). Musical expression should thus not be viewed as a homogenous scalar, but rather a multi-dimensional set of descriptors covering different modes of perceptual experience. This makes sense because we are able to perceive changes in performance parameters such as tempo, dynamics and timing independently from each other (Gabrielsson & Juslin, 1996).

Juslin (2003; Juslin & Timmers, 2010) decomposed musical expression into five “subcomponents that make distinct contributions to the aesthetic impact of a performance” (Juslin & Timmers, 2010, p. 454). Known as the GERMS model, the components refer to: *generative rules* (conveying the musical structure to a listener)*, emotional expression* (conveying expressive intentions)*, random fluctuations* (indeterminate fluctuations that make each performance unique)*, motion principles* (music perceived as movement)*, and stylistic unexpectedness* (violating musical expectations with the goal of conveying expressivity). These components work together to construct pleasing performances that are appropriate within a stylistic or social context.

Perhaps the component of the GERMS model most relevant to this thesis is the one that deals with motion principles. This aspect of musical expression draws from the assumption that music is perceived as motion. Some researchers have shown that the perception of motion in music is related to biological motion (Friberg & Sandberg, 1999; Shove & Repp, 1995). Juslin thus argues that aesthetically pleasing performances should “reflect the dynamic patterns of movements that are characteristic of humans” (Juslin, 2003, p. 283). In other words, performances are deemed to be expressive if the rhythmic motion throughout appears natural and within the constraints of biological motion (Shove & Repp, 1995). Models based on musical motion have revealed that the perception of music as motion (e.g., rhythm) originates in the “kinematic and dynamics characteristics of typical motor actions” (Palmer, 1997, p. 132).

Much of the empirical research carried out on musical expressivity in music performance stems from the performer’s ability to communicate his or her emotional and expressive intentionality to a listener. The research has focused as much on the production of emotional states as on the perception of emotional states. The communication of emotional intentions of music does not depend on *felt* emotions, but rather on the *portrayal* of emotional states on the part of the musicians and the *recognition* of portrayed emotional states on the part of the perceiver (Juslin & Timmers, 2010, p. 455). Although this frames emotional communication as not existing beyond a symbolic level, the inability to satisfyingly communicate emotions is considered a major detriment to an otherwise technically flawless performance. The aims of the GERMS model is thus to demonstrate how the components of musical expression function together to create aesthetically and stylistically appropriate performances.

A common framework used to examine the communication of emotional portrayals and expressive intentions has come to be known as the *standard paradigm* (Juslin & Timmers, 2010). Within this framework, musicians are requested to perform an excerpt while employing a specific emotional state or expressive intention. Participants then evaluate the performances based on their perceptual experience of the music using forced choice scales (Kotlyar & Morozov, 1976) or continuous responses (Schubert, 2004; Schubert & Dunsmuir, 2004). Commonly, musical features can be extracted either manually or computationally from a recording of the performance. If there are any differences between performances representing different emotional states, these can be tested statistically against the perceptual ratings.

The types of scales used in these studies depend on how the researchers have chosen to conceptualize musical emotions. The two main theoretical frameworks in which musical emotions has been studied are the discrete emotion model and the dimensional model of affect (Eerola & Vuoskoski, 2010).

The discrete emotion model is based on innate or basic emotions such as fear, anger, disgust, sadness and happiness (Ekman, 1992). This model is based on the assumption that each basic emotion has a supposed independent neural substrate—the evidence for which was first brought forth by Ekman and Friesian (1971). After observing that indigenous people of New Guinea were able to identify and differentiate between the emotional expressions of Caucasians, Ekman and Friesian argued that the ability to observe basic emotions was an innate skill across all humans. However, most music research on emotions uses a modified version of the basic emotion model to account for the emotions portrayed in music[2]. Indeed, in a recent study, it was found that musicians preferred to rate the emotional qualities of performances by using illustrative and detailed descriptors (selected from Hevner's (1936) Adjective Circle) than by using the five standard emotions (Huang & Krumhansl, 2011).

The dimensional model of affect, based on Russell's circumflex model (Russell, 1980) is built upon a different assumption, proposing "that all affective states arise from two independent neurophysical systems: one related to valence (a pleasure–displeasure continuum) and the other to arousal (activation–deactivation)" (Eerola & Vuoskoski, 2010, p. 3). In this model, different combinations of valence and activity represent the two scales from which all emotional states arise. A suggested third dimension to this model is power (weak–strong) (e.g., Osgood, Suci, & Tannenbaum, 1957). In this model, each dimension is a continuum that together describes a semantic space in which emotional polar opposites (e.g., happy–sad) appear suitably far apart. This implied dynamism is particularly beneficial for investigating music, as an excerpt's emotional inflection can change from moment to moment. Thus, researchers employing continuous response paradigms have found the dimensional model particularly beneficial to their research.

---

[2] The Geneva Emotion Music Scale (GEMS), for example, consists of wonder, transcendence, tenderness, nostalgia, peaceful, power, joyful activation, tension and sadness (Zentner,, Grandjean & Scherer, 2008).

For example, Schubert (2004) asked participants to continuously judge the emotion perceived in a musical excerpt by using a computer mouse to position the cursor at any point within a two-dimensional emotion space (the x-axis represented valence and the y-axis represented arousal), and the positions were then sampled once per second. In this way, the resulting data consisted of a time series representing the participant's real-time perception of the strength and quality of emotion in the music. Schubert then extracted features from the audio signal related to the music's loudness, tempo, melodic contour, texture and spectral centroid. These features were then entered into a linear regression model to determine which parameters might be predictors of the perceived emotion. Through this method, Schubert was able to tease out the prominent musical features that correlate with a musical piece's emotional intentionality.

Considering the crucial role body movements play in music performance, it is interesting to see that its role has been underestimated in research on musical expressivity and emotion. This thesis has made use of both the standard paradigm proposed by Juslin (Studies I & VI) and a continuous rating framework similar to Schubert's (Study II) to study expressive body movement.

1.4.2 Music performance: perceptual studies

Early work on the perception of expressive body movement in music performance was influenced by research on gait and biological motion. Johansson (1973, 1976) showed that common human motion patterns such as walking, running and dancing could be represented with as little as five points of light representing the motions of the main joints. Johansson's research employed a method whereby reflective ribbons were attached to the joints of actors. Their movements were then recorded while being flooded with lights mounted close to the camera lens. During playback, the TV monitor's contrast was set to maximum, so only the reflective ribbons were visible. Under these conditions, observers were able to identify the human motion patterns within 500 ms of exposure to the footage.

The method was adopted by Cutting and Kozlowski, who showed that point light displays of an actor's walking pattern provided enough information to identify both the actor's gender (Kozlowski & Cutting, 1977) and their identity (Cutting & Kozlowski, 1977). These studies revealed that biological motion was both gender specific and highly individualized.

Johansson's point light display method was later used to show that biological motion contains perceptual information regarding covert mental dispositions. Runeson and Frykholm's (1983) Kinematic Specification of Dynamics principle states that mental dispositions are perceivable to observers through the dynamics of their motions. For example, in an experiment, actors were requested to lift boxes and throw sandbags. Observers were shown point-light displays of these actions and were able to discern the weight of the box and how far the sandbag had been thrown. Additionally, they were able to detect whether the actor was being deceitful concerning the weight of the box. Runeson and Frykholm showed that biological motion was sufficient for perceiving the underlying dynamics (masses, forces, momentum) from the point-light displays.

The framework employed by those studying biological motion also influenced Davidson's seminal 1993 study on expressive body movement. Davidson asked musicians to play in three different performance manners: *deadpan, projected* and *exaggerated.* The methodology is reminiscent of Juslin's (Juslin & Timmers, 2010) notion of the standard paradigm, but this time applied to body movement. The performance manners were based on teaching devices. The deadpan manner encourages students to perform note-accurate renditions of a musical work while the exaggerated form encourages students to overstate expressive features such as timing and dynamics. Interestingly, the instructions are to alter their expressive and/or performance goals, and not their movement patterns—and movement is never mentioned in the instructions given to musicians. The study included two experiments. In the first, observers judged the performances of four violinists and in the second, the performances of a solo pianist. Participants observed performance excerpts under three conditions: sound and vision together, sound only, and vision only.

In both experiments, the pattern of expressivity ratings showed that participants correctly identified each expressive manner. The deadpan manner consistently received the lowest scores while the exaggerated received the highest scores. The gap between the projected and exaggerated scores tended to be more narrow; Davidson suggested performers might find it easier to "withhold expression from the piece than exaggerate the expressivity of a piece beyond its normal level" (Davidson, 1993, p. 109). Interestingly, the difference between the projected and exaggerated manners was largest in the vision only condition, and the overall difference in ratings was smallest in the sound only condition. The study concluded that the kinematic cues provided by the vision mode therefore play a substantial role in discriminating between a performer's expressive intentions.

In several follow-up studies, Davidson tried to ascertain which parts of the body most contributed to the perception of expressivity in movement, and in what manner. She found that larger movements performed with exaggeration were the key indicator of enhanced expressivity, resulting in higher expressivity ratings. Additionally, it was found that observing the head/torso region gave enough information to discriminate between the amounts of expression (Davidson, 1994). Later, she found that in the vision and sound conditions, visual kinematic cues were the prominent feature used by non-musicians to judge the performer's expressive intentionality (Davidson, 1995).

When focusing on the expressive movements of pianists, observers reported that their expressivity ratings were based on a pianist's tendency to sway their upper body, and on isolated gestures occurring throughout the performance (Davidson, 1991). Body sway is an interesting phenomenon as it appears to be quasi-periodic, but its exact relationship to the music is unclear. Clarke & Davidson (1998) argued that body sway had both time-keeping as well as phrase-linking functions (Davidson, 1994). Isolated gestures, on the other hand, relate to the localized features of the music being performed, as if the pianists were—through their gestures—drawing attention to a significant portion of the performance.

In a more recent study, Davidson (2007) made similar observations while observing a pianist performing a Beethoven *Bagatelle*. She identified locations in which the same expressive head and hand movements were repeated throughout each expressive manner. This suggested that expressive gestures are influenced by music's structure, as well as the biomechanical constraints involved in a piano performance. For example, the expressive movements in an exaggerated manner may have a functional origin, which can also be seen in the deadpan condition.

Many researchers have supported and expanded upon Davidson's findings by employing experimental designs that either used stimuli produced by motion capture systems (Nusseck & Wanderley, 2009; Bernardin et al., *in preparation*) or video recordings (Vines et al., 2006; Dahl & Friberg, 2007; Schutz & Lipscomb, 2007; Broughton & Stevens, 2009; Behne & Wöllner, 2011; Luang & Krumhansl, 2011). These studies have all contributed to uncovering the types of information that music-related movements convey.

For instance, in addition to expressive intentions, Dahl and Friberg (2007) found that distinct emotional intentions could also be perceived from musicians' body movements. Participants were able to recognize emotional intentions relating to sadness, happiness and anger, but not fear. The experimenters had several viewing conditions in which the video image was cropped so either only the head, torso or arms could be viewed. Participants were for the most part still successful at identifying emotional states throughout these conditions. The authors attributed the participants' success rate to their ability to reconstruct or imagine the movement cues absent in the filtered video footage. Nusseck and Wanderley (2009) presented similar results when they asked participants to watch short kinematic displays of clarinettists' gestures. The participants rated four different music-related dimensions: tension, intensity, fluency, and professionalism. The experimenters took advantage of the motion-capture-produced stimuli by manipulating the kinematic displays in various ways. Somewhat analogous to Dahl and Friberg, they produced kinematic displays in which the points representing the clarinettists' arms and torso were frozen, thus appearing not to move. Like Dahl and Friberg's study, the alterations did not affect the perceptual ratings across presentation mode. The experimenters also manipulated the size of the movement patterns. The stimuli containing smaller motions yielded lower perceptual ratings. From these findings, it was proposed that the clarinettists' expressive intentions were sufficiently communicated through overall holistic motion patterns, rather than in individual parts of the body. However, the size of the motion patterns contributed to the perceptual ratings.

Other studies have investigated body movement in music performance by comparing the perceptual differences between the audio and visual modes. Vines and colleagues (Vines et al., 2006) investigated cross-modal interactions between sight and sound. Using a continuous response paradigm similar to

Schubert (2004) and Study II of this thesis, participants provided ratings of perceived phrasing (structural content) and tension (emotional content) while viewing videos of clarinet performances (in the sound alone, vision alone and sound+vision conditions). Through the use of functional data analysis (Levitin et al., 2007), they found that vision influenced the experience of both phrasing and tension. For example, phrasing gestures performed by the musicians had a linking function, used to connect adjacent sections of the piece. The enhanced sense of musical structure and emotional content provided by the visual mode (facial gestures, postural adjustments) led the authors to suggest the existence of an emergent quality that appears only when musical performances were both heard and viewed. A recent follow-up study found that the vision mode also greatly influenced participants' ratings of a performance's emotional qualities (Vines et al., 2011).

Multi-sensory integration has further been studied by presenting participants conflicting information. Schutz and Lipscomb (2007) swapped a video recording's audio and visual components of a marimba player performing single notes. Participants were asked to judge the duration of the notes under audio-only and audio-visual conditions. The researchers found that the participants' perception of note duration was influenced by the visual component of the performer's gestures, rather than by the sound of the note, and that short notes were judged to be longer when accompanied by longer gestures. Schutz and Lipscomb suggested that sensory integration could act as a powerful perceptual tool for percussionists, since they have minimal control over acoustic note durations. Performers subconsciously use visual tricks very often. Investigations of famous musicians such as Glenn Gould (Delalande, 1988), Annie Lennox (Davidson, 2001), B. B. King (Thompson, Graham & Russo, 2005) and Keith Jarrett (Elsdon, 2006) have all emphasized the power of a performer's body movements on their audience.

The perceptual study included in this thesis (Study II) posits the notion that observers judge expressive movements based not only on the amplitude of the performer's movements, but also based on the movement's kinematic features. This study was based on previous research in which it was shown that observers are sensitive to the movement kinematics of conductors' time-beating gestures (Luck & Nte, 2008; Luck & Sloboda, 2007, 2008, 2009). Participants reported their perceived experience of emotional content while viewing kinematics displays of conducting gestures. Unlike previous research, the aim was not to have observers synchronize to the musical beat, but rather to find relationships between expressive content inherent in the conductor's gestures and the observer's perceptual experience.

1.4.3 Music performance: production studies

Perceptual studies indicate that music-related body movements have a deep impact on an observer's appreciation and understanding of the performance. Some researchers have circumvented the perceptual aspect of music performance and focused on the performances themselves. This has often entailed extracting relevant features from movement data and testing them computationally or statistically. Although this has been attempted by annotating the movements from video excerpts (Davidson, 1994), motion capture techniques have provided movement information at a much higher temporal resolution. Also, because the output of motion capture is stored as a timeseries, several researchers have adopted signal-processing techniques traditionally used outside the area of music performance (Luck & Toiviainen, 2006; Goebl & Palmer, 2009; Palmer et al., 2009).

Much of the research in music production has used paradigms similar to Davidson's expressive intentions. Palmer and colleagues (2009) asked clarinettists to perform without expression, with a regular amount of expression and with exaggerated expression. As the musicians increased their amount of expression, the movements of the instrument's bell increased in amplitude, as did the musician's expressive timing. Also, the bell seemed to embody several features from the musical structure: elevations of the bell appeared to coincide with areas of rubato and the beginning and ending of phrases.

In contrast to requesting musicians to alter their expressive intentions, Wanderley explicitly asked musicians to alter their amount of movement in three performance conditions: immobile, standard and expressive (Wanderley, 2002; Wanderley et al., 2005). Based on the analysis of multiple clarinet players, Wanderley suggested that the types of movements used in performance are related to three main factors. The first factor, physiological features, implies that the type of expressive movements musicians can engage in restricted by physical elements such as respiration, fingering and the ergonomics specific to the instrument being performed on. The second factor relates to the musical structure. Wanderley observed that expressive movements, or ancillary movements as he called them, were related to rhythmic patterns and coincided with the musical phrase structure. The third factor relates to the musician's own interpretive choices. Given the restrictions of the first two factors, the musician can use their individual 'mental model' of the piece to develop their own gestural language.

One interesting outcome of Wanderley's studies—which focus on instructions related to quantity of movement—was that even though the instructions given were different compared to Davidson's design, Wanderley observed many of the same results. Performances with more movements appeared to be more expressive; more related to the musical structure and made more use of expressive timing (immobile performances were mostly played at a faster tempo).

Study VI was greatly influenced by the work of both Davidson and Wanderley. There appears to be an interesting connection between performance instructions that focus on expressive intention and those that focus on amount of movement. Study VI was an attempt to synthesize both these sets of instructions in order to explore the relationships between amount of movement and expressivity.

1.4.4 Synchronization, entrainment and co-performer communication

Another area of research that is of interest to this thesis concerns music-related movements used by musicians to communicate expressive intentions amongst themselves. Musicians use movements to communicate musical ideas and synchronize tempo (Goebl & Palmer, 2009)—and gestural communication occurs in rehearsal settings almost as much as it occurs in performance settings (Williamon & Davidson, 2002).

Musical synchronization is possible because humans are able to perceive spatiotemporal properties of human movements as being significant. Gestures and actions are loaded with musically relevant information. The main example of this may be the perception of a musical beat through gestures performed by conductors (Luck & Nte, 2008; Luck & Sloboda, 2007, 2008, 2009). Conductors use a wide range of gestures to communicate expressivity within a composition's structure and style (Price & Byo, 2002). Their gestures must convey emotion as well as embody the current music's metrical hierarchy in order to synchronize their ensembles.

Vital to the understanding of music-related activities involving two or more individuals are the concepts of joint action and entrainment. Joint action refers to social interaction in which two or more individuals coordinate their actions in space and time towards a common goal. According to Sebanz et al. (2006), successful joint action relies on the following abilities:

- To share representations
- To predict actions
- To integrate predicted effects of own and others' actions. (p. 70).

For music-related activities, these three elements of joint action are largely accomplished through entrainment, which refers to the synchronization between two or more individuals (Clayton et al., 2005).

Traditionally, the paradigm to study entrainment and synchronization as related to musical phenomena is to have participants tap their fingers (see Repp, 2005). Although informative, this method is abstracted from realistic musical activities. This lack of ecological validity has been somewhat remedied by the growing number of studies that have used motion capture data to study synchronization and entrainment by investigating gross body movements (Keller, 2008; Goebl & Palmer, 2009; see also Keller & Rieger, 2009; Toiviainen & Keller, 2010).

In this thesis, joint action and entrainment were investigated within the context of a cross-cultural dance performance (Study V). Following the trend of studying entrainment in ecological settings, a choral group consisting of experts and novices was motion-captured. The paper's goal was to detect differences in beat perception between experts and novices by evaluating the synchrony within each group.

1.4.5 Music-induced movements

Research on music-induced movements investigates how people react corporeally to music. One does not need to be in a nightclub to move to music—most people self-report moving with music in the most ordinary of situations (Lesaffre et al., 2008). For instance, tapping one's foot and nodding one's head are common ways to move to music in social and private situations. These corporeal articulations have been described as the parsing of musical structure, and are often used by ensemble musicians while anticipating their entries, for example. In addition to movements being related to the music temporal organization, spontaneously occurring music-induced movements can also be related to the expressive intentions implied in the music. Just as a performer moves to the music they are making, so does someone listening to their music.

An obvious example of music-induced movements is dance. Commonly, dance consists of goal-directed actions synchronized to the metrical structure of the perceived music. Leman characterizes these actions as corporeal imitations of the local energy bursts perceived in the musical audio stream (Leman, 2008). These regular bursts of energy evoke our subjective sense of the musical beat, which represents the main perceived metrical level. Music frequently contains multiple beat levels that are organized in a metrical hierarchy. Meter represents the temporal grid in which the beat levels are perceived as having a strong or weak pulse (Lerdahl & Jackendoff, 1983). Usually, each beat level has a period that is an integer ratio of the most prominently perceived pulse (Palmer & Krumhansl, 1990).

There is some evidence to suggest that movement affects beat perception. Todd, Cousins, and Lee (2007) found that preferred beat rate was related to anthropometric factors, such as one's weight, and the width and length of body segments. Philips-Silver and Trainor (2005; 2007) found evidence that body movement influenced beat perception related to a specific pulse. Participants heard a six-beat rhythmic pattern with an ambiguous meter and synchronized their movements by bouncing on their knees on either every second or third beat of the pattern. The authors argued that participants were biased towards either duple or triple form based on which beat they bounced on, which highlighted a significant multi-sensory interaction between movement and auditory perception. A follow-up study attempted to tease out which aspects of movement were most critical for parsing sensorimotor information by examining passive movements (Philips-Silver & Trainor, 2008). The researchers induced passive movements by having participants lie down on a novel seesaw bed. The apparatus made it possible to induce passive movements to specific sections of the body in isolation. Isolated head movements were found to bias beat perception while isolated leg movements did not. This indicated a possible relationship between movement and auditory rhythm perception that is facilitated through the vestibular system (the sensory system responsible for spatial perception and the sense of balance).

Study III of this thesis was influenced by this literature and explored how different metrical levels might manifest themselves in different parts of the body. Based on previous work, it was expected to see movement patterns synchronized to several metrical levels. For example, based on the biomechanical properties of the body, it was expected to see higher (i.e., slower) metrical levels embodied as gross body movements (i.e., the torso), and lower (faster) metrical levels embodied as movements of the extremities.

As mentioned above, music-induced movements are not always related to the music's temporal structure, but instead mirror the music's expressive and emotional goals as envisioned by its composer or performer. Moving spontaneously or candidly to music is affiliated to notions such as *letting go* and *getting into the music.* Such occurrences are common in social situations such as concerts, parties and rituals, and since the rise of the personal music player, music-induced movements occur in complete isolation of others.

As with music performances, computational methods have been explored to help interpret the emotional and expressive content from body movements in dance performances. Camurri and colleagues, for example, developed tools for measuring and analysing the gestures of professional dancers and then classifying them according to a basic emotion (Camurri, Lagerlöf & Volpe, 2003), also see www.eyesweb.org). Other studies have targeted expressive dance movements by having modern dance students perform to music without a salient pulse, thus controlling for movements linked to the music's metrical hierarchy (Casciato, Jensenius and Wanderley, 2005). Casciato and colleagues performed an observational study, the results of which indicated that dancers' gestures tended to imitate the music's instrumentation, and that their movement trajectories correlated to low-level sound features (e.g., spectral centroid).

Besides dance, there are other types of music-induced movements. One movement category of note deals with so-called *air instrument performances*. An air instrument performance re-enacts or mimics a recorded performance. Such activity occurs spontaneously (for example, the person in the car next to you stopped at a red light, *rocking out* to some *wicked* guitar solo) or highly practiced (e.g., the Air Guitar World Championship that is held annually in Oulu, Finland). Rolf Inge Godøy, for example, has taken particular interest air instrument performances, for which he coined the term *motormimetic sketching* (Godøy, Haga & Jensenius, 2006). Godøy and colleagues argue that motormimetic sketching is indicative of the mental imagery of musical structure, and plays a role "in parsing and chunking musical sound, as well as in grasping rhythmical, textual, melodic, and harmonic patterns" (Godøy et al., 2006, p. 266).

Godøy and colleagues (2006) conducted observation studies in which participants ranging novice to expert pianists 'air performed' along to commercial piano recordings. The novices maintained synchrony with the recording and their improvised gestures matched the emotional qualities implied in the music. The experts on the other hand, performed goal-directed actions. For example, one pianist planned a hand lift towards the imaginary keyboard's higher register by first eyeing the target position and then lifting her hand towards that position. The authors observed that the gestures performed by the participants indicated their musicality and the amount of training they had had—the novices focused on emotional content and the experts on technical content. In both cases, gestures were used to *make sense of,* and parse the audio signal according to their individual musical background.

The studies mentioned here have all yielded interesting findings, but have not really addressed how people move to music in everyday situations. The actual movements induced tend to be highly individualistic and are influenced by a plethora of factors. Study IV of this thesis examined music-induced movements by relating the types of movements performed to the individual's personality traits.

Personality refers to the set of characteristics that influence our reasoning, motivation and behaviour in various situations (Ryckman, 2008). By this token, mannerisms and gestures, which occur naturally during speech, are related to and indicative of one's personality traits (Koppensteiner & Gammer, 2010). while personality judgments can be made based on gestures and non-verbal communication (Knapp & Hall, 2009). Large confident movements may indicate extroversion and Openness to experience while nervous unconfident movements might indicate introversion or Neuroticism. Perhaps the fact that music compels someone to move at all states something obvious about their personality. Individual movement patterns are overt demonstrations of one's personality characteristics.

1.4.6 Conclusion for this section

The goal of this section was to present literature that is most relevant to the types of music-related movements that have been studied during my doctoral studies. While this section dealt with the type of research, the next will focus on the technology used to conduct my research: optical motion capture.

1.5 Motion capture in music-related movement research

Motion capture, in the context of this thesis, refers to a technology in which biological motion is measured and stored for the purposes of movement analysis. Though there exist several types of motion capture technology, a focus will be placed on infrared optical motion tracking systems, particularly the Qualisys ProReflex MCU 120 system, which acted as the main data-collecting utility used in each of the studies presented in this thesis.

In the past ten years, motion capture has become an increasingly common way to store and investigate music-related movements. Early examples within this period appear in research areas studying musicians' gestures for the purposes of designing more intuitive music input devices (e.g., Wanderley, 2002). These studies were exploratory and focused on the movements of a small pool of participants. The scope and scale of research entailing motion capture has grown as more laboratories around the world have acquired their own systems. That is not to say that motion capture is an entirely democratized technology—optical infrared systems cost a great deal, which limits their access. Nonetheless labs focusing almost exclusively on music-related movements and music cognition issues are found in several countries including Finland, Belgium, Germany, Norway, Canada, USA, and Australia. An excellent demonstration of the kind of music research that has been using motion capture can be seen in recent special journal issues on music and synchronization (Keller & Rieger, 2009) and spatiotemporal music cognition (Toiviainen & Keller, 2010). Both volumes include several articles that used motion capture for projects ranging from modelling dance gestures (Naveda & Leman, 2010), through gauging co-performer interaction (Goebl & Palmer, 2009), to perceiving the facial expressions of singers (Livingstone, Thompson & Russo, 2009).

This section begins with a brief overview of the history of motion capture and is followed by a presentation of the motion capture techniques and some of the analyses used for the studies in this thesis.

1.5.1 Motion capture: a brief history

Throughout history, and depending on the needs of society at the time, interest in and systematic investigations about body motion have motivated several types of discoveries, innovations and new technologies across many fields. Additionally, the human body and its kinematic properties have inspired artists to create great works of beauty. Incidentally, developments in art, science and technology depicting or centring on human motion have consistently influenced each other over the past 2000 years. In an excellent historical overview, Klette and Tee (2008) trace the first known text concerning biomechanics to Aristotle's *On the Motion of Animals*, in which gait and locomotion is explained using mathematical properties. The authors point out that throughout classical antiquity, "an advanced level of understanding of human or animal poses" (Klette & Tee, 2008, p. 2) led to the production of realistic depictions of motion in art and sculpture. Although these depictions were static, the artists' knowledge of geometry, anthropomorphic properties and use of perspective led to the production of works that conveyed the subject's motion and intentionality.

During the Renaissance, Leonardo da Vinci was particularly interested in human anatomy and wrote in his sketchbooks on the importance of understanding the structure of nerves, bones, muscles and sinews in order to properly depict human motion in art (Klette & Tee, 2008). Profoundly attentive to the gestures and facial expressions of individuals he encountered, da Vinci developed ideas on the relationship between the mind, nervous system and motor responses (e.g., which nerves are responsible for which movements). He was convinced that one's gestural behaviour was indicative of the pureness of their soul and incorporated this idea into his paintings. This was seen for example, in the *The Last Supper*, of which it has been said: "the ebb and flow of movement along the table is the outward effect of the inner causes of motion and emotion, as the individual movements of each disciple speak to body language of their mind" (Kemp & Wallace, 2011, paragraph 4). The tremendous contribution of da Vinci's work represented not only a revolution in art but also a great advancement in the knowledge of human motion.

In the centuries after da Vinci, developments in human movement modelling were greatly influenced by advancements in physics and knowledge of the natural world. During the Baroque era, Giovanni Alfonso Borelli (1608-1679) was the first to investigate biological animal locomotion by using methods devised by Galileo in the field of mechanics—earning him the title of the 'father of biomechanics' (Klette & Tee, 2008). Meanwhile, René Descartes (1596-1650) was developing calculus and geometrical algebra, which lie at the core of human motion analyses. By the end of the Enlightenment, when Sir Isaac Newton (1642-1727) had published his three laws of motion (in 1687), most of the basic concepts commonly used today to study biomechanics had been developed.

Despite these major milestones, the modelling of human motion remained theoretical and representations of motion remained static. It was only in the 19th century with the development of photography, that methods of representing motion dynamically were devised. The accomplishments of the earliest pioneers of motion capture, Eadweard Muybridge (1830-1904) and Etienne-Jules Marey (1830-1904), developed systems that acted as predecessors to the technology currently being used for researching human motion and also influenced developments of early animation and cinema.

While working in California, the English photographer Muybridge was commissioned by the railroad magnate (and founder of Stanford University) Leland Stanford, to help settle a bet for $25,000. The wager concerned the locomotion of a horse's gallop, and whether there was a segment of a horse's trotting in which the animal was entirely suspended from the ground. After intermittently working on the problem for six years, Muybridge settled the wager by developing a system of 12 (later 24) cameras with shutters allowing exposure times of 0.002 seconds. The cameras were each triggered by the horse's feet (Kitagawa & Windsor, 2008), which galloped at quickly as 27 meters/second. The series of pictures provided evidence that there were indeed segments of gallop in which the horse was entirely off the ground. To add further compelling evidence, Muybridge developed the zoopraxiscope, a device that could project images in rapid successions. Muybridge's work soon extended to human locomotion and his output was published in several volumes including *Animals in Motion* (1899) and The *Human Figures in Motion* (1901). These books consist of wonderful stopped-action serial photographs of subjects participating in several activities.

Meanwhile, the French physiologist Marey had an interest in movement that ranged from human and animal locomotion to blood circulation. Marey contributed to developing chronophotography, a process in which successive photographs could be taken and stored on the same fixed plate. He invented a gun-shaped camera capable of storing the pictures with paper-based film (Kitagawa & Windsor, 2008). When photographing, some of his subjects wore dark suits with white strips very similar to modern motion capture suits, creating images of stick figures representing the motions of an actor. In his book *Le Mouvement* (1894), he examined human movement by having subjects wear special shoes, which tracked duration and phases of contact with the ground (Klette & Tee, 2008).

The research area of biomechanics in its modern form began near the end of the 19th century with contributions from Christian Wilhelm Braune (German, 1831-1892), whose work with human gait led to the development of prostheses. By the middle of the 20th century, biomechanics had become an established field with many graduate programs springing up in the United States and Europe. An early application of the research was for sports science. The motion tracking technology used during this period consisted of electronic devices capable of tracking local movement. For instance, devices such as goniometers (e.g., Finley & Karpovich, 1964) and accelerometers (e.g., Morris, 1973) were used in gait analysis studies to measure relative joint angles and acceleration, respectively. This period also saw the development of the modern electromyograph (EMG), which is used to study muscle activity. One advantage that emerging optical imaging systems had over these technologies (except for EMG) is that they allowed for the simultaneous capture of human dynamic activity over time (Winter, 2009, p. 53).

Although cinematography had been used previously for data acquisition (Eberhart & Inman, 1951), television cameras, which recorded video at a higher frame rate, proved to be a more efficient medium to analyse movement. The earliest systems were developed in-house at university gait laboratories throughout the 1960s (Winter et al., 1968). The video cameras were mounted with strobe lights projected onto markers attached to subjects, whose coordinates on the screen could be digitized and used for analysis.

Infrared cameras provided a vast improvement as they did not use visible light and recorded only the motion of reflective material. Using integrated software, calibrated infrared camera systems could track the three-dimensional displacement of markers attached to subjects. Proliferation of infrared motion capture was further motivated by commercial systems being produced by companies such as VICON (VIdeo CONvertor). Optoelectric systems, which use active as opposed to passive markers have also been used for work in biomechanics as well as music-related research (e.g., Wanderley, 2002; Livingstone, Thompson & Russo, 2009). In addition, some motion capture manufacturers are now centred around markerless motion capture, in which the subjects do not need to wear any equipment in order to be tracked (e.g., www.organicmotion.com). In these systems, three-dimensional representations of subjects are produced by software that uses dynamic algorithms to group pixels of the same kind. In this sense, the technology functions like a human looking at a complex scene in which moving parts hold our attention more than static ones.

All of the motion capturing techniques mentioned here currently target a variety of communities in both fields of research (i.e., rehabilitation, gait analysis, sports performance), and industry (i.e., underwater surveillance, computer animated feature films, video games). It is quite interesting to see that, just as it was 2000 years ago, the technology being used by creative minds working in artistic fields is the same as that being used by those dedicated to understanding how the human body works.

1.5.2 Data acquisition methods

The studies in this thesis made use of an eight-camera Qualisys ProReflex optical motion capture system housed at the University of Jyväskylä's Music Motorics laboratory. In this system (as in all optical infrared systems), each camera within the network captures two-dimensional coordinates of reflective markers (affixed to actors) by flashing quick pulses of infrared light within the capture volume. When a marker is tracked by at least two cameras, the system is able to triangulate the marker's spatial displacement in three-dimensions. The three-dimensional coordinates are constructed using a method called direct linear transformation (see Robertson et al., 2004, p. 37-38).

Temporal accuracy of the recording is determined by the rate at which the camera produces infrared flashes. Some systems can achieve rates up to several thousand frames per second. The maximum frame rate of the Jyväskylä cameras (ProReflex MCU120) was 120 frames per second. The rate used in the studies within this thesis was either 60 or 120 frames per second. Although this frame rate would have been unsuitable for capturing the fine motor skills of a drummer, for example, it accurately captured the general gross body movements performed by the participants of these studies (pianists, dancers, etc.).

Motion capture systems produce time-series data representing the position of markers with respect to the capture volume's point of origin. The volume's point of origin within the Cartesian coordinate system is determined when calibrating the cameras. Calibration is the process that takes place at the start of a motion capture session, in which the system software gains information about the orientation and position of each camera in order to create a three-dimensional model of recorded data. Markers placed on subjects are then tracked in three dimensions (x, y, z) within the global coordinate system. These three pieces of information describing the marker's current position are also referred to as the marker's independent parameters or its three degrees of freedom. It is sometimes useful, in addition to position, to know the orientation of the subject being tracked. In this case, three markers are attached to a rigid body. Rigid bodies have their own local coordinate system, and contain six degrees of freedom—"the (x, y, z) position of its centre of mass and the three angles that describe its orientation" (Robertson et al., 2004, p. 11).

In each study in this thesis, between 15 and 28 markers were attached to the participants. The markers were placed on a combination of joints around the body and at areas of interest relevant to the research question at hand. Before any movement analysis could take place, the motion capture data had to be annotated by identifying each movement trajectory under a label name. Much of this process was automated within the Qualisys Track Manager (QTM) software through the creation of a model containing the angle and distance ranges between each marker. Markers that were at times occluded by other markers resulted in labels with empty frames. The values of the missing frames were interpolated using a gap-filling module in QTM. However, manual annotation was usually required if the gaps in the data were too large. Once the annotation process was completed, the data were exported as text files with tab-separated values (*.tsv). In addition to the movement data, the text files contained headers listing the names of the markers, frame rate and other pertinent information regarding the measurement. When needed, an audio recording that had been either performed or played back to the participants during a capture had to be synchronized to the motion capture data. The synchronization between audio and movement was implemented offline after the motion capture sessions. This process is detailed in Study VI.

1.5.3 Analysis techniques

Throughout my doctoral studies, I made great use of the Motion Capture (MoCap) Toolbox, developed by Toiviainen and Burger (2011). It is a collection of MATLAB functions developed especially for the analysis of motion capture data. It is best suited for analysis with data output from Qualisys' .tsv format, but can function with other types of files (e.g., the industry standard .cp3).

Müller argues that a "central task in motion analysis is the design of suitable similarity measures to compare between two given motion sequences in a semantically meaningful way" (Müller, 2007, p. 189). He suggests that two movements can be considered similar if "they only differ by certain *global transformations.* For example, one may leave the absolute position in time and space out of consideration by using a similarity measure that is invariant under temporal and spatial translations" (p. 189). In other words, relative position is more important: the height of the subjects and their actual position within the capture space, or even the timing of their movements is far less relevant to for most research projects. Throughout each study therefore, analysis techniques were predominantly tied to the computational extraction of descriptors or *features* related to the overall kinematic and kinetic properties of the recorded movements.

These kinematic features were mainly derivatives of the position data, namely velocity, acceleration and jerk. Jerk is the rate of change of acceleration and the third derivative of position. Its sensation can be felt when, for example, pumping the breaks of a car. These features were, for instance, subjected to statistical testing using linear regression (Study II), ANOVA (Study VI) and correlational analysis (Study V). Calculating derivatives can dramatically increase the signal-to-noise ratio. Therefore, any derivation necessitates substantial filtering of the signal. This was implemented using the Savitzky-Golay smoothing FIR filter (Savitzky & Golay, 1964), which is used in the MoCap toolbox. The kinetic features were mechanical energy (kinetic and potential energy). The computation of these features was made possible by using Dempster's body segment model to obtain the inertial constants of the body (Dempster, Gabel, &

Felts, 1959; see also Robertson et al. 2004). The kinetic variables thus derived were used as input to principal component analysis in Study III. Another analysis technique was to *time warp* the data in order to control for timing discrepancies in performance studies (Study VI). This was implemented by altering an existing time warping algorithm by Verron (2005).

2 AIMS OF THE THESIS

The studies on which this thesis is based present several methodologies in which motion capture technology has been used to study music cognition. Using the embodied approach as a starting point, the studies herein centre on the movements, gestures and actions associated with producing and listening to music. The studies are independent from one another as they feature unique data sets, analysis techniques and hypotheses. Despite having varying subjects and goals, one shared commonality is their uses of motion capture technology as the main data-collecting utility. Much of the research within this thesis would not have been possible without the ready availability of such technology.

Between the six articles, four music-related activities are discussed: piano performance, conducting an ensemble, choreographed choral performance and moving freely while listening to music. The movements associated with these activities span different functions, from music-producing to music-perceiving. I have organized the studies according to the categories of music-related movements as outlined by Jensenius (2007) and Jensenius et al. (2010). Jensenius classifies movements performed by musicians into four categories based on their function within a musical performance (sound-producing, ancillary, communicative, sound-accompanying). I have extended these designations to include music-related activities in which the production of music is not the end-goal. Such activities include moving spontaneously while listening to music and dancing to music.

The categories of music-related movements could be seen as seen as existing on an axis whereby one extreme is occupied by movements whose function it is to interact with instruments (producing music) and the other extreme by movements whose function it is to accompany the music (perceiving music). The categories within the framework are not mutually exclusive—producing music inherently involves perceiving music as well. In practice, many movements do not belong to a single category but can be considered to be multifunctional. However, the immediate goal is simply to differentiate one musical activity from another. This can be done by considering the function of movements in music performance as being mainly sound-producing whereas the function of movements in other musical activities lean more towards being sound-accompanying.

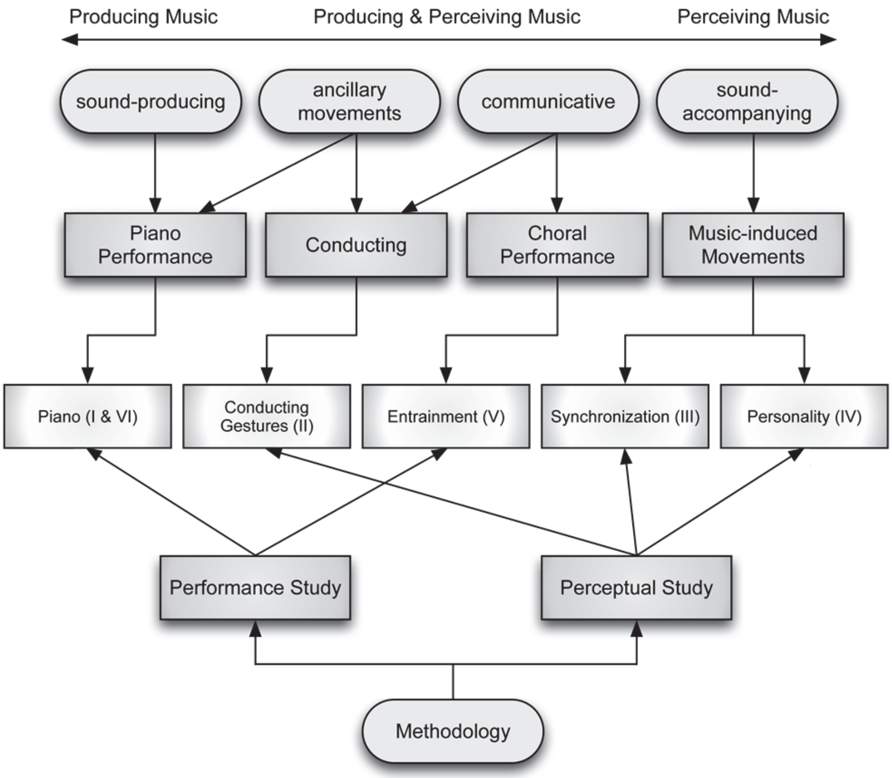


FIGURE 1: Organization of the thesis

Figure 1 presents the organization of the thesis. The top row of the schematic is an axis on which sound-producing movements occupy one extreme and sound-accompanying movements the other. Again, a single music-related movement can be multi-functional and thus labelled as belonging to either, some or all of the movement categories. The motivation behind this organization is to associate in a general sense a type of movement with a musical activity. The associations are marked by the arrows pointing down towards one ore more of the music-related activities investigated in the studies. In the third row down, the studies are referred to specifically by their subject. This level also organizes the studies according to their placement on an axis that ranges from music-producing to music-perceiving.

On the leftmost performance side of the axis appears the piano performance study, followed by conducting gestures—both of which I personally consider as being music-producing activities. While the categorization of piano performance is obvious, the inclusion of conducting gestures as a music-producing activity (performed using ancillary movements) should be clarified. Conductors use their movements to control the sound-producing movements of others. In this sense, conductors partake in the production of music, but at a level that is further abstracted from the actual sound production. At the same time this fits within a definition of ancillary movements (Wanderley, 2002), in that conductor's movements are simultaneously conveying their expressive intentionality.

In the centre, there is the entrainment study, which was investigated using a choir comprised of expert and novice performers. The South African folk song performance was accompanied by a choreography, which had to be taught to the novices of African music. In this respect, both sound-producing and sound-accompanying movements were used so that the co-performers could entrain the musical structure amongst themselves.

The studies on the rightmost side of this row, that involve music-induced movements, target two aspects of spontaneous movement. Study III deals with synchronization to music, which is a lower-level musical feature related to the music's temporal structure. Study IV deals with an individual's choice of movement. Specifically it investigated whether musical genre and personality could be reflected in the types of movements exhibited while listening to music, This could be seen as focusing more on the higher-level features of music related to form and genre.

The bottom row of the schematic indicates the methodology employed in each study. Studies I, VI and V focused on topics concerning musical performance, while II, V, and IV were perceptual studies in which participants were exposed to either visual (II) or auditory (II & IV) stimuli. Each study employed some form of statistical and/or computational analysis. Common methods included the extraction of kinematic and kinetic variables from the motion capture data and corresponding audio data. The themes of each study are further outlined below.

Studies I and VI may be seen as the bookends of the thesis. Both focus on the movements and actions performed while playing music. Study I was conducted at the very start of my doctoral studies. Three pianists participated in the study, in which their performances were motion captured and recorded. Its aim was to quantify the differences in movement when playing with three different levels of expressive intentionality. Study VI, which was conducted near the end of my doctoral studies, followed up on this research by including more participants and executing a more thorough analysis with refined methods for time-aligning performances of different durations.

Study II examined the role that movement kinematics play in perceiving emotional expressivity in conducting gestures. Participants provided continuous ratings of perceived emotion while viewing point-light animations of a conductor. The animations were rendered from motion capture data and were originally recorded while a conductor was leading an instrumental ensemble and choir. Regression models were obtained by entering kinematic movement features as predictors of the continuous ratings.

Studies III and IV investigated music-induced movements performed spontaneously while listening to music. The aim of Study III was to investigate how pulsations on different metrical levels manifest in music-induced movements. For this Study, participants were motion-captured while listening to a piece of music with a clear 4/4 beat. Signal processing methods and principal components analysis were applied to extract movement primitives synchronized with different metrical levels.

For Study IV, sixty participants completed the Big Five personality inventory (Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism), and had their movements recorded while listening to music. From the motion capture data, 55 postural, kinematic and kinetic movement features were computed. Multivariate analyses were then used to find the effects that personality traits and musical genre might have on the movements performed by the participants. This study is part of a larger project conducted by the Centre of Excellence that is entitled Music, Movement & Personality (MMP).

The aims of Study V were threefold: to investigate possible cross-cultural differences in movement, especially corporeal representations of beat and meter; to study group entrainment; and to examine factors contributing to synchronization accuracy. A visiting South African choir and a group of Finnish choir members came together for a two-day workshop. The study focuses on two performances in which the South African group taught choreography to the Finnish group. Synchronization between members and between groups was evaluated using cross-correlation and phase analysis.

3 STUDIES

3.1 Study I: Expressive movements in piano performance (a)

3.1.1 Study I: Background and Aims

Three pianists performed Brahms' *Intermezzo in A major, Opus 118 # 2* under three conditions denoted as minimum, normal and maximum expression. We were interested in analysing the contribution of the arms, torso, head etc., to the total amount of movement used for different levels of expression. The pianists attended multiple sessions and were expected to be more familiar with the piece each time they attended.

Based on previous research by Davidson (1994), it was expected that the shoulder and head would contribute most to altering the amount of expression. In effect, this meant that they were expected to show most variation in the amount of movement, while the fingers and hands would contribute the least. By utilizing motion capture, we hoped to investigate this problem from a quantitative perspective, with a higher degree of precision than had been done previously with video analysis and MIDI data (Davidson, 1994; Clarke & Davidson, 1998).

3.1.2 Study I: Methods

Fifteen reflective markers were affixed to the pianists' upper bodies (head (4), shoulders (2), centre of back (1), lower back (2), elbows (2), wrists (2), middle finger (2), and their movements were recorded at 120 Hz using an eight-camera Qualisys ProReflex motion capture system. To evaluate how much movement occurred in each body location for each performance condition, we calculated the Euclidian distance travelled by each marker. In addition, to control for disparate durations of the performances, the *adjusted* Euclidian distance was also calculated. The Euclidian distance was defined as the length of each vector using the components of marker location (x, y, z). These values, representing the norms of each vector were averaged across body location (head, shoulders, elbows wrists, and fingers). A distance matrix was then created by differencing the data so that each cell represented the distance in location between subsequent cells. Finally, the columns of the distance matrix were summed to calculate the total distance of each body location's trajectory. The adjusted Euclidian distance was also calculated in the same fashion, but with the added intermediate step of randomizing the order of the cells within the distance matrix, and truncating the matrix after the first 3500 cells (the duration of the shorted performance).

3.1.3 Study I: Results

Performances with more expression were found to contain more body movement for two of the three pianists (see Study I; figure 1; players 2 & 3). Player 1's total body movement was less affected by the level of expression. For example, in the first session, player 1's minimum performances contained more movement than the normal condition, and as much movement as the maximum condition. The other players, meanwhile, showed a clear gradient in movement for all the sessions they attended. As predicted, larger changes in amount of movement were confined to the head, shoulders and elbows while the lower back, wrists and fingers showed less variation.

Because the pianists attended multiple sessions, we expected to see a change in body movements as they acquired more familiarity with the *Intermezzo*. Using the adjusted Euclidean distance data, the effects of each session, marker location and expression on amount of body movements were examined by performing a three-way mixed ANOVA, using total distance travelled as the dependent variable. All main effects and interactions were statistically significant, and are summarized in Study I; Table 2. It can be seen from the overall effect sizes that the largest real-world effects were for expression and location, with session and all interactions having very small effects (although session seemed more important for player 1). The fact that all effects were statistically significant was most likely due to the huge number of data points analysed. Post-hoc tests (with Bonferroni correction) showed that all pairwise comparisons were statistically significant at $p \leq .001$, except for player 3 between elbows and fingers (ns) in sessions 3 and 4 (ns), and between wrists and elbows ($p < .05$) in sessions 1 and 2 ($p < .05$).

3.1.4 Study I: Discussion

This study was a preliminary attempt to computationally examine and statistically test differences in the amount of body movement when pianists play with varying amounts of expression. Our finding that movement changes were found in the upper torso and head support previous findings by Davidson (1994) and Clarke & Davidson (1998), that the head plays a major role in providing expressive movements. However, the changes in body movement across performance conditions were more prominent in only two of the three pianists analysed. This is indicative of the individual factors that no doubt play a role in music-related behaviour. There are many reasons why player 1 may not have moved in a similar fashion to players 2 & 3. These reasons could range from personality or mood, to instructions from her piano teachers. Another reason may have been lack of familiarity with the material. This would have resulted in more cognitive effort being dedicated to synchronizing the hands and memorizing or sight-reading the notes, and would have perhaps left less to focus on expression. In contrast, players 2 and 3 had memorized the score by the time they got to their last sessions and each played the piece with noticeable confidence. However, because we only investigated movement displacement, changes in articulation and the control of body movements across sessions were not detected.

This paper was presented at a conference but some shortcomings prevented it from being accepted to a peer-reviewed journal. First, the reviewers did not approve of the method used to prepare the data for statistical analysis. The adjusted Euclidian distance had been truncated so that the row of data from each performance would have the same length. But this also meant that our method eliminated data from the longer performances. Bearing this in mind, we circumvented this shortcoming in Study VI by time-warping all the data according to the onset of each measure. This method allowed proper time-aligned investigations across performances.

Secondly, the reviewers felt that analysing the movements of only three pianists was insufficient to make sufficiently meaningful generalizations about music-related movements. While the results of this study are meant to shed light on the performance movements of three individuals, this criticism is of course true—a greater number of participants will always allow for more perspective surrounding the research question. For this reason in Study VI, more pianists were invited to participate.

3.2 Study II: Conducting gestures

3.2.1 Study II: Background and Aims

The aim of Study II was to investigate the relationships between the kinematic features of a conductor's movements and the ratings of perceived expression. Twenty-four individuals provided continuous ratings of perceived expression, activity, valence and power in response to point-light animations of conducting gestures. This expanded upon previous work that had focused on the body's role in communicating expressivity in music performance. For instance, the presence of body movement has been found to enhance not only one's ability to perceive intended emotions (Dahl & Friberg, 2007), but also the intended level of expressivity (Davidson, 1993), and the musical structure and phrasing (Vines et al., 2006; Krumhansl & Schenck, 1997).

The employment of motion capture in this study served two uses. First, because the animations had been rendered from the original motion capture data, the conductors were simply seen as moving dots, which removed the influence of facial gestures, gender, etc. on ratings. Second, motion capture facilitated the computation of kinematic variables directly associated with the animations viewed by participants. These kinematic variables were then entered into linear regression models as predictors of the ratings of perceived expression.

3.2.2 Study II: Methods

Participants: Twenty-four individuals (mean age = 24.6 years, SD = 2.7, females = 16) provided continuous ratings of perceived activity, valence, power and general expressivity.

Stimuli: The stimuli were point-light animations of two conductors' gestures. The conductors had been motion-captured while rehearsing an instrumental ensemble and choir performing Mozart's Requiem. Reflective markers had been placed on the conductors' hand, wrists, elbow and shoulders. A total of six animations (three per conductor) displaying the widest range of varying gestures were produced for the study.

Procedure: The six stimuli were presented in four identical blocks, one each for ratings of perceived activity, valence, strength, and overall expression. Within-block presentation order was randomized each time, as was the order of the blocks. The animations were played within the Max/MSP environment, while participants manipulated the positions of a virtual slider via a computer mouse.

Movement features: The movement features extracted from the original motion capture data were the horizontal and vertical position of each hand (four features), the distance vector between each hand (one feature), and the instantaneous velocity, acceleration and jerk of each hand marker (six features), which were calculated using a Savitzky-Golay FIR smoothing filter (see Luck & Toiviainen, 2006).

3.2.3 Study II: Results

The rating data were collapsed across all participants and then cross-correlated against the movement features. The data were then subsequently time shifted so that the peak correlation occurred at the point where lag = zero. To model the experimental data, ordinary least squares linear regression was employed, in which the movement features were entered as predictors of the averaged and lagged ratings. Twelve separate linear regression analyses were carried out, one for each of the four rating dimensions on three datasets: (a) the whole dataset, (b) pooled data for excerpts 1–3, and (c) pooled data for excerpts 4–6. Overall, the regression models accounted for between 50% and 75% of the variance in participants' ratings of the four dimensions (mean variance: 56% for expression, 64% for valence, 65% for activity, and 61% for power). The sign and strength of the beta values revealed relationships between ratings and a number of aspects in the conductors' movements. To summarize the results: higher ratings of expression were characterized by accelerated but smooth movements, valence ratings were positively related to fast and jerky movements, higher ratings of activity by the hands being further apart, and finally power ratings were positively related to the right hand being held high and the left hand close to the body.

3.2.4 Study II: Discussion

This study demonstrated that participants' perception of expression correlated with aspects of a conductor's gestures. Observers rated the perceived expressiveness, valence, activity and power of two conductors' gestures. By applying multi-dimensional scales, we then aimed at relating the perception of expressive gestures with theories of emotion. However, it was found that some scales were redundant for this particular task (i.e., ratings for activity and power were correlated).

Differences in ratings for each conductor came down to their hand positions (e.g., conductor 1 was perceived as more expressive when the hands were spread out while conductor 2 was perceived as more expressive when the hands were closer together). While the main cues for perceived expression in general were mostly related to movement amplitude, we also found evidence that more fine-grained kinematic features also influenced observers' ratings.

Some ideas for future directions regarding this topic might include having conductors judge the expressiveness of their own movements and/or separating participants into groups (conductors vs. non-conductors). The continuous response paradigm allows observers to record their responses in real-time as their experience changes across time. However, it has been noted that this paradigm, which can result in an autocorrelated time series (i.e., the adjacent measurements are not independent of each other) is inappropriate for parametric statistical testing. Therefore, my future work concerning the perception of music-related movements will adopt functional data analysis (Ramsay & Silverman, 2005), Levitin et al, 2007)–a collection of analysis tools developed specifically for investigating time-dependent processes such as one's experience of music.

3.3 Study III: Embodied meter

3.3.1 Study III: Background and Aims

Study III investigated music-induced movement, focusing on how spontaneous movement patterns are synchronized to different beat levels of music. The ability to accurately synchronize to an external pulse has been researched extensively using finger-tapping studies (Repp, 2005). Some authors have stressed that this skill arises from sensorimotor capacities for the perception of meter and rhythm. From this view, beat induction is not a passive process but necessitates goal-directed actions involving the nervous system and musculoskeletal system (Todd, O'Boyle & Lee, 1999).

Most music contains hierarchically organized beat levels with periods having integer ratios (Palmer & Krumhansl, 1990). While the ability to synchronize to the tactus of music while dancing appears at a young age (Eerola, Luck & Toiviainen, 2006), adult professional dancers are able to synchronize to various beat levels at once (Naveda & Leman, 2010).

The study had two main aims: i) to investigate whether spontaneous music-induced movements display movement components that are simultaneously synchronized to different metrical levels and ii) to investigate whether the synchronized movement patterns differ between the metrical levels. From the motion-captured participants, we expected to see movement patterns synchronized to at least one of several possible metrical levels. We also expected inertial and biomechanical properties of the body to influence in which parts of the body the embodiments of different metrical levels occurred. These examinations were conducted by applying kinetic analysis, body modelling, dimensionality reduction and signal processing to the motion capture data.

3.3.2 Study III: Methods

A total of 18 participants, with an average of eight years musical training, were asked individually to 'move freely' to an instrumental piece of music composed especially for the study. The piece had a prominent 4/4 dance beat and was in a minor key blues form. The Qualisys ProReflex system tracked the movements of 28 reflective markers attached to the participants' bodies at a 60 Hz frame rate. The audio was played via loudspeakers. Both the audio and a synchronization pulse produced by the motion capture system were recorded into ProTools software. Synchronization between movements and audio was achieved offline within ProTools by registering the time difference between the start of the motion capture's pulse and start of the audio playback.

3.3.3 Study III: Results

Preprocessing: A set of 20 secondary markers, known as the joint markers, was derived from the original 28-marker set-up. This was done to make the movement data compatible with the body-segment model in subsequent analyses. The joint markers were either based on the locations of one the markers or the average of two or more markers. The joints were further divided into segments and established as belonging to one or more of five kinematic chains (see Study III; Figure 1b).

Mechanical energy: The kinetic and potential energy displayed by the participants was estimated using the joint locations and velocities. The inertial constants needed for this calculation were obtained using the Dempster body segment model (Dempster et al., 1959). When averaging across participants, the potential energy was found to be periodic at the one-beat level. Meanwhile, the averaged kinetic energy displayed a superposition of half-beat and two-beat periods.

Eigenmovements: Principal Components Analysis (PCA) was used to transform the joint position data into a matrix of uncorrelated variables, with each column representing a percentage of variance within the data. We coined the term eigenmovements as being the principal components projected onto the movement data. Thus, each eigenmovement represented a typical movement pattern.

Eigenmovement analysis: The structure of the eigenmovements was examined using a one-way analysis of variance (ANOVA), which identified significant differences between the three metrical levels in terms of the direction of the joint movements (mediolateral, anteroposterior, vertical). Post-hoc tests compared the individual differences in direction between each of the metrical levels. For instance, the main difference between the one-beat and two-beat eigenmodes was that one-beat eigenmodes exhibited more movement of shoulders in the vertical direction. (See Study III, Figure 6 for complete results).

Between-Subjects PCA: To investigate more closely the nature of different movements patterns at the different metrical levels, a second (between-subjects) PCA was carried out on the Eigenmovements described above. The data outputted from this secondary PCA were called the second-order PCs. To assess the degree to which each of the second-order PCs represents movement patterns synchronized to each of the metrical levels, we quantified the proportion of variance contained in the second-order PC component scores within each of the three metrical levels. Based on this analysis, we identified the following typical movements for each of the metrical levels (See Study III, Figure 7 & 8 for full results:

- *One-beat level:* (1) mediolateral arm movements; (2) vertical arm movements
- *Two-beat level:* (1) mediolateral arm movements; (2) rotation of the upper torso
- *Four-beat level:* (1) lateral swaying of the body; (2) rotation of the upper torso

3.3.4 Study III: Discussion

The analysis in this study took several steps to examine how metrical levels are embodied in music-induced movements. A kinetic analysis of peaks in mechanical energy revealed that participants embodied the musical stimuli on several metrical levels. A periodicity analysis of the movement data using autocorrelation techniques revealed the metrical levels to appear on the one, two and four beat level. Each metrical level was predominantly associated with a direction. For instance, the one-beat level tended to be embodied mostly as vertical movement while the four-beat level tended to be embodied as mediolateral movement. A more detailed kinematic analysis revealed that the tactus level often was associated with vertical hand and torso movements as well as mediolateral arm movements, the two-beat level with mediolateral arm movements and rotation of the upper torso, and the four-beat level with lateral flexion of the torso and rotation of the upper torso. This observation was in line with our hypothesis that faster metric levels are embodied in the extremities, and slower ones in the central parts of the body.

While this study was the first to tackle the question of music listeners' movement and embodiments of musical meter using quantitative methods, it did contain some flaws in experimental and methodological design. First, one limitation of the study was that is used only one musical stimulus. It may be that the observed movement patterns were characteristic of the musical excerpt but not generalizable to other musical stimuli. Other issues with this study deal with using PCA for dimensionality reduction. PCA produces components that are non-correlated. This makes the assumption that movement patterns synchronized to the different metrical levels are uncorrelated as well. Additionally, PCA assumes that the analysed data are stationary within the analysis window. Unlike walking, music-induced movements change over time, therefore the stationarity criterion may not be met. These issues listed could be overcome in the future by i) using a larger amount of musical stimuli, applying nonorthogonal dimensionality reduction methods such as Independent Component Analysis (Hyvärinen, Karhunen, & Oja, 2001) and iii) applying eigenmode extraction methods that do not assume stationary, such as Empirical Mode Decomposition (Huang et al, 1998).

3.4 Study IV: Influence of personality and genre on music-induced movements

3.4.1 Study IV: Background and Aims

Study IV is part of a larger on-going project called Music, Movement and Personality (MMP). The overarching aims of the project are to investigate the factors that influence what kind of movements individuals make when listening to music. Study IV focused on the effects of personality and musical genre on music-induced movements. Other factors taken into account throughout the MMP project have ranged from the participants' individual mood and self-expression (Saarikallio et al., 2010), to the low level audio features engrained in musical stimuli (Burger, et al., 2010). An additional direction of the project has been to use the motion capture data acquired during the trials as stimuli in a dance and attractiveness study. Study IV represents the first peer-reviewed article on the effect of personality on music-induced movements to use motion capture data. It also served to remedy some of the shortcomings on a previous proceedings paper on a similar topic (Luck, Saarikallio & Toiviainen, 2009).

Personality is related to individual differences in expressive behaviour (e.g. Gross, 1999). Perceptually, body movements and gesture patterns in speech have been shown to be reliable indicators of personality type (Ball & Breese, 2000; Koppensteiner & Grammer, 2010). One could therefore make the assumption that personality is also related to spontaneous music-related movements performed when listening to music.

3.4.2 Study IV: Methods

Participants: Participants were selected from a pool of 952 individuals who had previously completed the Big Five Inventory (BFI) online. The aim was to recruit the six highest- and six lowest-scoring individuals possible on each of the five dimensions. In the end, 64 individuals participated and 60 were retained for analysis.

Apparatus, stimuli, and procedure: Participants were presented with 30 randomly ordered musical stimuli representing jazz, latin, techno, funk, pop and rock. Twenty-eight reflective markers were attached to the participants and their movements were recorded using the Qualisys ProReflex System at 120 Hz. The movement data was synchronized to the audio stimuli using the pulse signal transmitted from the motion capture cameras. Participants were instructed to move in a way that felt natural with regards to the music.

Movement feature extraction: A total of 55 postural, kinematic, or kinetic features were extracted from the data of 60 participants. First, a set of 20 marker locations (subsequently referred to as joints) were derived from the original 28 to facilitate the calculation of kinetic features. The postural features were defined as the position of the joints with respect to the midpoint of the hip markers. The kinematic features were defined as the velocity (1st derivative), acceleration (2nd derivative), and jerk (3rd derivative) of the joint marker trajectories. The kinetic feature was instantaneous kinetic energy. The features were collapsed across time, resulting in 55 movement descriptors per participant.

Principal Component Analysis: The dimensionality of the data was reduced using Principal Component Analysis (PCA) whereby the data was transformed into uncorrelated variables differentiated by their amount of variability. The first five principal components (PC) amounted to 90.3% of the total variance and each represented uncorrelated movement attributes. PC 1 was labelled amount of Local Movement, PC 2 represented Global Movement, PC 3 was Hand Flux, PC 4 was Head Speed, and PC 5 was Hand Distance.

3.4.3 Study IV: Results

The effect of personality and genre on the five PCs was investigated by running a series of 2 (personality dimension: low-scorers vs. high-scorers) x 6 (genre) Multivariate Analyses of Variance (MANOVA), one for each of the five personality traits. To summarize the results, Extraversion was positively related to all five movement components while Neuroticism was positively related to Local Movement, but negatively related to the other movement PCs. For genre, Rock was related to Head Speed (especially for Extraverts), Techno was associated with high levels of Local Movement and Latin was related to Global Movement.

For both these genres, the results came from individuals scoring highly on the Extravert dimension.

3.4.4 Study IV: Discussion

Extraversion and Neuroticism were the only dimensions that had a significant effect (positive or negative) on all five movement PCs. The reasons for the clear effects may be that these two dimensions are strongly connected to emotional expressivity (e.g. Gross & John, 1995). Individuals scoring high in Extraversion tend to be energetic and in search of constant simulation. This may have contributed to their positive relationship to all PCs. Individuals scoring high on Neuroticism on the other hand tend to be more anxious or self-conscious; their results showed a positive relationship to local movement but negative relationships to the other PCs.

Because we were interested in the expressive gestures linked to music in the general population, we specifically avoided inviting trained dancers to participate in this study. However, the experimental setting was perhaps too synthetic to entice an average person to move in a manner in which we could link the gestural expressiveness to their personality trait. Therefore, a future paradigm should include more ecological settings. Individuals have been known to dance more when others are around (de Bruyn et al., 2009). One direction might therefore be to have multiple individuals scoring highly on the same personality trait move to music at the same time. Other directions may be to explore alternatives to the Big Five taxonomy of personality traits as well as consider musical preferences. Given the linkage between personality and musical preferences (Rentfrow & Gosling, 2003) and body motions as reliable indicators of others' personality type (Ball & Breese, 2000), the study of music-related movement and personality types could help inform us about how we use our sensorimotor capacities to perceive emotional content in music.

3.5 Study V: Learning and synchronizing dance movements

3.5.1 Study V: Background and Aims

Study V focuses on entrainment and coordination within a cross-cultural setting. In June 2008, a choir from South Africa and a group of Finnish choir singers were brought together for a two-day choral workshop in Jyväskylä. Songs with choreographed dance movements from various African cultures (e.g. Zulu, Sotho and Xhosa) were taught to Finnish participants. The unique gathering, the main goal of which was to share and experience musical cultures, provided an ecological setting in which to examine differences in corporeal representations of beat and meter between experts and novices of African music.

Singing in a choir while dancing requires vocal competence and a command over rhythmic choreography. In addition, it also requires cognitive resources dedicated to being aware of others and their actions. As in other music-related joint actions (Keller, 2008), choral singing requires an alignment of mental states through which performers share representations of tonal and metrical schemata (Tomasello & Carpenter, 2007) and choreography must be rehearsed and engrained within a motor program (Davidson & Correia, 2002). However, one must constantly monitor their output while taking into account that others are also making adjustments of their own behaviour (Repp & Keller, 2008). These joint actions may be further complicated in cross-cultural or learning situations, in which the performers have different conceptualizations of rhythm and meter.

The aim of this study was to investigate experts and novices performing traditional African and Finnish folksongs, in which both styles were accompanied by African-style choreography. Exploratory in nature, the performers were motion-captured and synchronization and entrainment issues were examined using signal-processing techniques. We were interested in evaluating the synchrony within each group and identifying metrical hierarchies embodied within the performers' movements.

3.5.2 Study V: Methods

Participants & materials: Participants were members of the Emmanuel Lutheran Choir from South Africa on a visit to Jyväskylä as well as selected Finnish participants from the University of Jyväskylä's extended community. The Finns had musical experience and musical training, mostly in Western music and choir singing. The South Africans and Finns were divided into groups along cultural lines, *Experts* and *Novices* respectively. Of all the performances throughout both days, two songs were chosen for this study. One performance was a South African song with a set choreography and the other a traditional Finnish song for which the South African participants prepared a novel choreography. In both songs, the South African participants taught the Finnish participants the choreography.

Apparatus & procedure: Movements were recorded using the Qualisys ProReflex motion capture system at 120 Hz. For the African song, eight participants (four experts and four novices) wore reflective markers (head + feet) and were positioned at the centre of the room in a two-by-four formation. In the Finnish song, there were ten performers, three of which were novices and markers were placed on the head only. The discrepancy in marker placement reflects the ecological setting in which the data was recorded: The Finnish song contained more performers within the capture volume, thus making it impossible to practically track the movements of the foot markers.

Analysis: The African song was analysed in more detail because a) more markers were used and b) our main interest was in the situation where true novices would learn a musical style they were not familiar with. As the focus was on timing, meter, and entrainment, a number of different synchronization analyses were performed. First, pair-wise cross-correlations were performed on acceleration of the heads and feet to investigate the coherence within and between groups and between all participants. The pair-wise temporal development was also analysed using windowed cross-correlations. The next part of the analysis focused on synchronicity within the entire group. The head marker position data was band-pass filtered using the music's tactus as the centre frequency. Then the phase of the vertical movement for each performer was calculated using the Hilbert Transform (Khvedelidze, 2002). A model for measuring synchronization (Acebron et al., 2005) was then applied to the phase data in which an indicator of coherence between each participant was calculated. The model's output could be plotted as a function of either the group's coherency as a whole, or an individual's deviations from the rest of the group.

3.5.3 Study V: Results

Pairwise analyses: Cross-correlations were calculated on the vertical acceleration of the foot markers in the South African song and on the horizontal acceleration of the head markers in the Finnish song (maximum lag of 500 ms). The maximum coefficient (*r*) in the cross-correlation function represents the strength of the correlation between two individuals, and the lag at which the maximum coefficient occurs signals the synchronization error between individuals. In this respect, the experts appeared more synchronized than the novices. With the windowed cross correlations, we were able to observe the temporal evolution of a pair's synchronization. Although some areas of higher synchrony were identified, the output did not identify a clear leader between novice and expert.

Phase analysis: The evolution of phase between groups was monitored using the order parameter, a unit-less phase coherence score between 0 (unordered) and 1 (perfect synchronization). For the Finnish song, the Finns appeared more synchronized as a group than the South Africans. In the South African song, synchronization was more stable among the Expert group. For the Finns, synchronization broke down at the boundaries of different sections in the choreography. This probably reflects the fact that experts do not have problems in remembering what the next pattern is, while novices might hesitate and need extra time to readjust their movements.

Metrical levels: When plotting the vertical feet and head movements of a Novice performer alongside an Expert performer, it was found that the Expert exhibited multiple levels of the music's metrical hierarchy. For the novice, the head marker's trajectory has a semblance to the feet: they appear to step in time together. For the expert, however, the head marker's trajectory captures not only the step-level periodicity, but also a slower movement occurring at a period that is four times longer than the step-level pulse. In other words, the expert's feet are locked to the beat of the music, while the head embodies the slower, bar-level rhythm of the music. This finding reinforced the findings in Toiviainen et al. (2010) and Study III, in which it was found that different parts of the body exhibit different levels of the metrical hierarchy.

3.5.4 Study V: Discussion

In general, the experts appeared more coherent and synchronized. However, they seemed less coherent when performing less familiar music. This was a predictable result, as theories about attentional resources (Keller, 1999) as well as about motor programs (Davidson & Correia, 2002) would suggest this to be the case.

In the phase analysis, the Novices appeared in the Finnish song more synchronized as a group than the Experts (even when controlling for the smaller amount of Novices taking part in the performance). This indicates that the Expert group, while more comfortable with the movements, focused their cognitive effort on the memorization of the lyrics and melody more than the Novices. The reverse is true for the African song, in which the expert group were more synchronized. The areas of transitions where the steps change appear to have been what cause the novice group to become less coupled with the experts. In both cases, it appears to have been a matter of dividing cognitive effort between different functions, the choreography for the novices and the melody and lyrics for the experts.

The data also indicated that the novices exhibited mainly one metrical level in their movements, which was locked to the most salient beat. The experts were able to embody multiple levels at once. This last observation may add support to the findings of Study III concerning metrical levels being parsed throughout different parts of the body.

Despite some methodological drawbacks, this study provided an ecological setting and novel methodology in which to examine synchronization and aspects of joint action. However, further studies on this topic should utilize a stricter experimental protocol (i.e. same number of participants and number of markers in each trial).

3.6 Study VI: Expressive movements in piano performance (b)

3.6.1 Study VI: Background and Aims

Eight pianists performed a Prelude by Chopin (E minor, op. 28, no. 4) under four conditions denoted as normal, deadpan, immobile and exaggerated. Serving as a follow-up to study I, this project implemented refined versions of our methods used to gauge changes in body movement, timing and dynamics across performance conditions. A secondary aim was to investigate if there was any significant difference between playing with limited movements (Immobile) and playing with limited expression (Deadpan). To this end, the study incorporated the performance conditions found in the work of Davidson (1993, 1994, 1995), Wanderley (2002) and Wanderley et al. (2005). The conditions used by Davidson emphasise a gradient of expression while Wanderley's conditions emphasise a gradient in the amount of movement. By synthesizing both sets, it was hypothesized that an interesting relationship might be revealed between the Deadpan and Immobile conditions. As in Study I, it was further expected that the parts of the body farthest from the keyboard would act as key purveyors of expressive intentionality, as their trajectories would be most affected by playing with more expression. Meanwhile, the parts of the body involved in sound production (the hands and fingers) would show less movement variation between performance conditions.

3.6.2 Study VI: Methods

Performers: eight pianists performed the Chopin Prelude in E minor Op. 28 # 4. The performers were sent a copy of the score and were requested to practice the piece at home prior to the recording sessions.

Apparatus and procedure: twenty-six markers were attached to the pianists' upper body and arms and their movements were captured at 120 Hz. They performed on a digital piano with weighted keys and the direct audio signal was recorded into Pro Tools 8. Synchronization between audio and motion capture data was implemented offline using the pulse signal generated from the motion capture cameras, which had also been routed into Pro Tools. The pianists performed the Prelude a total of twelve times, by cycling three times through the four conditions—normal, deadpan, exaggerated, and immobile.

Pre-processing audio data: the audio data was segmented manually according to the onset of each measure in the piece. The durations of each segment served as reference points for time-warping the data. Also, the dynamic range of each audio segment was determined by calculating its root mean square (RMS). This value was used to observe differences in dynamic range across performance conditions.

Pre-processing motion capture data: the original marker positions were transformed into a subset of markers, which are referred to as the areas of interest. The motion data for these areas were then temporally aligned by applying a nonlinear time-warping algorithm.

3.6.3 Study VI: Results

Performance duration and timing: The total durations of each of the pianists' performances were averaged together to examine how the performance conditions differed in terms of general timing. The deadpan condition generally had the shortest duration while the exaggerated condition had the longest duration (see Study VI; table I). This was also true when investigating each measure individually (Study VI, figure 2). Interestingly, the durations for the normal and immobile conditions were similar overall at the individual measure level. Deadpan performances contained the least amount of timing variations and exaggerated performances the most, which were most prominent at the climax of the piece. A one-way analysis of variance (ANOVA) indicated that overall timing differences between performance conditions were not significant, $F(3,92) = 1.77$, $p = 0.158$. However, when one-way ANOVAs were conducted for each measure individually, significant timing differences were found for measure 12, $F(3,92) = 4.68$, $p<0.01$, measure 18, $F(3,92) = 3$, $p<0.05$; measure 22, $F(3,92) = 3.38$, $p<0.05$; measure 23, $F(3,92) = 7.79$, $p<0.001$; and measure 24, $F(3,92) = 4.84$, $p<0.01$. Post-hoc tests with a Tukey-Kramer correction (Study VI, Table 3) indicated that discrepancies in duration between the deadpan and exaggerated performance types accounted for most of the statistical significance.

Evaluating amount of movement per measure across performance conditions: using the time-warped data, the total distance per measure was calculated and averaged across participants for each performance condition. The amount of movement within the deadpan and immobile conditions were not statistically different from one another, demonstrating that the pianists associated low expression with low amount of movement. Statistical differences between deadpan and exaggerated were highest in the parts of the body farthest from the keyboard.

Correlations between the velocity and acceleration of each performance type: velocity and acceleration features were computed by applying a differentiation algorithm to the time-warped movement data. The velocity and acceleration curves were correlated with each other in order to see how these features differed within performance type. The normal and exaggerated performances were the most correlated while the deadpan and exaggerated were the least correlated.

Performance dynamics: based on the RMS values computed individually for each measure and then averaged, the exaggerated performances yielded the most dynamically varied performances. Dynamics (changes between low RMS and high RMS) fluctuated more near the end of the piece.

Interview data: after the experiment, the pianists were asked to write about the strategies they had used for playing in each performance condition. With the exception of the immobile condition, the strategies tended to be guided by mental imagery rather than choices regarding physical movements. For example, in normal performances, they imagined performing for a friend or their teacher, and in the deadpan condition, they imitated a dull MIDI performance.

3.6.4 Study VI: Discussion

Deadpan performances were on average the shortest in duration while exaggerated performance were the longest in duration. Overall, these differences in duration were statistically non-significant. However, when each measure was evaluated individually, it was found that the difference in duration between deadpan and exaggerated performances was significant for five of the total 25 measure in the piece. These five measures occurred at the end of phrases—in which the harmonic structure is more tense—and corresponded with leaps occurring in the bass. This finding highlights two points. Firstly, the use of expressive timing occurred in areas of harmonic tension, signalling that the piece's musical structure plays a role in affecting when the performance's regular timing is perturbed. Secondly, the variation in timing that is actualized is in fact a physical necessity—the left hand leaps from the low end of the keyboard to the middle register.

While the greatest differences in duration were between deadpan and exaggerated, the normal and immobile performances had very similar timing profiles. This finding reflects the nature of the instructions, in which the pianists were, not surprisingly, told to move as little as possible during the immobile performance, as compared to the others. Although the participants commented on the difficulty of playing with diminished movement, the timing profile was essentially the same as if playing normally. This finding also contrasted with Wanderley et al. (2005), in which it was reported that movement attenuation led to faster performances. The disparity between studies was probably due to their piece of music being more technically challenging then the one used in our Study VI.

The analysis dealing with the total amount of movements per measure revealed that discrepancies in the amount of movement across the deadpan, normal and exaggerated conditions were found to be mostly in the parts of the body farthest from the keyboard—namely the head, neck and shoulders. The largest differences were again seen in the measures featuring harmonic tension and necessitating physical left hand leaps. In these measures, the head movements co-articulated two purposes. First, the head acted as an agent of expressive intentionality (as indicated by its increase in movement as the performance increased in expression). Second, the head served to guide the left hand while leaping from the bass register to the middle of the keyboard.

Another finding from this analysis was that the deadpan and immobile conditions employed similar amounts of movement. From the statistical analysis, the amount of movements per measure was never significantly different between immobile and deadpan performances, which indicates a clear association between playing without expression and playing without extraneous movements.

Pearson correlations between the performances' velocity curves indicated that the left elbow, wrist and finger were the most correlated across the performance types. This again showed that the parts of body whose movements have a functional role in the production of sound are less prone to change as expressive intentionality is increased.

Generally, these findings reinforce the notion that corporeal behaviour in music performance is influenced by musical structure, and the physiology of the instrument, as well as interpretive choices (Wanderley, 2002; Wanderley et al., 2005). For example, head movement trajectories increased in amplitude at harmonically tense sections of the score (musical structure), and were highest in performances where the expressive intentions were exaggerated (interpretive intentions). The head changed its movement the most because it acted specifically as a conveyor of expressive intentions—a function less suited for the parts of the body needed to produce sound (physiology of the instrument).

4 CONCLUSIONS

Embodied cognition views the human body as being central to our experience of the world. The interaction between our sensorimotor facilities and the environment—and most importantly for this thesis, musical stimuli—is brought about through goal-directed corporeal articulations, which allow us to engage in music-related activities. The theory of embodied music cognition posits that such corporeal articulations give rise to musical meaning that is communicable to others via observation. The corporeally based transference of musical meaning is a powerful attribute that allows the conveyance of expressive intentions and structural cues within musical activities. In addition, it is the mechanism that facilitates entrainment and shared sensorimotor representations among co-performers.

The studies included in this thesis each used the embodied approach to study music-related movements within four music-based activities: (1) piano performance (I & VI), (2) conducting an ensemble (II), (3) choreographed choral performance (V) and (4) music-induced movements while listening to music (III & IV). All of the studies employed a motion capture system that was used to collect positions of the subjects with high degree of temporal and spatial precision. The following sections highlight the main contributions each study made to the overall view of music-related movement. For each study, I have included some methodological considerations and indicated how future research might improve upon what has happened here. I then present some plans for future research endeavours I would like to participate in, upon completion of my doctoral studies.

4.1 Summary of findings

4.1.1 Expressive movements in piano performance

Study I looked at the performance gestures of three pianists. The study adopted the standard paradigm from musical emotion research (Juslin, 2010), in so far as emotional expressivity was examined under three performance conditions. These conditions (minimum, normal and maximum) had, in turn, been adopted from earlier work by Davidson (1993).

The analysis investigated the distance travelled by markers that were attached to several key points of the body. It was revealed that the points which were most affected by differing amounts of expression were the parts farthest from the keyboard (the head and torso). This finding highlighted that for pianists (and possibly for other types of musicians with restrained mobility, i.e., marimbas, Broughton & Stevens, 2009) the parts of the body not involved in sound production are most used to convey expressive intentions, while the parts of the body focused on sound production are more restrained to their functional role.

However, one pianist, for the maximum expression rendition, did increase the size of his hand gestures at certain locations in the score. These locations required the left hand to make a large and fast jump from the low end of the keyboard to the mid-range. This score location apparently provided a structural anchor from which the pianist chose to articulate a feature from the score. It can also be seen as an example of sonic gestures articulated as corporeal gestures.

Despite these interesting findings, only two of the three pianists significantly altered their movements when employing the three levels of expression. During interviews with the pianists, it was revealed that each participant interpreted the instructions differently, highlighting the individual nature of expression in music performance and possibly a shortcoming of the experimental design. Many of the methodological issues concerning this study were taken into account in study VI, which is discussed below.

Study VI again investigated the expressive movements in piano performance and was largely a follow-up to study I. However, several new issues were investigated and the methodology improved. Pianists performed under four conditions (normal expression, deadpan, exaggerated expression and immobile). This scheme was meant to synthesize performance manners used both by Davidson (1993), as well as by Wanderley et al. (2005). In addition to observing how movements were affected by different expressive intentions, a secondary goal was to investigate the relationship between playing without expression (deadpan) vs. playing without extraneous movements.

The results largely supported the previous findings of study I, Davidson (2007) and Wanderley and colleagues (2005). Movements in the exaggerated condition were more varied in the head and torso than in other parts of the body. Additionally, structural elements of the score such as harmonic cadences seemed to predict when the variations in movement would occur, as previously highlighted by others (Clarke & Davidson, 1998: Palmer et al., 2009).

As with the other studies, one could relate this finding to Leman's level of engagement theory. In the exaggerated expression renditions, the pianists appear to be more attuned and empathic to the emotional qualities of the music being performed. The pianists' goal-directed actions are at this point giving rise to musical meaning. Their movements draw attention to their expressive intentions, which are in turn being formulated in real-time while performing (in terms of an action/perception coupling).

Study VI also provided evidence of an association between playing without expression and playing without extra movements. The results showed no significant differences in amount of movement between immobile and deadpan performances. Interestingly when comparing the normal and immobile conditions, their performance durations were very similar. This finding highlights the fact that expressive body movement can be manipulated just as any other expressive parameter. In the immobile condition, though movements were restrained, the pianists still employed an expressive timing profile similar to that found in their normal performances.

Future work on the topic of expressive movements in piano performance will involve a perceptual study similar to the conducting gesture study (II) in which pianists' expressiveness will be investigated through the perception of their kinematic cues. This potential study may take an embodied approach whereby continuous ratings are recorded through some sort of embodied response (motion capture or game controller).

4.1.2 Conducting gestures

Study II investigated the perception of expressive and emotional content ingrained within conducting gestures. Participants rated dynamic gesture displays via a virtual slider in a bid to identify which components of conducting gestures convey emotion expressivity.

Among other findings, the results of the study showed that accelerated gestures corresponded to high expressivity ratings, while spatial positioning of the gestures (particularly time-keeping gestures) corresponded to high power ratings. By having each participant provide a continuous response to the perceived expression of visual stimuli, the study adopted a framework that is common to studies on musical emotion (e.g. Schubert, 2004). Furthermore, participants rated the movements on four separate scales (valance, activity, power and general expressivity), based on the multi-dimensional theory of emotion (Osgood et al., 1957; Russell, 1980). The study’s method thus adopted a framework that had been previously used for investigating auditory musical cues, in order to study visual musical cues. In addition, it also produced regression models using kinematic features extracted from motion capture, and these became predictors of the ratings. This last point is significant as past work on the perception of musical visual cues employing continuous responses, used video stimuli, and not the higher definition of motion capture (i.e. Krumhansl & Schenk, 1997; Vines et al., 2006).

As study II was a preliminary attempt to incorporate several methodological elements into one (continuous response, multidimensional theory of emotion, motion capture), there are several ways in which follow-up studies could be improved. First in terms of analysis, it is known that linear regression can be inappropriate for time series data as it is a procedure that assumes independent data points (time series is often autocorrelated). One potential remedy, suggested by Schubert (2004) has been to difference the time series, thus evaluating the ratings' gradient as opposed to the ratings themselves. A drawback is that by differencing the data, noisy artefacts are introduced into the time series, which further affect the regression models. Although autocorrelation does not appear to have affected the regression models presented in this paper, I have faced such issues with other time series data sets. An alternative to traditional parametric testing would be functional data analysis (Ramsay & Silverman, 2002; Levitin et al., 2007), which is a specialized branch of statistics targeted towards time series data. This method has been used in several studies on music-related movements and is an avenue I am interested in exploring in my future work on continuous response studies.

4.1.3 Choreographed choir performance

Study V investigated embodied entrainment between performers in a South African choral performance, which simultaneously featured an accompanying choreography. An additional aspect was that the members of the choir were from different cultures. The expert group (South Africans) had to teach a step sequence to the novice group (Finns). This study investigated possible cross-cultural differences in movement, especially corporeal representation of beat and meter by studying group entrainment and factors contributing to synchronization accuracy.

Results showed differences in embodiment of rhythm and synchronization between the novice and expert groups. The expert group was more unified, which was undeniably based on a familiarity factor. However, an analysis of the temporal dynamics within each group revealed a continuous, mutual adaptation to achieve accurate entrainment (Keller, 2008). Interestingly, looking at the vertical movement of the head marker, the novice group appeared more cohesive than the expert group

In terms of individual synchronization to the beat of the music, novices appeared to embody only the most salient metrical level while the experts exhibited several metrical levels. Besides the fact that the novices might have been preoccupied with recalling lyrics and choreography, there is evidence that the groups had divergent sensorimotor representations with respect to the beat. The experts explained in an interview that the source of movement occurred at the core of the body, rather than the feet. In their view, this allowed a decoupling between their head and feet movements, thus permitting them to represent different metrical levels in different parts of the body.

Because the groups showed different corporeal behavior, this could be an indication that they were engaging with the music at different levels of imitation. The novices seemed more concerned with synchronizing with the musical beat, a low-level feature while the expert group was engaging with the music at a higher-level, which would mean they were attuned to the lyrics, affect and expression of the music (Leman, 2008).

Although the study presented some interesting preliminary results, future work on group entrainment would benefit from several methodological improvements. The main concern for this study was a lack experimental control. Although the workshop provided a favourable ecological environment for a joint action study, the handling of participants as well as conducting a performance with different marker placements and a different number of participants each time, unnecessarily complicated the analysis.

4.1.4 Music-induced movements

Study III was a novel attempt at investigating the relationships between music's metrical hierarchies and emerging music-induced movements. Participants were asked to move freely while listening to a musical excerpt, but the analysis centred only on movements that were synchronized to a pulse in the audio signal. As it turned out, most participants appeared to have a natural tendency to move with the beat. This alone demonstrates the innate tendency to imitate the sonic pulsations and vindicates beat induction theories (Todd et al., 1999).

Movements tended to be synchronized on either the one-, two- or four-beat level. This finding supported the hypothesis (and supported previous research by Todd et al., 2007) that biomechanical properties (inertial properties of the body segments) influence which parts of the body are more likely to embody each metrical level. For instance, faster beat levels (i.e. the tactus) were embodied in arm movements while full torso movements embodied the slowest beat level (one-beat level).

Synchronization represents a form of imitation and engagement with a low-level feature of music. It is a low level feature in the sense that it is the musical features we are most prone to imitate corporeally (Knuf, Aschersleben & Prinz, 2001). The study was a good example of applying the embodied approach to music research as it revealed how the perception of multiple beat levels is parsed through individual, and sometimes co-occurring goal-directed actions. These observed participatory actions in response to environmental stimuli shed light on the behavioural outcome of our online tracking mechanisms (Wilson, 2002; Anderson, 2003).

While beat induction is usually investigated using a finger-tapping framework, this study provided a more ecological framework by investigating freely performed movements in response to a musical stimulus. Methodological weaknesses in this study included i) the use of only one musical stimulus; and ii) the use of principal components analysis to reduce the dimensionality of the movement data into uncorrelated variables (when it could be the case that movements occurring in different parts of the body are correlated). The first issue is easy to solve, as future studies on this topic could use a variety of music (as was done for Study IV). The second issue could be overcome by implementing different analytical methods, such as Independent Component Analysis and Empirical Mode Decomposition.

Study IV investigated how individual personality factors as well as musical genre effect music-induced movements. The approach was similar to Study III, however, for this study, the aim was to find relationships between individual factors and movement patterns. Personality traits were determined by having participants fill out the Big Five personality questionnaire. A total of 30 musical excerpts of various genres were then played to the participants.

An extended set of kinematic and kinetic features were extracted from the movement data of 60 participants. Principal Component Analysis (PCA) was used to decrease the dimensionality of the data set and produce uncorrelated variables. This allowed features with similar amounts of variability to be grouped together. Multivariate analyses then showed that individuals rating highly on Extroversion and Neuroticism (both emotionally charged personality traits) showed the clearest relationships between the movement variables. For example, extroversion was related to head speed and global movement. Relationships between musical genre and movement patterns were also found. Rock was associated to head speed (picture *head-banging*) while Latin elicited high levels of global movement.

This study was a first attempt at linking personality traits with freely performed movement patterns and is significant for several reasons. First, unlike other studies on music and dance, it drew participants from the general population and not professional dancers (Camurri et al., 2003; Casciato et al., 2005). This is a crucial point to establish for embodied music cognition, as it is supposed to apply to musical phenomena as they occur naturally. The ability to move spontaneously to music either at a low level of engagement (synchronization) or at a high level of engagement (empathic) is common among all humans. Researchers may benefit from taking into account the various and individual roles music plays (e.g. mood regulation) within the general population. A second implication for this line of research is that it may provide a reference point for music retrieval systems that are designed to take gestures as input. One challenge for such systems would be that gestures articulated by individuals vary from person to person. If a theoretical database were able to associate the input of a typical gesture pattern with a musical genre, a personality trait, or even the evocation of a mood, this would then be a starting point for navigating though databases using corporeal articulations. Of course, for this research to be relevant for this application, a future experiment on this subject may restrict behaviour to specific localized movements, such as arm gestures. Future work on the individual factors determining movement patterns should therefore take into account more forms of music (and possibly sub-genres) and explore alternatives to the Big Five taxonomy of personality traits by considering other forms of individual factors such as musical preference, education, age, etc.

4.2 Future directions

My four years of doctoral studies have been enlightening and challenging. I have had the opportunity to conduct multidisciplinary research, learn a wide range of technical skills and develop a strong sense of critical thinking. I look forward to applying this skill set and knowledge to new and exciting research projects.

At the current time, I plan to build upon my doctoral studies through a post-doctoral position at the Centre of Excellence in Interdisciplinary Music Research at the University of Jyväskylä. My future research interests consist of both short-term projects as well longer-term participation within research projects, which will culminate in a post-doctoral position.

One endeavour I am interested in pursuing immediately after finishing my doctoral studies is to publish a literature review concerning the application of motion capture technology in music-related research. During completion of this thesis, I have located many studies within the literature on music perception and musical gestures that have made use of motion capture throughout the, and so came to the conclusion that this research community would benefit from a state of the art literature review. To the best of my knowledge, no such work has emerged in any of the prominent music journals (although there exist some reviews of music and movement, e.g. Schutz, 2008). The review currently being planned would act as a point of entry for any researcher interested in using motion capture for their research (thus it would be targeted to a general audience) while also acting as a general point of reference for those with experience in this area.

In the longer-term, there is at least one specific project in which I plan to participate. The project deals with active music therapy and post-stroke recovery. Active music therapy is a form of treatment in which clients engage in musical improvisations with their therapists using instruments or song. It is used to treat a variety of conditions ranging from depression (Erkkilä et al., 2008) to brain injuries (Formisano et al., 2001). The Music Department of the University of Jyväskylä, together with several collaborators, is currently running a major project involving active music therapy and post-stroke recovery. The aim of this project is to develop a neuropsychologically-informed clinical music therapy model which caters for middle cerebral artery (MCA) stroke patients and to examine the effects of active music therapy on post-stroke recovery. The active music therapy sessions consist of rhythmic motor activities where the client plays a series of patterns on a MIDI drum set while being directed by the therapist, who is playing the piano, and also through improvisation using djembe drums. Motion capture is used to observe longitudinal improvements in motor skills and dexterity over the course of several sessions throughout an extended period of time. My role in this project would be to participate in the data gathering and analysis of the motion capture data obtained from each session. This will consist of developing measures for sensing changes in movement fluidity and in the range of motion for each client from session to session.

4.3 Closing remark

In the introduction, I highlighted that one of the reasons that interest in studying the body's role in music-related activities has escalated in the past few decades has been because of new technologies allowing for more accurate tracking and the instantaneous analysis of body movements. This technology ranges from elaborate motion capture systems, to much simpler motion-based gaming controllers. In addition, we also possess technology such as brain scanners, which we can use to track musical engagement at the cerebral level. Indeed, we now possess a wide breadth of technology from which musical engagement and corporeal imitation can be understood on multiple levels. The issue now is to develop the methodologies and experimental paradigms with which to study music-related movements—to use Godøy's term, we must conceive of a conceptual apparatus. Thus, the aim of this thesis has been to present novel research methods for the young yet burgeoning field of embodied music cognition. While it is unclear whether these methods will thrive, the preliminary studies included within this thesis represent an excellent starting point for future work.

YHTEENVETO

Liikkeenkaappausteknologian soveltaminen kehollisen musiikkikognition tutkimuksessa

Ilmaisulliset kehonliikkeet ovat erottamaton osa musiikkia. Muusikot käyttävät kehonliikkeitä emotionaalisessa ilmaisussa, kapellimestarit käyttävät eleitä välittääkseen musiikin rakenteita, ja yhtyesoittajat luottavat kehollisen ilmaisuun soittaakseen yhtenäisesti. Lisäksi musiikin kuuntelu voi aiheuttaa kuulijassa spontaaneja, musiikin tempoon luontaisesti tahdistuvia liikkeitä. Vaikka musiikkiin liittyvät kehonliikkeet ovat yleinen ilmiö, liikkeet itsessään vaikuttavat olevan hyvin yksilöllisiä ja saattavat liittyä musiikissa oleviin ilmaisullisiin piirteisiin. Musiikkiin liittyvien ilmaisullisten kehonliikkeiden merkityksellisyyden tiedostaminen on myötävaikuttanut kehollisen kognition soveltamiseen musiikin kontekstissa. Kehollisen kognition näkemyksen mukaan tietoa ympäröivästä maailmasta saadaan sensomotoristen representaatioiden kautta, jolloin keho toimii eräänlaisena välittäjänä ajatusten ja ympäristön välillä. Musiikillinen toiminta kuten tanssiminen tai esiintyminen ovat erinomaisia esimerkkejä kehollisesta kognitiosta, sillä ne vaativat jatkuvaa aisti-informaation ja oman toiminnan yhteensovittamista. Kehollisen musiikkikognition näkemyksen mukaan musiikillista vuorovaikutusta tapahtuu useilla eri tasoilla; voimme esimerkiksi synkronoida keholiikkeemme musiikin tempon tahtiin, tai eläytyä musiikin ilmaisemiin tunteisiin. Molemmissa tapauksissa musiikillisia merkityksiä syntyy tavoitteellisten, kehollisten toimintojen kautta, jotka sekä reagoivat että vaikuttavat musiikillisiin kokemuksiin.

Tämänkaltaisten ilmiöiden tutkimus on hyötynyt suuresti optisten liikkeenkaappausteknologioiden kehityksestä, joiden avulla kehonliikkeitä voidaan tallentaa hyvin tarkasti. Kehonliikkeiden analyysissa merkitykselliset liikemuuttujat erotetaan liikkeenkaappausdatasta. Näillä muuttujilla voidaan kuvata muusikoiden ja tanssijoiden liikkeiden kinemaattisia, kineettisiä sekä behavioraalisia piirteitä, ja niitä voidaan analysoida käyttämällä useita eri menetelmiä.

Tämän väitöstutkimuksen tavoitteena on havainnollistaa erilaisia liikkeenkaappausteknologian sovelluksia kehollisen musiikkikognition kokeellisessa tutkimuksessa. Kuudessa erillisessä tutkimuksessa tarkasteltiin neljää erilaista musiikillista toimintaa: pianistin esiintymistä, kapellimestarin eleitä, kuoron esiintymistä (koreografia kera) sekä musiikin aikaansaamaa liikettä.

Avainsanat: kehollinen musiikkikognitio, liikkeenkaappaus, eleet, musiikilliset kehonliikkeet

REFERENCES

Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F., & Spigler, R. (2005). The kuramoto model: A simple paradigm for synchronization phenomena. *Review of modern physics*, *77*(1), 137-185.

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

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (2000). *Hilgard's introduction to psychology* (13 ed.). Fort Worth, TX: Harcourt College Publishers.

Ball, G., & Breese, J. (2000). Relating personality and behaviour: Posture and gestures. In A. Paiva (Ed.), *Affective interactions: Lecture notes in computer science* (p. 196-203). Berlin: Springer.

Bayne, T. (2004). Closing the gap? Some questions for neurophenomenology. *Phenomenology and the cognitive sciences*, 3(4), 349–264.

Behne, KE., & Wöllner, C. (2011). Seeing or hearing the pianists? A synopsis of an early audiovisual perception experiment and a replication. *Musicae Scientiae*, *15*(3), 324-342.

Bernardin, D., M., T., Wanderley, M., Bardy, B., & Stoffregen, T. (in preparation). Visual perception of musicians' instrument from kinematic displays.

Blacking, J. (1973). *How musical is man?* Seattle, USA: University of Washington Press.

Bregman, A. S. (1990). *Auditory scene analysis. the perceptual organization of sound*. Cambridge, MA: MIT Press.

Broughton, M., & Stevens, C. (2008). Music, movement and marimba: an investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, *37*(2), 137-153.

Bruyn, L. de, Leman, M., Moelants, D., & Demey, M. (2009). Does social interaction activate music listeners? In S. Ystad, R. Kronland-Martinet, & K. Jensen (Eds.), *Computer music, modeling and retrieval: Genesis of meaning of sound and music. Lecture notes in computer science.* (Vol. 5493, p. 93-106). Berlin: Springer-Verlag.

Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2010). Influence of musical features on characteristics of music-induced movements. In S. Demorest, S. Morrison, & P. Campbell (Eds.), *Proceedings of the 11th international conference for music perception and cognition, (ICMPC '11)* (p. 425-428). Seattle, USA: University of Washington.

Cadoz, C., & Wanderley, M. (2000). Gesture-music. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music.* (CD-ROM). Paris: IRCAM/Centre Pompidou.

Camurri, A., De Poli, G., Friberg, A., Leman, M., & Volpe, G. (2005). The mega project: Analysis and synthesis of multisensory expressive gesture in performing art applications. *Journal of New Music Research*, *34*(1), 5–21.

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*(1-2), 213-225.

Casciato, C., Jensenius, A., & Wanderley, M. (2005). Studying free dance movement to music. In *Proceedings of ESCOM 2005 Performance Matters! Conference.* Porto, Portugal: .

Cayley, D. (Producer). (2011, April 11). The malaise of modernity: part one. *CBC Ideas*, Podcast retrieved from http://www.cbc.ca/ideas/episodes/2011/04/11/the-malaise-of-modernity-part-1—5/.

Chandler, D. (2009). Semiotics for beginners. *Semiotics for beginners: Signs*, Retrieved November 9, 2011, from http://www.aber.ac.uk/media/Documents/S4B/sem02.html.

Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the perception of musical performances: Physiological correlates. *Cognition*, *108*, 639-651.

Clark, A. (1997). *Being There: putting brain, body, and world together again*. MIT Press, Cambridge.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Science*, *3*(9), 345-351.

Clarke, E. (1995). *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford, England: Oxford University Press.

Clarke, E., & Davidson, J. W. (1998). The body in performance. In W. Thomas (Ed.), *Composition-performance-reception* (p. 74-92). Aldershot: Ashgate Press.

Clayton, M., Sager, R., & Will, U. (2004). In time with the music: The concept of entrainment and its significance for ethnomusicology. *ESEM Counterpoint*, *1*, 1-45.

Condon, W. (1982). Cultural microrhythms. In M. Davis (Ed.), *Interaction rhythms* (p. 53-77). New York, U.S.A.: Human Resources Press.

Cook, N., & Clarke, E. (2004). Introduction: What is empirical musicology? In E. Clarke & N. Cook (Eds.), *Empirical musicology: Aims, methods, prospects* (p. 3-14). Oxford, England: Oxford University Press.

Cooper, D. E., editor (2001). *Epistemology: The Classic Readings*. Blackwell Publishers Ltd.

Cowart, M. (2005). Embodied cognition. In *Internet Encyclopedia of Philosophy.* Retrieved from http://www.iep.utm.edu/embodcog/: .

Cross, I. (2003). Music and biocultural evolution. In M. Clayton, T. Herbert, & R. Middleton (Eds.), *The cultural study of music: A critical introduction* (p. 19-30). New York: Routledge.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception withou familiarity cues. *Bulletin of the Psychonomic Society*, *9*, 353-356.

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

Davidson, J. W. (1991). *The perception of expressive movement in music performance.* Unpublished doctoral dissertation, City University, London.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*(2), 103-113.

Davidson, J. W. (1994). What type of information is conveyed in the body movements of solo musician performers? *Journal of Human Movement Science*, *6*, 279-301.

Davidson, J. W. (1995). What does the visual information contained in music performances offer the observer? some preliminary thoughts. In R. Steinberg (Ed.), *Music and the mind machine: Psychophysiology and psychopathology of the sense of music,*. Berlin: Springer Verlag.

Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of annie lennox. *Musicae Scientiae*, *5*(2), 235–256.

Davidson, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: a case study approach. *Psychology of Music*, *35*, 381–401.

Davidson, J. W., & Correia, J. S. (2002). Body movement. In R. Parncutt & G. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 237–250). New York, NY: Oxford University Press.

Delalande, F. (1988). Le gestique de Gould: Élements pour une sémiologie du geste musical. In *Glenn Gould pluriel* (p. 85-111). Québec, PQ: Louise Courteau.

Dempster, W. T., Gabel, W. C., & Felts, W. J. L. (1959). The anthropometry of manual work space for the seated subject. *American Journal of Physical Anthropology*, *17*, 289-317.

Dubal, D. (1985). *The world of the concert pianist: Conversations with 35 internationally celebrated pianists*. London: Victor Gollancz.

Eberhart, H. D., & Inman, V. T. (1951). An evaluation of experimental procedures used in a fundamental study of human locomotion. *Annals of the New York Academy of Sciences*, *5*, 1213-1228.

Eerola, T., Luck, G., & Toiviainen, P. (2006). An investigation of pre-schoolers' corporeal synchronization with music. In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th international conference on music perception and cognition.* Bologna, Italy: .

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.

Eitan, Z. (1997). *Highpoint: A study of melodic peaks*. Philadelphia, PA: University of Pennsylvania Press.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–553.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124-129.

Elsdon, P. (2006). Listening in the gaze: The body in Keith Jarrett's solo piano improvisations. In A. Gritten & E. King (Eds.), *Music and gesture* (pp. 209–225). Aldershot, England: Ashgate Publishing Limited.

Erkkilä, J., Gold, C., Fachner, J., Ala-Ruona, E., Punkanen, M., & Vanhala, M. (2008). The effect of improvisational music therapy on the treatment of depression: protocol for a randomised controlled trial. *BMC Psychiatry*, *8*, 50. doi:10.1186/1471-244X-8-50.

Fink, P. W., Foo, P. S., & Warren, W. H. (2009). Catching fly balls in virtual reality: A critical test of the outfielder problem. *Journal of Vision*, *9*(13), 1-8.

Finley, F. R., & Karpovich, P. V. (1964). Electrogoniometric analysis of normal and pathological gaits. *Research Quarterly (supplement)*, *35*, 379-384.

Formisano, R., Vinicola, V., Penta, F., Matteis, M., Brunelli, S., & Weckel, J. W. (2001). Active music therapy in the rehabilitation of severe brain injured patients during coma recovery. *Ann. Ist. Super. Sanita*, 37(4), 627-630.

Frame, J. M. (Lecturer). (2009, September 14). Phenomenology, pragmatism and existentialism - 01. *History of Philosophy and Christian Thought*. Podcast retrieved from feed://deimos3.apple.com/WebObjects/Core.woa/Feed/rts–public–dz.4126928765.04126928767.

Friberg, A., & Sundberg, J. (1999, March). Does music performance allude to locomotion? a model of final ritardande derived from measurements of stopping runners. *Journal of the Acoustical Society of America*, 105(3), 1469-1484.

Gabrielsson, A., & Juslin, P. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of music*, *24*(1), 68.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593-609.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3/4), 455-479.

Gibson, J. J. (1979/1986). *The ecological approach to visual perception*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Godøy, R. I. (2010). Gestural affordances of musical sound. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement and meaning* (p. 103-125). New York: Routledge.

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, & JF. Kamp (Eds.), (Vol. 3881/2006, p. 256-267). Berling Heidelberg: Springer Verlag.

Godøy, R. I., & Jensenius, A. R. (2009). Body movement in music information retrieval. In *Proceedings of the 10th international society for music information retieval conference (ISMIR 2009)* (p. 45-50).

Godøy, R. I., & Leman, M. (Eds.). (2010). *Musical gestures: Sound, movement and meaning*. New York: Routledge.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, *26*, 427-425.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: The Belknap Press of Harvard University Press.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences*, *1156*, 97-117.

Gritten, A., & King, E. (Eds.). (2006). *Music and gesture*. Aldershot, England: Ashgate Publishing Limited.

Gritten, A., & King, E. (Eds.). (2011). *New perspectives on music and gesture*. Farnham, England: Ashgate Publishing Limited.

Gross, J. J. (1999). Emotion and emotion regulation. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality. theory and research* (2nd ed., p. 525-552). New York, NY: Guliford.

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

Hedley, A. (1947). *Chopin*. London: J.M. Dent.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, *48*, 246–268.

Hewlett, W. B., & Selfridge-Field, E. (Eds.). (1998). *Melodic similarity: Concepts, procedures, and applications* (Vol. 11 of Computing in Musicology). Cambridge, MA: MIT Press.

Himberg, T., & Thompson, M. (2009). Group synchronization of coordinated movements in a cross-cultural choir workshop. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & PS. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 08).* Jyväskylä, Finland: University of Jyväskylä.

Himberg, T., & Thompson, M. (2010). Dyadic entrainment and interaction in african dance. In S. Demorest, S. Morrison, & P. Campbell (Eds.), *Proceedings of the 11th international conference for music perception and cognition, (ICMPC '11)* (p. 425-428). Seattle, USA: University of Washington.

Huang, J., & Krumhansl, C. L. (2011). What does seeing the performer add? It depends on musical style, amount of stage behavior, and audience expertise. *Musicæ Scientiæ*, *15*(3), 343-364.

Huron, D. (1999). *Methodology: On finding field-appropriate methodologies at teh intersection of the humanities and the social sciences.* (1999 Ernst Bloch lectures, University of California at Berkley)

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, *930*, 43-61.

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. New York, NY: Wiley-Interscience.

Jensenius, A. R. (2007). *Action–sound: Developing methods and tools to study music-related body movement*. Unpublished doctoral dissertation, University of Oslo. Available at http://www.arj.no/research/phd/.

Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). Musical gestures: Concepts and methods in research. In *Musical gestures: Sound, movement and meaning* (p. 12-35). New York, NY: Routledge.

Johansson, G. (1973). Visual perception of biological motion and a modell for its analysis. *Perception & Psychophysics*, *14*(2), 201-211.

Johansson, G. (1977). Studies on visual perception of locomotion. *Perception*, *6*(4), 365-376.

Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (p. 453-489). New York, U.S.A.: Oxford University Press.

Keller, P. E. (1999). Attending in complex musical interactions: the adaptive dual role of meter. *Australian Journal of Psychology*, *51*(3), 66-175.

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

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

Kemp, M., & Wallace, M. (2011). Motions of the mind. *Universal Leonardo*, Retrieved October 22, 2011, from http://www.universalleonardo.org/essays.php? id=564.

Kendon, A. (1982). The study of gesture: Some remarks on its history. *Recherches Sémiotiques/Semiotic Inquiry*, *2*, 45-62.

Khvedelidze, B. V. (2002). Hilbert transform. In M. Hazewinkel (Ed.), *Encyclopaedia of mathematics.* Berlin: Springer Verlag. Retrieved from http://eom.springer.de/H/h047430.htm.

Kitagawa, M., & Windsor, B. (2008). *MoCap for artists*. Burlington, MA: Focal Press.

Klette, R., & Tee, G. (2008). Understanding human motion: A historic review. In B. Rosenhahn, R. Klette, & D. N. Metaxas (Eds.), *Human motion: Understanding, modeling, capture and animation* (p. 1-22). Dordrecht, The Netherlands: Springer.

Knapp, M. L., & Hall, J. A. (2009). *Nonverbal communication in human interaction*. Boston, M.A.: Wadsworth Pub Co.

Knuf, L., Aschersleben, G., & Prinz, W. (2001). An analysis of ideomotor action. *Journal of Experimental Psychology: General*, *130*(4), 779-798.

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

Kotlyar, G. M., & Morozov, V. P. (1976). Acoustic correlates of the emotional content of vocalized speech. *Soviet Physics*, *22*, 370-376.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*, 575-580.

Krumhansl, C. L., & Schenck, D. L. (1997). Can dance reflect the structural and expressive qualities of music? A perceptual experiment of Balanchine's choreography of Mozart's Divertimento No. 15. *Musicæ Scientiæ*, *1*(1), 63-85.

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

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. New York, NY: Basic Books.

Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.

Leman, M., Desmet, F., Styns, F., Noorden, L. van, & Moelants, D. (2007). Embodied listening performances reveal relationships between movements of player and listeners. In *Proceedings of ENACTIVE/07: 4th international conference on enactive interfaces* (p. 141-144). Grenoble, France: .

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lesaffre, M., Voogdt, L., Leman, M., Baets, B., Meyer, H., & Martens, J. (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.

Levitin, D. J. (2008). *The world in six songs: How the musical brain created human nature*. New York: Dutton.

Levitin, D. J., McAdams, S., & Adams, R. L. (2002). Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, *7*(2), 171-189.

Levitin, D. J., Nuzzo, R. L., Vines, B. W., & Ramsay, J. O. (2007). Introduction to functional data analysis. *Canadian Psychology*, *48*(3), 135-155.

Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *Annals of the New York Academy of Sciences*, *1156*(1), 211-231.

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

Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, *20*(1), 23-47.

Livingstone, S., Thompson, W., & Russo, F. (2009). Facial expressions and emotional singing: A study of perception and production with motion capture and electromyography. *Music Perception*, *26*(5), 475–488.

Luck, G. (2000). Synchronising a motor response with a visual event: The perception of temporal information in a conductor's gestures. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the 6th international conference on music perception and cognition (ICMPC6).* (CD-ROM). Keele: Keele University.

Luck, G., & Nte, S. (2008). An investigation of conductors' temporal gestures and conductor musician synchronization, and a first experiment. *Psychology of Music*, *36*(1), 81-99.

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.

Luck, G., Saarikallio, S., & Toiviainen, P. (2008). Personality traits correlate with characteristics of music-induced movement. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & PS. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 08)* (p. 276-279). Jyväskylä, Finland: University of Jyväskylä.

Luck, G., & Sloboda, J. (2007). An investigation of musicians' synchronization with traditional conducting beat patterns. *Music Performance Research*, *1*, 26-46.

Luck, G., & Sloboda, J. (2008). Exploring the spatio-temporal properties of simple conducting gestures using a synchronization task. *Music Perception*, *25*(3), 225.

Luck, G., & Sloboda, J. (2009). Spatio-temporal cues for visually mediated synchronization. *Music Perception*, 465–473.

Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: an automated feature-extraction analysis. *Music Perception*, *24*(2), 189-199.

Marchand, D. J. (1975). A study of two developing approaches to developing expressive performance. *Journal of Research in Music Education*, *23*(1), 14-22.

McBeath, M., Shaffer, D., & Kaiser, M. (1995). How baseball outfielders determine where to run to catch fly balls. *Science*, *268*(5210), 569-573.

McNeil, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeil, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.

Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Paris: Gallimard.

Morais Ribeiro, H. de. (1998). *On the philosophy of cognitive science.* Retrieved August 26, 2011 from http://www.bu.edu/wcp/Papers/Cogn/CognRibe.htm

Morris, J. R. W. (1973). Accelerometry—a technique for the measurement of human body movements. *Journal of Biomechanics*, *6*, 729-736.

Müller, M. (2007). *Information retrieval for music and motion*. Berlin: Springer Verlag.

Naveda, L., & Leman, M. (2010). The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA). *Music Perception*, *28*(1), 93–111.

Nusseck, M., & Wanderley, M. (2009). Music and motion-how music-related ancillary body movements contribute to the experience of music. *Music Perception*, *26*(4), 335–353.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Champaign, IL: University of Illinois Press.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*(1), 115–138.

Palmer, C., Koopmans, E., Carter, C., Loehr, J., & Wanderley, M. (2009). Synchronization of motion and timing in clarinet performance. In A. Williamon, S. Pretty, & R. Buck (Eds.), *Proceedings of the international symposium on performance science* (p. 159-164). Utrecht: European Association of Conservatories (AEC).

Palmer, C., & Krumhansl, C. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 728.

Pfeifer, R., & Scheier, C. (1999). *Understanding intellligence*. Cambridge, MA: MIT Press.

Phillips-Silver, J., & Trainor, L. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430.

Phillips-Silver, J., & Trainor, L. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533–546.

Phillips-Silver, J., & Trainor, L. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*(1), 94–102.

Price, H. E., & Byo, J. L. (2002). Rehearsing and conducting. In R. Parncutt & G. E. McPherson (Eds.), *The science & psychology of music performance* (p. 335-351). New York, U.S.A.: Oxford University Press.

Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis*. New York, NY: Springer Verlag.

Reed, A. (2011). *Wii-bändi: designing an expressive digital musical interface for group therapeutic use with disabled people.* Master's thesis, University of Jyväskylä, Jyväskylä, Finland.

Rentfrow, P., & Gosling, S. (2003). The Do Re Mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*, 1236-1256.

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

Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science*, *27*(3), 423-456.

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

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis fore person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585-615.

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

Ryle, G. (2001). What is mind? In D. Donderi (Ed.), *Psychology: the science of mind and behavior* (p. 1-9). Dubuque, Iowa: Kendall/Hunt Publishing Company.

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

Schoen-Nazzaro, M. B. (1978). Plato and Aristotle on the ends of music. *Laval théologique et philosophique*, *34*(261-272).

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, *21*(4), 561–585.

Schubert, E., & Dunsmuir, W. T. M. (2004, August). Introduction to interupted time series analysis of emotion in music: the case of arousal, valence and points of rest. In S. Lipscomb, R. Ashley, R. Gjerdingen, & P. Webster (Eds.), *Proceedings of the 8th international conference on music perception and cognition (icmpc8)* (p. 445-448). Adelaide, Australia: Causal Productions.

Schutz, M. (2008). Seeing music? what musicians need to know about vision. *Empirical Musicology Review*, *3*(3), 83-108.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*, 888-897.

Seashore, C. E. (1936). The psychology of music. ii. approaches to the experimental psychology of music. *Music Educators Journal*, *22*(6), 22-23.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Science*, *10*(2), 70-76.

Shapiro, L. (2007). The embodied cognition research programme. *Philosophy Compass*, *2*, 338-346.

Shove, P., & Repp, B. H. (1995). The practice of performance. In J. Rink (Ed.), (p. 55-83). Cambridge University Press.

Spiegelberg, S. (2006). Book review: Eric clarke, ways of listening: An ecological approach to the perception of musical meaning. *Empirical Musicology Review*, *1*(2), 127-129.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thompson, E. (2007). *Mind in life: Biology, phenomonology and the sciences of mind*. Cambridge, MA: Harvard University Press.

Thompson, M. R., & Luck, G. (2008). Effect of pianists' expressive intention on amount and type of body movement. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th international conference on music perception and cognition, (ICMPC '10)* (p. 540-544). Sapporo, Japan: Hokkaido University.

Thompson, W., Graham, P., & Russo, F. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *156*(1/4), 203–227.

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

Todd, N. P. M., O'Boyle, D. J., & Lee, C. S. (1999). A sensory-motor theory of rhythm, time perception and beat induction. *Journal of New Music Research*, *28*(1), 5-28.

Toiviainen, P., & Burger, B. (2009). *Mocap toolbox manual*. Jyväskylä, Finland: University of Jyväskylä. Available at http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual.

Toiviainen, P., & Keller, P. E. (2010). Special issue: spatiotemporal music cognition. *Music Perception*, *28*(1), 1-2.

Toiviainen, P., Luck, G., & Thompson, M. R. (2009). Embodied metre in spontaneous movement to music. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & PS. Eerola (Eds.), *Proceedings of the 7th triennial conference of european society for the cognitive sciences of music, (ESCOM '09)* (p. 526-529). Jyväskylä, Finland: University of Jyväskylä.

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

Trevarthen, C., Kokkinaki, T., & Fiamenghi Jr, G. (1999). What infants' imitations communicate: With mothers, with fathers and with peers. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (p. 127-185). Cambridge: Cambridge University Press.

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

Verron, C. (2005). *Traitement et visualisation de données gesturalles captées par optotrak [processing and visualizing gesture data captured by optotrak].* (Unpublished lab report. McGill University, Input Devices and Music Interaction Laboratory.)

Vines, B., Krumhansl, C., Wanderley, M., & Levitin, D. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113.

Wanderley, 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*(1), 97–113.

Wanderley, M. M. (2002). Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and sign language in human-computer interaction* (p. 241-253). Berlin: Springer Verlag.

Wanderley, M. M., Depalle, P., & Warusfel, O. (1999). Improving instrumental sound synthesis by modeling the effects of performer gesture. In *Proceedings of the 1999 international computer music conference* (p. 418-421). Beijing: .

Weiss, P., & Taruskin, R. (Eds.). (1984). *Music in the Western world: A history in documents*. New York, NY: Simon & Schuster.

Williamon, A., & Davidson, J. W. (2002). Exploring co-performer communication. *Musicæ Scientiæ*, *6*(1), 53-72.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625-636.

Wilson, R. A., & Foglia, L. (2011). Embodied cognition. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2011 ed.). Retrieved from http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition/.

Winter, D. A. (2009). *Biomechanics and motor control of human movement*. Hoboken, NJ: John Wiley & Sons Inc.

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

ORIGINAL PAPERS

I

EFFECT OF PIANISTS' EXPRESSIVE INTENTION ON AMOUNT AND TYPE OF BODY MOVEMENT

by

Marc R. Thompson & Geoff Luck

In Miyazaki, K., Hiraga, Y., Adachi, M., Nakajima, Y. & Tsuzaki, M. (Eds.), Proceedings of the 10th International Conference on Music Perception and Cognition, (ICMPC10) (pp. 540-544). Sapporo, Japan: Hokkaido University.

Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10). Sapporo, Japan.
Ken'ichi Miyazaki, Yuzuru Hiraga, Mayumi Adachi, Yoshitaka Nakajima, and Minoru Tsuzaki (Editors)

Effect of pianists' expressive intention on amount and type of body movement

Marc R. Thompson*[1], Geoff Luck*[2]

* *Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, University of Jyväskylä, Finland*
[1]`mathomps@jyu.fi`, [2]`luck@campus.jyu.fi`

ABSTRACT

Body movement displayed in music performance is said to be an overt manifestation of the musician's expressive intentions and goals regarding the music being played. We are interested in studying if different levels of expression result in different amounts of body movement and gestures. For this study, musicians were asked in multiple sessions to play an excerpt from the same piece using three different levels of expression while their movements were recorded using an optical motion capture system. Statistical tests show that an increasing amount of expression resulted in more body movement and that the amount of expression and physical movement were inter-twined. Also, we present data suggesting that the head and shoulders travelled a further distance overall, and showed bigger differences between performance manners, compared to the fingers, wrists and lower back. We hypothesize that this is related to the contrasting roles these parts of the body play in piano performance.

Keywords: Musical expression, embodied music cognition, Motion Capture

I. INTRODUCTION

In music performance, the musician's embodiment of musical events demonstrates the complex yet quantifiable mappings between musical expression and motor movement.

Our current research speculates that the quantity of body movement in a performance is driven by the amount of expression exploited by the performer. For example, if the music lacks specific expressive goals, the musician will restrain him/herself from effectuating what an audience would perceive as being expressive body movement or gestures. Past studies have shown audiences are able to differentiate between different levels of expression during performances in both the auditory (Kendall & Carterette, 1990) and visual domains (Davidson, 1993 & 1995) or recognize different emotions from watching performances without sound (Dahl & Friberg, 2003).

Perception aside, we are interested in quantifying the different amounts of movement employed by musicians when asked to perform using different levels of expression. We have chosen to study pianists because unlike smaller wind and string instruments who may be performing while standing, pianists are restricted in the amount of expressive movement they can execute during performances.

If one takes the embodied view of human cognition (e.g., Varela, Thompson & Rosch 1991; Port & van Gelder, 1995), that cognitive processes are governed by an organism's sensorimotor capacities, body, and environment, one can see that musical expression and bodily movement are inextricably connected. There is no music without movement, no musical expression without expressive movement. Similarly, when we hear music, we parse the elements of the music through, for example, body movement, such as foot tapping or body-sway. At times, our comprehension of the actions responsible for producing music is undetected at a conscious level, but activation of so-called 'mirror neurons' in the brain (e.g., Rizzolatti, Fadiga, Gallese & Fogassi, 1996) reveal its presence nonetheless.

This concept has opened up a new way of studying music expression and performance. Marc Leman (2008) notes that in the last twenty years, there has been 'growing awareness that gesture and motor-based processing of musical content play an important role in connecting musical mind and matter (pg. 27)'. Therefore, studying the relationships between musical expression and body movement can be incorporated into the wider research framework of mind-matter relationships.

Corporal movement as an overt manifestation of the musician's expressive intentions and goals has been the subject for various case studies of famous musicians known for expressive, sometimes eccentric physical display during performances (e.g. Thompson, Graham & Russo, 2005; Eldson, 2006; Davidson, 2006) or even during studio recorded performances (e.g. Delalande, 1988, 1995, etc.). These studies demonstrate that the body movement attains meaning by adding an extra-musical dynamic to performances because they are presented in a social or idiosyncratic context. Meaningful body movement has also been examined in studies involving co-performer communication (Williamon & Davidson, 2002) and conductor-musician interaction (Luck & Toiviainen, 2006). Finally, embodied music cognition can be used in developmental studies where much can be learned about how humans learn to synchronize to rhythms and music from an early age (Eerola, Luck & Toiviainen, 2006).

In perception studies, body movements and gestures are said to play a significant role in being able to recognize the expressive intentions (Davidson 1993, 1995) or emotional states (Dahl & Friberg, 2003) of musicians, or musical phrasal boundaries as interpreted by a dancer (Krumhansl & Schenck, 1997).

It has been noted that different types of gestures convey different types of expressivity. For example, in a study where observers were asked to continuously rate tension and phrasing of bodily movements in a series of clarinet performances (Vines, Krumhansl, Wanderley & Levitin, 2006), it was found that gestures associated with tension were related to expressivity while gestures associated with phrasing indicated musical structure. Related to this, musical gestures can be categorized as being either instrumental (having to do with the production of sound) or ancillary (an accompanying gesture) (Wanderley, 1999; Cadoz & Wanderley, 2000).

In piano performance particularly, motor movements can be broken down into two loose dichotomies: movements used for function and movements used for expressivity.

ISBN: 978-4-9904208-0-2

The category in which a specific movement falls into may depend on the degree of freedom allocated to that body part during performance. The fingers, wrists and lower back, for example, are physically required to play the notes on the piano and their total movement is restricted by the specific music being played. The shoulders and the head, meanwhile, are freer to move and not as involved in the production of sound. Therefore, we might expect a pianist to use their head and shoulders, as opposed to their fingers, wrists, or lower back, to embody musical expression.

In the present study, three pianists were asked to play the same piece (Brahms: Intermezzo in A major, Opus 118 # 2) in three different pre-determined expressive dispositions. For the purposes of this study, we were interested in analysing the contribution of the arms, torso, head, etc., to the total amount of movement used for different levels of expression. To do this, we applied computational and statistical analysis techniques to motion capture data obtained from the pianists' performances. It was predicted that increased expression would result in an increased amount of body movement. In addition, it was predicted that the shoulders and head would contribute most to different amounts of expression, showing the most variation in amount of movement, while the fingers and hands would contribute the least.

II. METHOD

A. Participants

Three pianists volunteered to take part in the present study. All three were competent pianists, and each had at least 10 years of playing experience. Player 1 (b. 1982, Finnish, 15 years playing experience) was a student in Piano Pedagogy at the Jyväskylä University of Applied Sciences. Player 2 (b. 1978, Finnish, 22 years playing experience) had received her Bachelors in Piano Performance from the same university one year earlier. Player 3 (b. 1977, Hungarian, 17 years playing experience) was a visiting researcher in Musicology at the University of Jyväskylä's music department. All participants received an honorarium for their participation in the form of gift vouchers.

B. Design and procedure

Each participant attended between two and four recording sessions in a professional recording studio, during each of which they were instructed to play the first sixteen measures of the Brahms Intermezzo in A major Opus 118 #2 three times, each time using one of three levels of expression: Minimum expression, normal expression and maximum expression. The words movement and gesture were avoided as far as possible. The goal was to let each pianist interpret for him or herself what was meant by different levels of expression.

Audio recordings of all performances were made using a high quality microphone and ProTools recording software. In addition, fifteen reflective markers were attached to key locations on the body (four on the head, one on each shoulder, one at the centre of the back, two on the lower back, two on the elbows, two on the wrists, one on each middle finger), and their three-dimensional spatial position recorded at 120 fps using an eight-camera optical motion capture system (Qualisys ProReflex). Additionally, two markers were placed at each end of the keyboard to act as reference points.

III. RESULTS

Once the motion capture data was imported into MATLAB, we calculated the Euclidean distance between subsequent Cartesian (x, y, z) coordinates for each marker, resulting in a matrix of distance values. The data was then simplified by averaging the columns that corresponded to the same parts of the body. For example, the four head markers were averaged, as were the left and right shoulders, elbows, etc. The last step was to sum the distance matrix, resulting in six values for each performance, which represent the total Euclidean distances traveled for the head, shoulders, lower back, elbows, wrists and fingers.

Figure 1a demonstrates the total Euclidean distance values for the head, shoulders, lower back, elbows, wrists and fingers. From this figure, one can recognize that for Player 2 and Player 3, the amount of body movement increased as more expression was employed to the performances. In session 1, for example, the markers attached to Player 3's body traveled a total distance of 1.6 meters in the Minimum Performance, 2.9 meters in the Normal Performance and 4.7 meters in the Maximum Performance.

Because of the intended ambiguity of the instructions, there was no control over the tempi chosen by the participants. This resulted in performances of varying length (*Normal Expression:* mean length = 53.4 sec., std = 7.9; *Minimum Expression:* mean length = 53.6 sec., std = 15.3; *Maximum Expression:* mean length: 55.5 sec., std = 14.5 sec.). Evidently, the longer performances contained markers with greater distance values. We found a large variance from player to player and from session to session. For this reason we thought it useful to normalise the data so that the actual length of the performance would be irrelevant. We randomized the vertical order of the cells in each Euclidean distance matrix and truncated the data after the first 3500 cells (the length of the shortest performance). Figure 1b shows the normalized data. The differences in relationships between both figures are minimal but reflect the effectiveness of having the data normalised.

For example, Player 1, in session 1, played the Minimum Performance with a total distance value of 4.3 metres (78 seconds long) while in Session 2, the total distance value during her minimum performance was 3.8 meters (70 seconds long). The normalised data, however, shows that in Session 1, the distance travelled was comparatively less then in Session 2. Normalising the data in this way was useful because it removed differences in performance length, thus revealing the true relationship between level of expression and amount of movement.

Figure 1c shows the mean inter-frame distance, which can additionally be interpreted as the average inter-frame *speed*. For example, on average, Player 1 in Session 1's minimum performance travelled a distance of 4 millimetres within 120th of a second or a speed 4.8 centimetres per second.

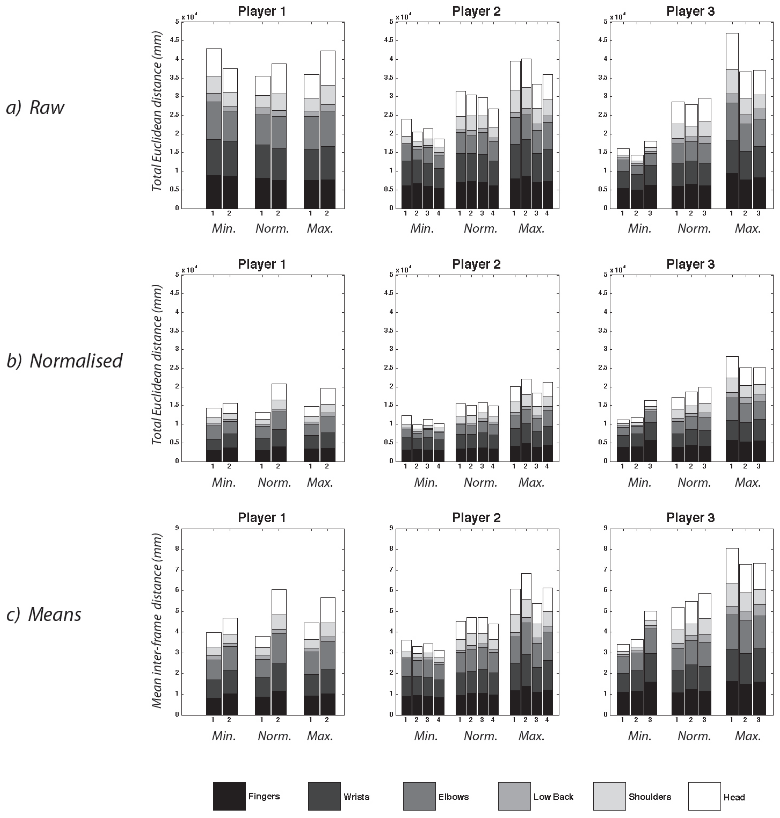


Figure 1. a) The Euclidean distance values for the raw data. b) The Euclidean distance for the normalized data. c) The average inter-frame distance shows the average speed for each body part (e.g. Player 2' s head markers, in Session 1, minimum performance traveled at an average speed of 0.6 mm per 120th of a second, or 72 mm/second).

C. Functional versus expressive movements

From Figure 1, we can discern changes in the amount of movement between the levels of expression and also eyeball which part of the body had the highest change. It appears that the total Euclidean distance values for the lower back, wrists and fingers have less variance then the total Euclidean distance values for the head, shoulders and elbows. This claim is strengthened by the data in Table 1, which show that the head and shoulders had more variance in inter-frame distance values than did the lower back and wrists (i.e. higher standard deviations). The head, shoulder and elbows contributed to most of the differences in movement between levels of expression, which is revealed by the higher standard deviations in Table 1. This may have occurred because the body parts with the least difference in movement between expression levels have movements that are related to the functionality of playing the piano. In other words, the amount of expressive movement possible for the fingers, wrists and lower back are limited by the physical act of playing the piano. For pianists, the head, shoulders and elbows have a higher degree of freedom. It would seem that these body parts are used more habitually for expressive body movement then are the wrists and fingers.

D. Statistical Analysis

Using the normalized data matrices, the effects of session, location, and expression on amount of body movement were examined by performing a three-way

Table 1. The average Euclidean distance traveled for all Players, all Sessions and all performance manners.

Markers (all)	Averages of Total Euclidean distances values (meters)	Standard Deviation (meters)
Head	3.0	1.3
Shoulders	1.8	0.888
Lower back	0.766	0.418
Elbows	3.3	1.1
Wrists	4.1	0.794
Fingers	4.0	0.823

Table 2. Summary of all main effects and interactions from the three ANOVA's. Estimations of effect size are based on Partial Eta Squared.

	df	*F*	*Sig.*	*Effect size*
Participant 1				
Expression	2, 83976	295.10	<.001	.007
Session	1, 41988	2867.83	<.001	.064
Location	5, 41988	4781.85	<.001	.363
*Expression*Session*	2, 83976	302.31	<.001	.007
*Expression *Location*	10, 83976	27.43	<.001	.003
*Expression*Session *Location*	10, 83976	45.10	<.001	.005
*Session*Location*	5, 41988	155.78	<.001	.018
Participant 2				
Expression	2, 167952	10769.68	<.001	.114
Session	3, 83976	174.739	<.001	.006
Location	5, 83976	14419.11	<.001	.462
*Expression*Session*	6, 167952	206.42	<.001	.007
Expression Location*	10, 167952	333.20	<.001	.019
*Expression*Session *Location*	30, 167952	18.438	<.001	.003
*Session*Location*	15, 83976	41.10	<.001	.007
Participant 3				
Expression	2, 125964	7150.22	<.001	.102
Session	2, 62982	245.28	<.001	.008
Location	5, 62982	7361.15	<.001	.369
*Expression*Session*	4, 125964	279.84	<.001	.009
*Expression *Location*	10, 125964	420.98	<.001	.032
*Expression*Session *Location*	20, 125964	42.09	<.001	.007
*Session*Location*	10, 62982	48.16	<.001	.008

mixed ANOVA, using total distance travelled as the dependent variable. All main effects and interactions were statistically significant, and are summarised in Table 2. It can be seen from the overall effect sizes that the largest real-world effects were for expression and location, with session and all interactions having very small effects (although session seemed more important for participant 1). The fact that all effects were statistically significant was likely due to the huge number of data points analysed. Post-hoc tests (with Bonferroni correction) showed that all pairwise comparisons were statistically significant ($p \leq .001$) except for participant 2 between sessions 3 and 4 (ns), and between elbows and fingers (ns), and for participant 3 between sessions 1 and 2 ($p < .05$), and between wrists and elbows ($p < .05$).

IV. DISCUSSION

This study explored novel ways in which the intricate mappings between musical expression and body movement displayed in piano performance could be quantized using computational feature extraction and statistical methods.

Our statistical analysis determined that having participants replicate the procedure in multiple sessions had little real-world effect on the amount of movement used in performances. At the start of the study, all participants needed the score to play the Brahms Intermezzo. But by the end of the study, both Player 2 and Player 3 had memorized the score and played the piece with noticeable confidence. We had expected these changes to be reflected in the movement data. But because our analysis was focused on the total amount of movement and not fluency, we were not able to show how the musicians, within on-going sessions, might have developed more articulated and controlled expressive body movements. It would be interesting in future studies to explore ways in which fluency and accuracy features could be extracted from movement data.

In this study, we were limited to analyzing the behaviour of only three pianists. However, we recognize that musical expression itself is driven by the performer's temperament, education, background, etc. For this reason, interviews were conducted with the participants. Unsurprisingly, each participant had quite contrasting views of what it meant to play with or without expression. Player 1, who varied the least between performance manners, defined a minimum expression as playing with strong adherence to the score's dynamic markings. Player 3, whose amount of movement varied the most within performance manners, played with minimum expression by simply ignoring all of the score's dynamic markings and playing in effect, the most mechanic way possible. Observations such as these add real-world personality driven perspective to our data. Future research is planned which will focus more fully on the qualitative data collected during our trials.

Our computational analysis has shown that increasing the amount of expression may incline a pianist to use a greater amount of body movement in a performance. Of the parts of the body analyzed, it was found that the head and shoulders contribute the most to conveying expression in piano performance while the amount of finger and wrists movement showed smaller deviations from the average amount of movement.

We are inclined to relate this observation to Wanderley's gesture categories. Instrumental gestures are related to functional body movements such as wrist and finger movements, which are required for the physical act of playing the piano thus having lower degrees of freedom. By contrast, the head and shoulders are less involved in the production of sound and could be said to be ancillary. Their higher degree of freedom promotes them to be apt conveyers of expression. Therefore, the expressive movements accomplished by the head and shoulders in piano performance could be said to have more significance within the framework of embodied music cognition, which is a musically oriented abstraction of the far broader field of mind-matter relationship studies.

Our research has practical implications for music perception and cognition. By exploring the intricate relationships between expressivity and body movement, we gain a deeper understanding of what it means to be musical. For this reason, our research has applications in areas including piano pedagogy, performance modelling and gesture controlled musical interfaces.

REFERENCES

Cadoz, C. & Wanderley M.M., (1999). Gesture - music. In M. Wanderley, M. Battier, (Eds.), *Trends in Gestural Control of Music*. Ircam.

Davidson, J. (1993). Visual Perception of Performance Manner in the Movements of Solo Musicians. *Psychology of Music, 21*, 103-113.

Davidson, J. (1995). What Does the Visual Information Contained in Music Performances Offer the Observer? Some Preliminary Thoughts. *Music and the Mind Machine: The Psychopathology of the Sense of Music XXX*: pp. 105-113.

Davidson, J. (2006). 'She's the one': Multiple Functions of Body Movement in a Stage Performance by Robbie Williams. In A. Gritten & E. King (Eds.), *Music and Gesture* (pp. 209-225). Hampshire: Ashgate Publishing, Ltd.

Dahl, S., A. Friberg (2003). What Can The Body Movements Reveal About a Musician's Emotional Intentions. *Proceedings of the Stockholm Music Acoustics Conference, (SMAC 03)* (pp. 599-602). Sweden: Stockholm.

Delalande, F. (1988). La gestique de Gould. In G. Guertin (Ed.), Glenn *Gould: Pluriel*, (pp. 85-111). Montréal: Courteau.

Delalande, F. (1995). Meaning and behavior patterns: The creation of meaning in interpreting and listening to music. In E. Tarasti (Ed.) *Musical Signification Essays in the Semiotic Theory and Analysis of Music*. (pp. 219-228). Berlin: Mouton de Gruyter.

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

Eldson, P. (2006). Listening in the gaze: the body in Keith Jarrett's solo piano improvisations. In A. Gritten & E. King (Eds.), *Music and Gesture* (pp. 209-225). Hampshire: Ashgate Publishing, Ltd.

Kendall, R. A. & Carterette, E. C. (1990). The communication of musical expression. *Music Perception, 8,* 129-164.

Krumhansl, C.L., & Schenck, D.L. (1997). Can dance reflect the structural and expressive qualities of music?, *Musicæ Scientiæ, 1*(1), 63-85.

Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge, Mass.: MIT Press.

Luck, G., Toiviainen, P., (2006). Ensemble Musicians' Synchronization with Conductor's Gestures: An Automated Feature-Extraction Analysis, *Music Perception, 24*, pp. 189-199.

Port, R. F., & Van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Mit Press.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research, 3*(2), 131-141.

Thompson, W.F., Graham, P., & Russo, F.A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica 156*(1), 203–227.

Wanderley, M.M. (1999). Non-obvious performer gestures in instrumental music. *Proceedings of the International Gesture Workshop (GW'99)* (pp. 37-48), Gif-sur-Yvette, France.

Williamon, A. & Davidson, J. W. (2002). Exploring co-performer communication. *Musicæ Scientiæ, 1*(1), 53-72.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind.* MIT Press Cambridge, Mass.

Vines, B.W., Krumhansl, C.L., Wanderley, M.M., & Levitin D.J. (2006). Cross-modal interactions in the perceptions of musical performance. *Cognition, 101*, 80-103.

II

PERCEPTION OF EXPRESSION IN CONDUCTORS' GESTURES: A CONTINUOUS RESPONSE STUDY

by

Geoff Luck, Petri Toiviainen & Marc R. Thompson

Music Perception vol. 28(1), 47-57.

PERCEPTION OF EXPRESSION IN CONDUCTORS' GESTURES: A CONTINUOUS RESPONSE STUDY

GEOFF LUCK, PETRI TOIVIAINEN,
& MARC R. THOMPSON
University of Jyväskylä, Jyväskylä, Finland

THE VISUAL CHANNEL HAS BEEN SHOWN TO BE MORE informative than the auditory channel in perceptual judgments of a performer's level of expression. Previous work has revealed a positive relationship between amplitude of music-related movement and ratings of expression, for example, and observers have been shown to be sensitive to kinematic features of music-related movement. In this study, we investigate relationships between the kinematics of a conductors' expressive gestures and ratings of perceived expression. Point-light representations (totalling 10 minutes) of two professional conductors were presented to participants who provided continuous ratings of perceived valence, activity, power, and overall expression using a virtual slider interface. Relationships between these ratings and 11 kinematic variables computationally extracted from the movement data were subsequently examined using linear regression. Higher levels of expressivity were found to be conveyed by gestures characterized by increased amplitude, greater variance, and higher speed of movement.

Received November 23, 2009, accepted April 7, 2010.

Key words: expressiveness, movement kinematics, regression, continuous response, point-light display

PREVIOUS RESEARCH HAS HIGHLIGHTED relationships between corporal behavior and perception of musical expression. It has been shown, for example, that observers are able to perceive a musician's expressive intention through the visual mode alone (Davidson, 1993, 1995), while expressive body movement has been found to influence observers' level of interest in the performance (Broughton & Stevens, 2008). Expressive body movement can also be a determining factor in discriminating between specific emotional intentions (Dahl & Friberg, 2007), while a musical structure that represents musical form, such as phrase structure, can also be perceived through movement alone (Krumhansl & Schenck, 1997). In addition, emotional characteristics of dancers can be perceived in their movements (e.g., Dittrich, Troscianko, Lea, & Morgan, 1996).

In addition to identifying specific emotions, observers can also perceive varying *magnitudes* of expression through the visual channel alone. While the ability to discriminate between different magnitudes of musical expression had already been studied in the auditory mode (Kendall & Carterette, 1990), Davidson (1993) reported that the visual channel also played a significant role in recognizing different performance manners. Further, Davidson found that for nonmusicians, it was the visual mode that most clearly specified the different expressive manners (deadpan, projected, and exaggerated).

The way the body conveys different magnitudes of expression may be related to the amplitude of the gestures produced. Davidson (2007) studied a video of a pianist performing a Beethoven Bagatelle in the deadpan, projected, and exaggerated expressive intentions. She identified expressive locations in which the same body movements in the head and hands were repeated throughout every intention. Although the movements were similar in shape throughout all performance manners, the movements were smallest in the deadpan intention and largest in the exaggerated intention.

In a related study, two of the present authors investigated whether playing with different magnitudes of expression would result in different amounts of body movement (Thompson & Luck, 2010). Pianists performed in multiple sessions an excerpt from the same piece using three different levels of expression ("minimum", "normal", and "maximum"), and their movements were recorded using an optical motion capture system. Data showed that the head and shoulders traveled a further distance overall, and showed bigger differences between expressive intentions, compared to the fingers, wrists and lower back, indicating a positive relationship between range of motion of a body part and its use in conveying expression.

Music Perception VOLUME 28, ISSUE 1, PP. 47–57, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 © 2010 BY THE REGENTS OF THE UNIVERSITY OF CALIFORNIA. ALL RIGHTS RESERVED. PLEASE DIRECT ALL REQUESTS FOR PERMISSION TO PHOTOCOPY OR REPRODUCE ARTICLE CONTENT THROUGH THE UNIVERSITY OF CALIFORNIA PRESS'S RIGHTS AND PERMISSIONS WEBSITE, HTTP://WWW.UCPRESSJOURNALS.COM/REPRINTINFO.ASP. DOI:10.1525/MP.2010.28.1.47

Changes in amplitude of movement, then, may be a cue that observers use when judging the expressivity of a performance. An additional hypothesis is that observers react to more detailed features of the performer's movements; that is, the movement kinematics. It has been shown, for instance, that observers are sensitive to the movement kinematics of conductors' time-beating gestures. Several studies have identified the features of conducting gestures that offer cues for synchronization, employing a range of paradigms to do so, including tapping (Luck & Sloboda, 2008, 2009), temporal adjustment (Luck, 2008), and live performance (Luck & Toiviainen, 2006).

In Luck and Sloboda's (2008) study, participants synchronized finger taps with point light stimuli derived from simple conducting gestures. A series of linear regression analyses identified acceleration along the trajectory as the main cue for synchronization, with this movement feature accounting for up to 68% of the variance in participants' responses. A follow-up study, in which participants synchronized with extended sequences of beats in the form of traditional three-beat patterns (Luck & Sloboda, 2009) again identified absolute acceleration along the trajectory as the main cue for synchronization. A study in which participants adjusted the timing of an auditory stimulus until it was perceived as being in synchrony with a series of point-light conducting gestures identified acceleration along the trajectory as the sole cue for synchronization (Luck, 2008). In a more ecological study (Luck & Toiviainen, 2006), a series of cross-correlations identified acceleration along the trajectory as the primary feature with which ensemble musicians synchronized with a conductor in a live setting, while also showing that the ensemble tended to lag behind the conductor. In several of the above studies, speed of movement was also found to be a relevant kinematic feature, but less so than acceleration.

Given, then, that observers are sensitive to certain kinematics features of temporal conducting gestures, might they also be sensitive to the kinematics of more expressive gestures? The aim of the present study was to collect some preliminary data on this issue. We investigated relationships between the kinematics of conductors' expressive gestures and ratings of perceived expression in order to identify whether expression is conveyed simply by amplitude of the gestures, or if observers are, in addition, sensitive to more subtle features of the movements. Since the expression conveyed in conductors' gestures does not remain static, but rather changes and evolves over time, we employed a continuous response paradigm. This methodology is used widely in studies of emotion and music because of its dynamic nature of data collection – participants' responses can be collected throughout the duration of a temporally extended stimulus, such as a conductor conducting a musical passage.

We combined the continuous response paradigm with a computational feature-extraction approach in which selected movement features were extracted from the gestures at regular intervals, and relationships between these features and the continuous ratings subsequently investigated. Two of the present authors have used this combination of methods in the investigation of emotional responses to musical features in music therapy improvisations (Luck et al., 2008), and in diagnosing level of mental retardation from music therapy improvisations (Luck et al., 2006). While these two studies examined perception of various types of auditory (as opposed to movement) features, the underlying principle of examining relationships between continuous ratings and computationally extracted features was the same. Wöllner and Auhagen (2008) have employed the continuous response paradigm in investigating conductors' expressive gestures, but without detailed kinematic analysis of the gestures themselves. Instead, video recordings were analyzed using the EyesWeb (Camurri & Volpe, 2004) Quantity of Motion (QoM) module, and relationships between amount of movement and ratings of expression examined.

The movement features extracted in the present study related to position, speed, acceleration, and jerk of different body parts. The first three of these were included because of their relevance in previous studies (position indicating magnitude of movement, and speed and acceleration being important in temporal conducting gestures). Jerk, meanwhile, was also included because of the potential relationship between jerkiness or smoothness of movement and emotional expression.

The present study adopted the dimensional concept of emotion in order to investigate emotional expression in a more detailed fashion (as opposed to asking observers to rate only perceived expression as a single concept). It has been suggested that the dimensional concept of emotion is particularly well-suited to studies that examine the dynamic changes in music-related emotional expression (Juslin & Sloboda, 2001). Dimensional theories of emotion hold that emotional meaning can be described within a multidimensional emotion space comprised of a small number of dimensions, most frequently cited as relating to valence, activity, and power (e.g., Osgood, Suci, & Tannenbaum, 1957). Each dimension is assumed to be anchored by semantic terms representing polar opposites, such as happy–sad for valence, active–inactive for activity, and weak–strong for potency. Schubert (2001) notes that references to the first two of these dimensions are frequently found in the music-emotion literature. Power, however, is somewhat less frequently described, and its role in the emotion space not so clearly defined. Nonetheless, the concept of power seems apt for the description or experience of conducting gestures since physical movement can be easily described as being "powerful."

In the present study, participants were presented with point-light displays of conducting gestures, and asked to provide continuous ratings of perceived emotional expression on four scales: *valence*, *activity*, *power*, as well as overall amount of *expression*. In line with the findings of Davidson (2007) and Thompson and Luck (2010), it was predicted that ratings of perceived expression would be positively related to gesture amplitude. It was also expected that the different scales would tease out different combinations of movement features, revealing more specific relationships between movement features and ratings of expression. One might suppose, for example, that activity would be positively related to at least movement amplitude and speed, and that power would be positively related to movement amplitude and perhaps jerk. Precise relationships between the different scales and movement features were, however, hard to predict given the exploratory nature of this study. Nonetheless, it was further predicted that ratings of expression would be more strongly associated with movement features of the left hand given the traditional role of the left and right hands in conducting. In other words, while the right hand is primarily reserved for time-keeping duties, it is the left hand that is traditionally used to convey expression (Rudolf, 1980). Thus, we would expect the left hand to be more important to observers when rating the expressiveness of conducting gestures.

Method

Participants

Twenty-four individuals (mean age = 24.6 years, *SD* = 2.7, females = 16) provided continuous ratings of perceived expression. All participants were students attending the University of Jyväskylä, Finland. While a majority of the participants had received five or more years of music training (54%), only four participants had any formal conducting training. The participants were recruited using a university mailing list for students and were compensated for their time with a free cinema pass.

Gesture Recording and Stimulus Generation

The gestures of two early-career professional conductors (both enrolled on the conducting course at the Sibelius Academy, Finland) were recorded with an optical motion-capture system (Qualisys ProReflex) while they directed an ensemble of approximately 40 instrumentalists and singers. The piece being rehearsed was Mozart's Requiem Mass in D minor (K. 626). The three dimensional position of eight reflective markers attached to the hands (base of middle finger), wrists, elbows, and shoulders of each conductor was tracked at 60 Hz. Six excerpts (totaling 10 minutes) covering a range of expression and movement types (i.e., varying in terms of movement amplitude, speed, smoothness, etc.) were selected as stimuli, and the corresponding movement data then transformed into six QuickTime movies. To create the QuickTime movies, the movement data were imported into MATLAB, where animation frames of the point-light displays were created using the animation-related functions included in the MoCap Toolbox (Toiviainen, 2008). Each animation cell was saved as a JPEG file and then opened in QuickTime 7 as an image sequence and saved as a self-contained QuickTime movie. The animations had a frame rate of 30 Hz and segments were created to connect relevant markers to each other.

Procedure

A Max/MSP patch was created to play back the QuickTime movies, provide the slider interface that participants used to give their ratings, and save the ratings to file. A screen shot of the patch is shown in Figure 1. The six stimuli were presented in four identical blocks, one each for ratings of perceived activity, valence, strength, and overall expression. Within-block presentation order was randomized each time. Further, the order in which the blocks were presented was unique for each of the 24 participants. This was possible because four blocks could be ordered in a maximum of 24 unique permutations. Each participant initially practiced using the interface by providing continuous ratings for stimuli similar to the experimental stimuli. Once the participant expressed confidence that he/she understood the task, the experimenter left the room. Each block was prefaced by instructions on how each perceptual dimension should be rated and that the participant should use the full range of the slider when making perceptual judgements. The slider was controlled by clicking and dragging with the mouse. The instructions for each block were as follows:

> *Expression*. Move the slider to the right when you feel the gestures are more expressive and move it to the left when you feel they are less expressive.
>
> *Valence*. Move the slider to the right when you feel the gestures are more pleasant and move it to the left when you feel they are less pleasant.

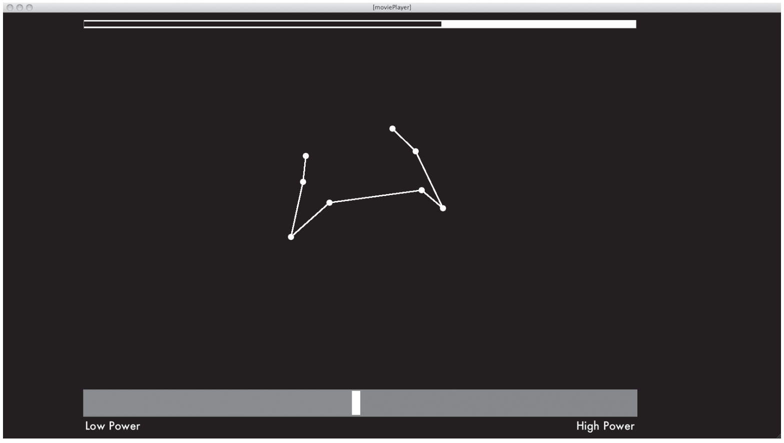


FIGURE 1. A screenshot of the Max/MSP patch displaying the stimuli. The user watches the point-light display while dragging the vertical white bar across the horizontal grey bar as their perception of expression, power, valance, and activity changes. A yellow status bar at the top of the screen indicates the temporal progression of the animation.

> *Activity*. Move the slider to the right when you feel the gestures have a high level of activity and move it to the left when you feel they have a low level of activity.
>
> *Power*. Move the slider to the right when you feel the gestures are more powerful and move it to the left when you feel they are less powerful.

Within a block, the QuickTime movies were played back automatically with an interval of 6 s between each. The positions on the slider ranged from –100 to +100. The Max/MSP patch recorded the position of the slider at a frequency of 4 Hz, allowing a high level of accuracy and detection of abrupt changes in the slider's position. Once the participant had been exposed to an entire block, they were given the option to take a break before starting the following block. Data collection lasted approximately 45 minutes per participant.

Movement Feature Extraction

In order to examine relationships between the perceptual ratings and the conductors' movements, 11 hand-related movement features were extracted from the raw movement data.

The first four movement features were simply the position trajectories relating to the horizontal (x-axis) and vertical (y-axis) movement of the conductors' hands. Movement depth (z-axis) was not considered for analysis as the participants were only exposed to animations displaying horizontal and vertical movements.

The fifth feature was a single vector representing the frame-by-frame distance between the hand markers. This feature was selected for analysis as one might assume that the distance between the hands was a distinguishing feature of large and small movements and would have influenced participants' perceptual ratings.

Features six to eleven were related to the kinematics of the hand movements. The absolute values of velocity, acceleration, and jerk were estimated along the horizontal and vertical dimensions of the movement data by applying numerical differentiation to the position data. The time derivative at each frame was estimated using a Savitzky-Golay FIR smoothing filter, which fitted a second order polynomial to seven consecutive frames centred at the frame in question and computed the derivative of the thus obtained polynomial at the centre frame (Luck & Toiviainen, 2006). From these column vectors, the instantaneous velocity, acceleration, and jerk were derived as the length of the vector consisting of the components of velocity, acceleration, and jerk respectively.

The final step was to make all of the movement variable column vectors the same length as the slider position vectors. This was achieved by taking the average of 30 consecutive frames occurring prior to each reading of the slider, thus down-sampling the movement data from 60 Hz to 4 Hz. This averaging method allowed for direct examination between the slider position ratings and the average of the movement data occurring 250 ms prior to the rating.

To sum up, the eleven movement features from each conducting performance were the horizontal and vertical positional data of each hand (4), the distance vector between each hand (1), and the instantaneous velocity, acceleration and jerk of each hand marker (6).

Results

Calculating Mean Ratings

Mean ratings (in the form of mean time series) were first calculated for each excerpt. Each participant's individual rating was then correlated with each of these means, and participants with either nonsignificant or negative coefficients were regarded as outliers and removed from further analyses. Mean ratings were then recalculated from remaining participants' ratings. For the six different stimuli, an average of 4.6 participants were removed from the initial expression ratings, 7.2 from the valence ratings, 3 participants from the activity ratings, and 5 participants from the power ratings.

Next, in order to see how ratings of the four dimensions related to each other, interdimension correlations, based upon the mean rating for each of the four dimensions, were calculated. The correlation matrix for the averaged data is shown in Table 1 and reveals that ratings for the four dimensions were generally moderately correlated with each other, except for activity and power which were quite highly correlated. The lowest correlation was between expression and valence, and the highest correlation between activity and power.

TABLE 1. Correlation Matrix Showing the Interdimension Correlations Between Expression, Valence, Activity, and Power.

	Expression	Valence	Activity	Power
Expression	1.00	.37	.45	.49
Valence	—	1.00	.55	.43
Activity	—	—	1.00	.79
Power	—	—	—	1.00

Note: All coefficients were significant at $p < .01$, and N for all time-series was 2222.

Lag Analysis

Mean ratings for each excerpt were then crosscorrelated with the corresponding movement variables. In each case, the maximal cross-correlation within the range −20 samples to +20 samples (−5 s to +5 s) indicated the lag between the musical feature and participants' response to it. Variables were subsequently time shifted so that the peak correlation occurred at lag = zero. This was done separately for each of the 24 mean ratings (six excerpts × four dimensions), and it was these final time-shifted datasets that were used in subsequent regression analyses. Table 2 shows the mean lag for each of the 11 movement features for each rating scale.

Regression Analysis

To model the experimental data, we employed ordinary least squares linear regression, in which the movement features were entered as predictors of participants' emotion ratings. Excerpts 1–3 were produced by Conductor 1, while excerpts 4–6 were produced by Conductor 2. Preliminary inspection of the data revealed some differences in terms of relationships between ratings of perceived emotion and their movement features such that

TABLE 2. Mean Lags (s) for Each of the 11 Movement Features and the Four Rating Scales. Positive Values Indicate the Amount of Time Elapsed Between a Movement Feature Occurring and Participants Responding to It.

Movement feature	Expression	Valence	Activity	Power	Overall
Distance between hands	3.21	1.96	2.33	2.54	2.51
Right hand x position	0.58	0.79	1.50	2.08	1.24
Right hand y position	1.38	1.38	1.33	1.46	1.39
Right hand speed	1.88	2.04	1.33	3.25	2.13
Right hand acceleration	2.42	2.25	3.04	2.33	2.51
Right hand jerk	2.29	1.50	1.79	1.67	1.82
Left hand x position	2.79	1.71	1.58	1.46	1.89
Left hand y position	2.29	1.54	1.42	1.42	1.67
Left hand speed	2.79	1.63	1.71	1.33	1.87
Left hand acceleration	2.71	1.33	1.54	1.21	1.70
Left hand jerk	2.67	1.58	1.83	1.92	2.00

TABLE 3. Summary of the Eight Regression Models.

Dimension	Conductor	*F* ratio	*df*	R^2	R^2_{adj}
EXPRESSION	Both	260.71	11, 2231	.562	.560
	1	105.57	11, 1161	.500	.495
	2	160.61	11, 1058	.625	.622
VALENCE	Both	314.18	11, 2244	.606	.604
	1	173.79	11, 1169	.621	.617
	2	226.53	11, 1063	.701	.698
ACTIVITY	Both	282.38	11, 2253	.580	.578
	1	174.62	11, 1173	.621	.617
	2	284.24	11, 1068	.745	.743
POWER	Both	265.42	11, 2222	.568	.566
	1	118.89	11, 1154	.531	.527
	2	272.42	11, 1056	.739	.737

Note: All *F* ratios were significant at $p < .001$.

TABLE 5. Regression Models for VALENCE.

	Beta values		
Movement feature	Both	Conductor 1	Conductor 2
Distance between hands	−.007†	.023†	.050†
Right hand x position	.040†	.085***	−.029†
Right hand y position	.134***	.165***	.063**
Right hand speed	.803***	.577***	.584***
Right hand acceleration	−1.407***	−.115†	−1.138***
Right hand jerk	.987***	.149**	.855***
Left hand x position	−.026†	−.071**	.091***
Left hand y position	.008†	.259***	−.029†
Left hand speed	1.040***	.422***	1.108***
Left hand acceleration	−1.745***	−.253**	−1.683***
Left hand jerk	1.103***	.127**	1.028***

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean valence rating; * < .05, ** < .01, *** < .001, †not significant.

analyzing all six excerpts as a whole could obscure true effects. Thus, in addition to analyzing all excerpts together, excerpts 1–3 (Conductor 1) and excerpts 4–6 (Conductor 2) were pooled and analyzed separately. In order to examine larger structures and overall themes in terms of relationships between movement features and ratings of expression, all variables were smoothed using a running mean with a 10-point-span (equivalent to 2.5 s) prior to analysis.

Twelve separate linear regression analyses were carried out, one for each of the four emotion dimensions on: (a) the whole dataset, (b) pooled data for excerpts 1–3, and (c) pooled data for excerpts 4–6. In each analysis, the 11 movement variables were entered simultaneously. Significant models emerged for all twelve analyses, and are summarized in Table 3. Beta values and significance levels for the movement variables are shown in Tables 4–7.

Expression. It can be seen from Table 4 that, when both conductors were analyzed together, ratings of perceived expression were positively related to right-hand y position (greater perceived expression when the right hand was held higher), and acceleration of both hands, and negatively related to left-hand x position (greater perceived

TABLE 4. Regression Models for EXPRESSION.

	Beta value		
Movement feature	Both	Conductor 1	Conductor 2
Distance between hands	.006†	.082***	−.094**
Right hand x position	.024†	.094***	.013†
Right hand y position	.109***	.027†	.065**
Right hand speed	.002†	.607***	−.057†
Right hand acceleration	.664***	.326**	.883***
Right hand jerk	.052	−.293***	.032†
Left hand x position	−.142***	−.172***	−.085***
Left hand y position	.024†	.145***	.143***
Left hand speed	−.001†	−.075†	−.024†
Left hand acceleration	.696***	.176**	.900***
Left hand jerk	−.695***	−.057†	−.894***

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean expression rating; * < .05, ** < .01, *** < .001, †not significant.

TABLE 6. Regression Models for ACTIVITY.

	Beta values		
Movement feature	Both	Conductor 1	Conductor 2
Distance between hands	.082***	.104***	.129***
Right hand x position	.020†	.056**	−.007†
Right hand y position	.108***	.107***	.015†
Right hand speed	.249***	.735***	−.312***
Right hand acceleration	.032†	−.031†	1.247***
Right hand jerk	.291***	.031†	−.353***
Left hand x position	−.083***	−.032†	−.034†
Left hand y position	−.011†	.276***	.013†
Left hand speed	.378***	.049†	.319***
Left hand acceleration	−.307***	.002†	−.124†
Left hand jerk	.127*	.092*	−.015†

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean activity rating; * < .05, ** < .01, *** < .001, †not significant.

TABLE 7. Regression Models for POWER.

	Beta values		
Movement feature	Both	Conductor 1	Conductor 2
Distance between hands	.036†	–.011†	.149***
Right hand x position	–.019†	.058**	–.072***
Right hand y position	.168***	.185***	.107***
Right hand speed	.927***	.617***	.995***
Right hand acceleration	–1.554***	–.126*	–.1487***
Right hand jerk	.937***	.041†	.922***
Left hand x position	–.061***	–.027†	–.038†
Left hand y position	–.040*	.119***	.028†
Left hand speed	.736***	.162***	.652***
Left hand acceleration	–.949***	–.046†	–.939
Left hand jerk	.614***	.174***	.665***

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean power rating; * < .05, ** < .01, *** < .001, †not significant.

expression when the left hand was held closer to the body) and jerk.[1] When Conductor 1 was analyzed separately, ratings of perceived expression were positively related to distance between the hands, right-hand x position, speed, and acceleration, and left-hand y position and acceleration, and negatively related to right-hand jerk, and left-hand x position. Regarding Conductor 2, ratings of perceived expression were positively related to right-hand y position and acceleration, and left-hand y position and acceleration, and negatively related to distance between the hands, and left-hand x position and jerk.

Valence. Table 5 reveals that, when both conductors were analyzed together, ratings of perceived valence were positively related to right-hand y position, speed, and jerk, as well as left-hand speed and jerk, and negatively related to both right- and left-hand acceleration. For Conductor 1, meanwhile, perceived valence was positively related to right-hand x position, y position, speed, and jerk, and left-hand y position, speed, and jerk, and negatively related to left-hand x position and acceleration. For Conductor 2, perceived valence was positively related to right-hand y position, speed, and jerk, and left-hand x position, speed, and jerk, and negatively related to both right and left-hand acceleration.

Activity. Table 6 shows that, when both conductors were analyzed together, ratings of perceived valence were positively related to distance between the hands, right-hand y position, speed, and jerk, and left-hand speed and jerk, and negatively related to left-hand x position and acceleration. For Conductor 1, ratings of perceived activity were positively related to distance between the hands, right-hand x position, y position, and speed, and left-hand y position and jerk. Regarding Conductor 2, perceived activity was positively related to distance between the hands, right-hand acceleration, and left-hand speed, and negatively related to right-hand speed and jerk.

Power. Table 7 shows that, when both conductors were analyzed together, ratings of perceived power were positively related to right-hand y position, as well as speed and jerk of both hands, and negatively related to left-hand x position, and acceleration of both hands. For Conductor 1, meanwhile, perceived power was positively related to right-hand x position, y position, and speed, and left-hand y position, speed, and jerk, and negatively related to right-hand acceleration. For Conductor 2, perceived power was positively related to distance between the hands, right-hand y position, speed, and jerk, and left-hand speed and jerk, and negatively related to right-hand x position and acceleration.

To help visualize the success of the 12 models in predicting participants' ratings, the predicted values of expression, valence, activity, and power resulting from these models were plotted against the actual mean rating for each of these dimensions, for each of the six excerpts. These plots are shown in Figure 2. It can be seen that the predicted ratings generally correspond quite closely to the actual mean ratings. The fit between predicted and actual ratings, however, varies from excerpt to excerpt, and, since the different models explain different amounts of variance, between both the conductors and the four dimensions.

Finally, potential issues of multicollinearity were investigated in order to check the accuracy and stability of the models. An examination of two indices of multicollinearity revealed no serious concerns related to this phenomena. More specifically, mean variance inflation factors (VIFs)—which indicate whether a predictor has a strong linear relationship with the other predictors—for the expression, valence, activity, and power models (by conductor) were 6.5/11.2, 7.2/27.1, 6.2/13.3, and 3.4/34.9 respectively. These figures suggest that small levels of multicollinearity may be present (see Bowerman & O'Connell, 1990). However, tolerances for most variables were at least .1, and in the majority of cases were between .4 and .8. Tolerance is the percentage of

---

[1] The coordinate system was such that larger values on the y axis indicated higher vertical position, while larger values on the x axis indicated a greater degree of movement towards the left of the conductors' body. Consequently, a positive relationship between each dimension's rating and y and x position indicated that, when the rating had a higher value, position was high and shifted towards the left (e.g., high up, and further away from the body for the left hand but closer to the body for the right hand), and vice versa for a negative relationship (e.g., low down, and closer to the body for the left hand but further away from the body for the right hand).

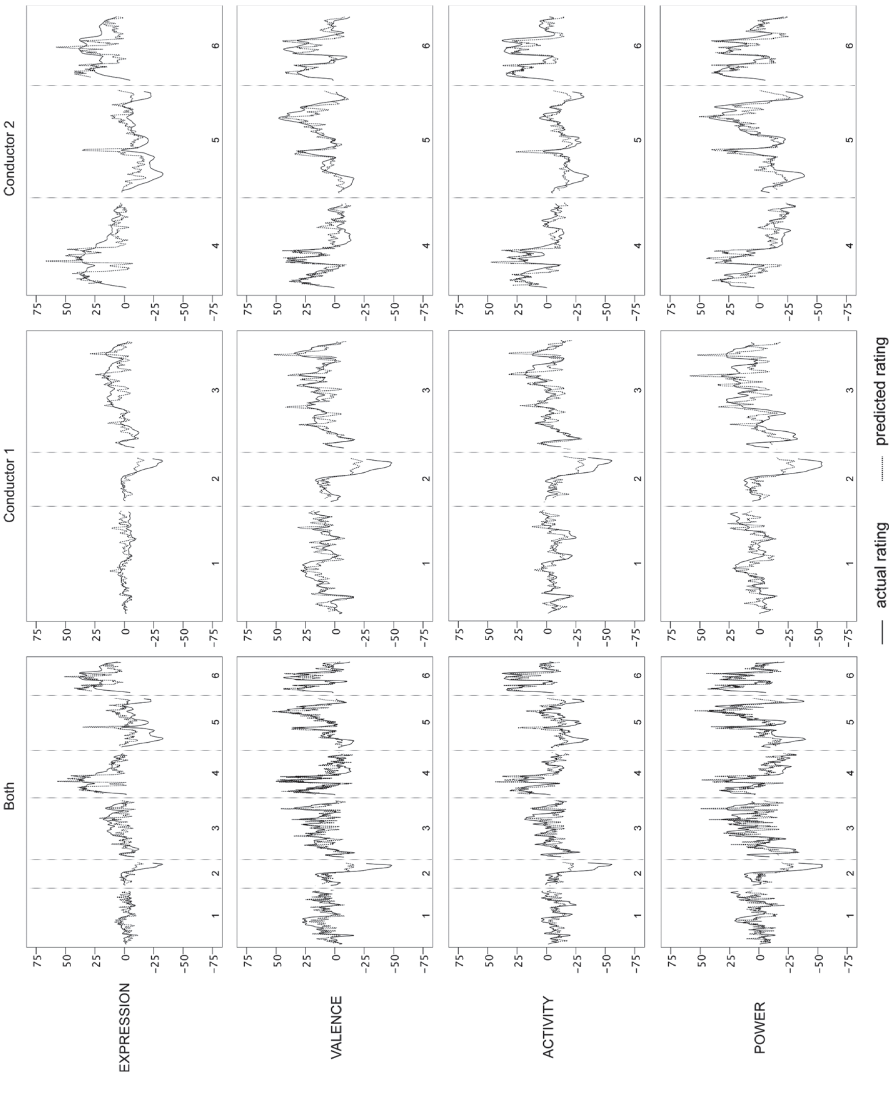


FIGURE 2. Actual and predicted ratings (vertical axes) of each of the four emotion dimensions plotted again time (horizontal axes) for all six excerpts.

the variance in a given predictor that cannot be explained by the other predictors (values should be multiplied by 100 to obtain the percentage). Only values below .1 indicate serious problems (see Field, 2005), and the fact that most of the variables had high tolerances indicated no problems of multicollinearity. In sum, then, we were confident that multicollinearity was not unduly affecting the accuracy or stability of the regression models.

Discussion

Overall, the regression models accounted for between 50 and 75 percent of the variance in participants' ratings of the four dimensions (mean variance = 56% for expression, 64% for valence, 65% for activity, and 61% for power). When both conductors were analyzed together, gestures rated as being more expressive overall were characterized by vertically extended and accelerated right-hand movement, and smooth but accelerated, and less horizontally extended, left-hand movement. Gestures receiving higher ratings of valence were characterized by vertically extended right-hand movement, and less accelerated but fast and jerky movement of both hands. Gestures receiving higher ratings of activity were characterized by vertically extended right-hand movement, greater distance between the hands, fast and jerky movement of both hands, and less horizontally extended and accelerated left-hand movement. Finally, gestures rated as being more powerful were characterized by vertically extended right-hand movement, less accelerated but fast and jerky movement of both hands, and both horizontally and vertically less extended left-hand position. These descriptions are summarized in Table 8.

A few points regarding these descriptions are worth noting. First, the similarity between the movement descriptions for activity and power, and the dissimilarity between expression and valence, mirror the highest and lowest inter-dimension mean rating correlations shown in Figure 1 above. Second, while the right hand is vertically extended for all dimensions (likely due to it being consistently held high to convey primarily temporal information using traditional time-beating gestures), the beta value for this feature was highest for power, indicating that this feature was most strongly associated with conveying power. Third, speed, acceleration, and jerk appear to be important for all four measured dimensions, the direction of the relationships being identical for all three components of Osgood et al.'s (1957) dimensional theory (valence, activity, and power). Fourth, and related to the previous point, hand position seems to be the feature that differentiates between Osgood et al.'s dimensions, activity being represented in the same way as valence but with hands held further apart, and power further specifying the right hand be held highest, and the left hand close to the body. Finally, these results provide partial support for the hypothesis that higher ratings of expression will be related to greater movement amplitude insasmuch as higher ratings of activity were related to the hands being held further apart, and higher ratings of power were related to the right hand being held the highest.

TABLE 8. The Primary Movement Features Associated with Mean Ratings of Expression, Valence, Activity, and Power When Both Conductors Were Analyzed Together.

Dimension	Movement features
Expression	Accelerated but smooth
Valence	Fast, jerky, but less accelerated
Activity	Hands further apart, fast and jerky, less accelerated
Power	Fast and jerky, less accelerated, right hand held highest, left hand close to body

One problem with generalizing the results this way is that it can obscure stylistic differences. That is, if we look for commonalities between the conductors, we may miss more important relationships between the movement features and ratings of expression that are conductor-specific. An alternative approach, therefore, is to look at each conductor individually.

Regarding perceived expression, the main differences in terms of significant movement features between the conductors were the direction of the relationship with distance between the hands (positive for Conductor 1, negative for Conductor 2), horizontal (Conductor 1) versus vertical (Conductor 2) extension of the right hand, higher speed and smoother movement of the right hand for Conductor 1, and smoother left-hand movement for Conductor 2. For perceived valence, the main differences were horizontally extended right-hand for Conductor 1, significantly less accelerated right-hand movement for Conductor 2, the direction of the relationship with horizontal extension of the left hand (negative for Conductor 1, positive for Conductor 2), and the significance of vertically extended left-hand position. Regarding perceived activity, the main differences were more horizontally and vertically extended right hand position for Conductor 1, the direction of the relationship with right-hand speed (positive for Conductor 1, negative for Conductor 2), more accelerated but smoother right-hand movement for Conductor 2, vertically extended left-hand position for Conductor 1, and left-hand speed for Conductor 2. Finally, for perceived power, the main differences were distance between the hands for Conductor 2, the direction of the relationship with horizontal extension of the right-hand (positive for Conductor 1, negative for Conductor 2), smoother movement for Conductor 2, and vertically extended left-hand position for Conductor 1.

It is interesting that most of these differences relate to hand position, with fewer differences between the conductors in terms of speed, acceleration, and jerk of movement. It seems, therefore, that differences in hand position help differentiate not only between different expressive dimensions, but also between different conductors. It also seems that the relatively stable relationships between speed, acceleration, and jerk, and the expression dimensions seen when both conductors were analyzed together break down when the conductors are analyzed separately.

In addition, it is worth noting that, in several cases, the direction of the relationship between movement features and the expressive dimensions are at odds with each other (such as distance between the hands being positively related to the expression dimension for Conductor 1, but negatively related for Conductor 2). These differences have the effect of canceling each other out when both conductors are analyzed together, and thus limits the extent to which the results of the overall analysis (both conductors together) can be generalized.

Analyzing the results by conductor, then, helps to reveal stylistic differences in conducting technique. Moreover, Table 3 reveals that the regression models for Conductor 2 consistently "outperformed" those for Conductor 1: mean variance explained for Conductor 1 was 57%, as compared to 71% for Conductor 2. This implies that relationships between movement features and perceived expression were clearer for Conductor 2 than Conductor 1, and suggests that Conductor 2 was able to express emotion in a more consistent manner.

Both stylistic differences between conductors, as well as more general relationships between movement features and perceived expression, might be clarified in future work by replicating this study with more conductors. The investigation of two conductors in the present study reduces the extent to which our results can be generalized, particularly in light of the idiosyncrasies they demonstrated.

Another interesting approach would be to have the conductors rate each other (and perhaps themselves) in terms of perceived expression. This might be further extended by having other skilled conductors rate these conductors. Since musicians with previous conducting experience have been shown to synchronize more consistently with temporal conducting gestures compared to musicians with synchronization experience only or nonmusicians (Luck & Nte, 2008), one might speculate that a similar effect of domain-specific experience would be evident in ratings of perceived expression. That is, skilled conductors might have a deeper understanding of how expression is conveyed by movement, and thus be able to make better use of the kinematic information available.

This also raises the question of potential connections between expressivity and timing. Might there be a connection between expressivity and timing, for example, such that observers can synchronize better with more expressive conductors? Or vice versa? There is also the issue of how the expressiveness of conductors' gestures affects timing within an ensemble, not just between the ensemble and the conductor. One might suppose that gestures containing smaller peaks in acceleration (in the present study, those characterized by, for example, high levels of valence) would be harder to synchronize with accurately compared to gestures containing more marked peaks in acceleration (in the present study, gestures characterized by higher levels of overall expression).

Furthermore, there is the issue of the use of other body parts in expressive conducting. A number of studies have shown, for example, that the conductor's face conveys important expression-related information (e.g., Dan, 2005; Mayne, 1992; Wöllner, 2008; Wöllner & Auhagen, 2008). Wöllner and Auhagen argue that, while a reduction of conducting gestures to the movement of a limited number of points in space (such as the tip of the baton, or the hands, for instance) allows a controlled experimental setting, validity could be increased if other aspects, such as facial expression, are included in stimuli. We agree with this assertion, but chose in the present study to focus on the more general role of gross body movements in conveying expression.

In sum, this study provides preliminary data on relationships between characteristics of conductors' gestures and perceived expression. We found that observers (including nonmusicians) were sensitive not only to movement amplitude but also more fine-grained kinematic features when rating perceived expression. In conclusion, we found that higher levels of expressivity can be conveyed by gestures characterized by increased amplitude (greater distance between the hands, lower and more horizontally extended left-hand position, higher right-hand position), greater variance (jerkier movement of both hands), and higher speed of movement (faster movement of both hands).

Author Note

Correspondence concerning this article should be addressed to Geoff Luck, Academy of Finland Research Fellow Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, P.O. Box 35(M) 40014, University of Jyväskylä, Finland. E-MAIL: geoff.luck@jyu.fi

References

BOUGHTON, M., & STEVENS, C. (2008). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, *37*, 137–153.

BOWERMAN, B. L., & O'CONNELL, R. T. (1990). *Linear statistical models: An applied approach.* Boston, MA: PWS Publishing Co.

CAMURRI, A., & VOLPE, G. (Eds.). (2004). *Gesture-based communication in human-computer interaction*. Berlin: Springer.

DAHL, S., & FRIBERG, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, *24*, 433–454.

DAN, K. (2005). Facial expression and eye contact used by instrumental conductors: Practical applications and exercises. *Dissertation Abstracts International*, *66*(02), 399 (UMI No. 3166080).

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

DAVIDSON, J. W. (1995). What does the visual information contained in music performances offer the observer? Some preliminary thoughts. In R. Steinberg (Ed.), *Music and the mind machine: The psychopathology of the sense of music XXX* (pp. 105–113). Berlin: Springer-Verlag.

DAVIDSON, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music, 35,* 381–401.

DITTRICH, W. H., TROSCIANKO, T., LEA, S. E. G., & MORGAN, D. (1996). Perception of emotion from dynamic light-point displays represented in dance. *Perception*, *25*, 727–738.

FIELD, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

JUSLIN, P. N., & SLOBODA, J. A. (2001). *Music and emotion: Theory and research.* London: Oxford University Press.

KENDALL, R. A., & CARTERETTE, E. C. (1990). The communication of musical expression. *Music Perception*, *8*, 129–164.

KRUMHANSL, C. L., & SCHENK, D. L. (1997). Can dance reflect the structural and expressive qualities of music? *Musicae Scientiae*, *1*, 63–83.

LUCK, G. (2008, August). Spatio-temporal cues for synchronization with conductors' gestures. *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10)*, Sapporo, Japan.

LUCK, G., & NTE, S. (2008). A new approach to the investigation of conductors' gestures and conductor-musician synchronization, and a first experiment. *Psychology of Music*, *36*, 81–99.

LUCK, G., RIIKKILÄ, K., LARTILLOT, O., ERKKILÄ, J., TOIVIAINEN, P., MÄKELÄ, A., ET AL. (2006). Exploring relationships between level of mental retardation and features of music therapy improvisations: A computational approach. *Nordic Journal of Music Therapy*, *15*, 30–48.

LUCK, G., & SLOBODA, J. A. (2008). Exploring the spatio-temporal properties of the beat in simple conducting gestures using a synchronization task. *Music Perception*, *25*, 225–239.

LUCK, G., & SLOBODA, J. A. (2009). Spatio-temporal cues for visually-mediated synchronization. *Music Perception*, *26*, 465–473.

LUCK, G., & TOIVIAINEN, P. (2006). Ensemble musicians' synchronization with conductors' gestures: an automated feature-extraction analysis. *Music Perception*, *24*, 189–200.

LUCK, G., TOIVIAINEN, P., ERKKILÄ, J., LARTILLOT, O., RIIKKILÄ, K., MÄKELÄ, A., ET AL. (2008). Modelling the relationships between emotional responses to, and musical content of, music therapy improvisations. *Psychology of Music*, *36*, 25–46.

MAYNE, R. G. (1992). An investigation of the use of facial expression in conjunction with musical conducting gestures and their interpretation by instrumental performers. *Dissertation Abstracts International*, *53(08),* 2729 (UMI No. 9238229)

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. (1957) *The measurement of meaning*. Champaign, IL: University of Illinois Press.

RUDOLF, M. (1980). *The grammar of conducting.* New York: Schirmer Books.

SCHUBERT, E. (2001). Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.) *Music and emotion: Theory and research* (pp. 393–414). Oxford, UK: Oxford University Press.

THOMPSON, M. R., & LUCK, G. (2010). *Exploring relationships between pianists' body movement, their expressive intentions, and structural elements of the music.* Manuscript submitted for publication.

TOIVIAINEN, P. (2008). *MoCap toolbox manual*. University of Jyväskylä: Jyväskylä, Finland. Available at http://www.jyu.fi/music/coe/materials/ mocaptoolbox/MCTmanual.

WÖLLNER, C. (2008).Which part of the conductor's body conveys most expressive information? A spatial occlusion approach. *Musicae Scientiae*, *12*, 249–272.

WÖLLNER, C., & AUHAGEN, W. (2008). Perceiving conductors' expressive gestures from different visual perspectives. An exploratory continuous response study. *Music Perception*, *26*, 129–143.

III

EMBODIED METER: HIERARCHICAL EIGENMODES IN MUSIC-INDUCED MOVEMENT

by

Petri Toiviainen, Geoff Luck & Marc R. Thompson

Music Perception vol. 28(1), 59-70.

EMBODIED METER: HIERARCHICAL EIGENMODES IN MUSIC-INDUCED MOVEMENT

PETRI TOIVIAINEN, GEOFF LUCK,
& MARC R. THOMPSON
Department of Music, University of Jyväskylä,
Jyväskylä, Finland

LISTENING TO MUSIC OFTEN IS ASSOCIATED WITH spontaneous body movements frequently synchronized with its periodic structure. The notion of embodied cognition assumes that intelligent behavior does not emerge from mere passive perception, but requires goal-directed interactions between the organism and its environment. According to this view, one could postulate that we may use our bodily movements to help parse the metric structure of music. The aim of this study was to investigate how pulsations on different metrical levels manifest in music-induced movement. Musicians were presented with a piece of instrumental music in 4/4 time, played at four different tempi ranging from 92 to 138 bpm. Participants were instructed to move to the music, and their movements were recorded with a high quality optical motion capture system. Subsequently, signal processing methods and principal components analysis were applied to extract movement primitives synchronized with different metrical levels. We found differences between metric levels in terms of the prevalence of synchronized eigenmovements. For instance, mediolateral movements of arms were found to be frequently synchronized with the tactus level pulse, while rotation and lateral flexion of the upper torso were commonly found to exhibit periods of two and four beats, respectively. The results imply that periodicities on several metric levels are simultaneously present in music-induced movement. This could suggest that the metric structure of music is encoded in these movements.

Received October 30, 2009, accepted April 7, 2010.

Key words: music, movement, synchronization, meter, embodiment

THERE IS A STRONG CONNECTION BETWEEN music and bodily movement. When we listen to a piece of music, we often tap our foot, nod our head, or move our body along with it. In most cultures, music and dance have evolved together (Arom, 1991; Cross, 2003; Wallin, Merker, & Brown, 2000). Most people report that they move with music (Lesaffre et al., 2008). When listening to music, people tend to walk faster (i.e., take longer strides) than with metronomic stimuli (Styns, van Noorden, Moelants, & Leman, 2007).

The notion of embodied cognition (Leman, 2008) assumes that intelligent behavior does not emerge from mere passive perception, but requires goal-directed interactions between an organism and its environment. In the field of linguistics, the motor theory of speech perception (Liberman & Mattingly, 1985) draws from this notion, assuming that speech perception is based on the cognitive system's ability to compute the articulatory gestures that could have produced the received acoustical signal. The notion of embodied music cognition has its roots in the early 20th century and has recently been elaborated by Leman (2008). Leman emphasizes the role of goal-directed actions in music perception. He sees synchronization to music as a form of corporeal imitation, and postulates: "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" (Leman, 2008).

There is evidence that movement is more strongly connected to the auditory system than to the visual system. For instance, tapping to a visual rhythm is more efficiently disrupted with an auditory distractor than vice versa (Repp & Penel, 2004). Moreover, Patel, Iversen, Chen, and Repp (2005) found that in a tapping task participants failed to synchronize to metrical nonisochronous visual stimuli, while they had no difficulty in synchronizing to similar auditory stimuli.

Music often contains recurring temporal patterns that give rise to a percept of beat. The beat refers to the subjective sense of periodicity in music evoked by temporal regularities in the acoustical signal. The ability to synchronize with musical beat has to date been mostly investigated with finger tapping studies. According to these studies, synchronization ability is spontaneous and accurate (Drake, Penel, & Bigand, 2000; Large, Fink, & Kelso, 2002; Snyder & Krumhansl, 2001; Toiviainen &

Music Perception VOLUME 28, ISSUE 1, PP. 59–70, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2010.28.1.59

Snyder, 2003) within a relatively wide range of periods (300–900 ms), and is associated with a preferred pulse period near 500 ms (Fraisse, 1982; Parncutt, 1994; van Noorden & Moelants, 1999).

When listening to music, we frequently perceive more than one beat level. These beats are often hierarchically organized, with their periods having integer ratios (Palmer & Krumhansl, 1990). The interaction between these different pulse sensations results in a percept of periodically alternating strong and weak beats, corresponding to the generally accepted definition of meter (Lerdahl & Jackendoff, 1983). Metrical structure is normally derived from the accent structure of music, which again emerges from various sources such as the duration, loudness, and pitch of tones, and harmonic changes. Most music has either a duple (every second beat accented) or a triple meter (every third beat accented).

While music certainly induces movement, there also is some evidence to suggest that movement affects beat perception. Todd, Cousins, and Lee (2007) found that 16% of variation in preferred beat rate can be predicted from anthropometric factors, such as weight as well as length and width of certain body segments. Phillips-Silver and Trainor found that both infants' (2005) and adults' (2007) encoding of meter can be affected by movement. Todd, O'Boyle, and Lee (1999) proposed that pulse is an inherently sensorimotor phenomenon in the sense that pulse perception necessarily involves motor system activity. Phillips-Silver and Trainor (2008) investigated whether passive movement of the head and legs affects the perception of rhythm, and found that head movement biases meter perception while movement of legs has no effect. Moreover, Trainor, Gao, Lei, Lehtovaara, and Harris (2009) discovered that galvanic stimulation of the vestibular system can be used to disambiguate an ambiguous metric pattern. These results suggest that the vestibular system has a central role in the perception of rhythm.

While there exists a few studies on synchronization with a musical beat (Drake et al., 2000; Large et al., 2002; Repp, 2005a, 2005b; Snyder & Krumhansl, 2001; Toiviainen & Snyder, 2003), the kinematic aspects of this activity have been investigated to a lesser extent. However, Eerola, Luck, and Toiviainen (2006) investigated toddlers' corporeal synchronization with music using a high resolution motion capture system. They found that 2–4 year old children exhibited periodic movement, and that this movement was at times synchronized with the tactus of music. The tactus refers to the basic pulse of music, i.e., the pulse to which listeners typically entrain when tapping. The periodic movement was most clearly visible in the vertical dimension. The embodiment of higher metrical levels in musical movement has received less attention. Naveda and Leman (2010), however, studied how the metric hierarchy of Samba and Charleston is reflected in professional dancers' repetitive gestures. To this end, they propose a method based on principal components analysis and periodicity transform (Sethares & Staley, 2001). With respect to musical movement by ordinary music listeners, the presence of metrical levels has, to date, not been studied.

The present work investigates the nature of music-induced movements, focusing on how pulsations on different metrical levels manifest in these movements. We were interested in two main questions. First, whether music-induced movements display movement components that are simultaneously synchronized to different metrical levels. Second, whether the synchronized movement patterns differ between the metrical levels. To this end, we applied kinetic analysis, body modeling, dimensionality reduction, and signal processing to data acquired using a high resolution motion capture system to identify the most typical movement patterns, or eigenmovements, synchronized to different metrical levels.

Based on the theoretical considerations presented above, in particular the notion of embodied cognition, we expect to see simultaneous synchronization of movement patterns to several metrical levels. Moreover, due to inertial and biomechanical properties of the body, we expect to see the higher (i.e., slower) metrical levels to be embodied as gross body movements, in particular movement of the torso, and lower (faster) metrical levels embodied as movements of the extremities. This behavior can be expected based on the simple argument that the specific period of oscillation of a rigid physical body is proportional to the square root of its moment of inertia and the moment of inertia is proportional to the mass and the physical dimensions of the rigid body. For instance, the torso has larger physical dimensions and larger mass than the arm and therefore has a higher moment of inertia and consequently a longer specific period of oscillation.

Method

Participants

A total of 18 participants (12 females, 6 males, mean age = 24.2, *SD* of age = 3.9) took part in the study. These participants were predominantly Finnish university students (75%) with the remainder comprising international exchange students of various nationalities. The average number of years of formal music training was 8.0 (*SD* = 6.7); 89% of those with music training had trained for 10 years or more. Most participants stated that they enjoyed dancing at home or in a nightclub (56%), with an average time spent dancing of 2.0 hours per week (*SD* = 2.6 hours), while a minority (25%) of the participants had had formal dance or aerobics training. Additionally, participants were asked how many hours per week they spent doing physical activities or sports, of which the average time was 5.1 hours (*SD* = 4.1 hours). Participants received an honorarium in the form of a movie ticket for their participation in the study.

Apparatus

Participants' movements were recorded using an eight-camera optical motion capture system (Qualisys ProReflex). The system tracked the movement of reflective markers placed on each participant at a frame rate of 60 Hz. The trajectories were interpolated together to create a three dimensional point-light display of each participant. The motion capture data were synchronized with the musical stimulus using the synchronization pulses transmitted by the Qualisys cameras. The musical stimulus was played back from a Pure Data (Pd) patch running on an Apple computer. Both the audio and the synchronization pulses were routed to the inputs of a Digidesign Mbox and recorded using ProTools software. For reference purposes, the sessions were additionally videotaped using a Panasonic Mini DV camcorder.

Materials and Procedure

Participants were recorded individually, and, prior to each motion capture session, a total of 28 reflective markers were attached to the participant's body. The locations of the markers were as follows (L = left, R = right, F = front, B = back): 1: LF head (L frontal eminence); 2: RF head (R frontal eminence); 3: LB head (L dorsal parietal bone); 4: RB head (R dorsal parietal bone); 5: L shoulder (L scapular acromion); 6: R shoulder (R scapular acromion); 7: spine (midpoint between the superior angles of the scapulae); 8: breastbone (sternum); 9: LF hip (L anterior superior iliac spine); 10: RF hip (R anterior superior iliac spine); 11: LB hip (L posterior superior iliac spine); 12: RB hip (R posterior superior iliac spine); 13: L elbow (L olecranon); 14: R elbow (R olecranon); 15: L inner wrist(L distal radius); 16: L outer wrist (L distal ulna); 17: R inner wrist (R distal radius); 18: R outer wrist (R distal ulna); 19: L middle finger (metacarpophalangeal joint of L digitus medius); 20: R middle finger (metacarpophalangeal joint of R digitus medius); 21: L knee (L lateral distal femur); 22: R knee (R lateral distal femur); 23: L ankle (L lateral malleolus); 24: R ankle (R lateral malleolus); 25: L little toe (L 5th proximal phalanx); 26: L big toe (distal interphalangeal joint of L hallux); 27: R little toe (R 5th proximal phalanx); 28: R big toe (distal interphalangeal joint of R hallux). The locations of the markers are depicted in Figure 1a.

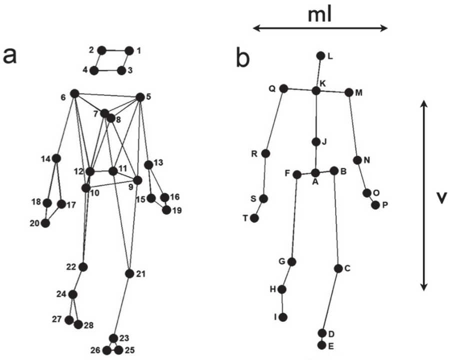


FIGURE 1. (a) Anterior view of the location of markers attached to participants' bodies. (b) Anterior view of the locations of the secondary markers used in the analysis (see text). Neighboring joints in kinematic chains are connected with lines. The directions of the mediolateral (ml) and vertical (v) axes are indicated by arrows. The anteroposterior axis is vertical to the surface of the page.

Once the markers were attached and well placed, we ensured that they were nonintrusive to the participant's mobility, and that the participant was comfortable with the apparatus and the setting. The participant stood in the centre of the capture area and was instructed to 'move freely' to the music.

Participants were presented with an instrumental 12-bar blues progression performed in a minor key and 4/4 meter. During the piece, the progression was repeated at four different tempi chosen randomly from the tempi 92, 103, 115, 126, and 138 beats per minute (bpm), and presented in random order. The total amount of motion capture data collected and subjected to analysis was ca. 37.5 min.

Results

A schematic representation of the entire analysis procedure is provided in Figure 2.

In what follows, the steps of the analysis process are described in detail.

Preprocessing

To start the computational analysis, a set of secondary markers, subsequently referred to as the joints, was

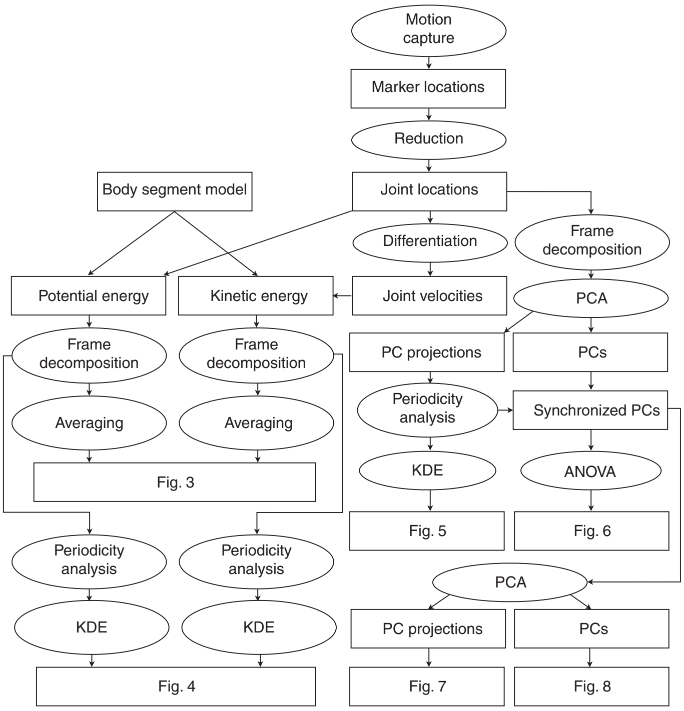


FIGURE 2. Overview of the analysis of the data. PCA = principal components analysis, PC = principal component, KDE = kernel density estimation.

derived from the marker locations. This was done in order to make the movement data compatible with the body-segment model used in subsequent analyses. The new set of twenty joints is depicted in Figure 1b. The locations of joints C, D, G, H, M, N, P, Q, R, S, and T are identical to the locations of one of the markers, while the locations of the other joints were obtained by averaging the locations of two or more markers. In addition to the joints obtained by the calculation described above, Figure 1b displays as lines the body segments that connect the joints. The following nomenclature will be used in subsequent discussion of the joints: A: root; B: left hip; C: left knee; D: left ankle; E: left toes; F: right hip; G: right knee; H: right ankle; I: right toes; J: midtorso; K: manubrium; L: head; M: left shoulder; N: left elbow; O: left wrist; P: left fingers; Q: right shoulder; R: right elbow; S: right wrist; T: right fingers.

The body segments considered in this study comprise five kinematic chains of body segments connected by joints. These are (see Figure 1b): ABCDE (left hip; left thigh; left leg; left foot), AFGHI (right hip; right thigh; right leg; right foot), AJKL (abdomen; thorax; head and neck), AJKMNOP (abdomen; thorax; left shoulder; left upper arm; left forearm; left hand), and AJKQRST (abdomen; thorax; right shoulder; right upper arm; right forearm; right hand).

Mechanical Energy

Next, we estimated two kinetic variables from the joint locations and velocities: potential energy and kinetic energy. Potential energy of a body of mass is a form of mechanical energy that depends on its height in the gravitational field. For instance, bending the knees while standing reduces the total potential energy of the body. Kinetic energy, on the other hand, is energy due to motion and thus depends on the speed of body of mass. The potential energy of the body was calculated using the formula

$$E_{pot}(t) = M\sum_{s} m_s g y_{s,c}(t), \tag{1}$$

where M stands for the total mass of the body, m_s the mass of body segment s relative to the total mass of the body, g the acceleration of gravity, and $y_{s,c}$ the vertical location of the center of mass of body segment s. The kinetic energy was calculated using the formula

$$\begin{aligned} E_{kin}(t) &= E_{trans}(t) + E_{rot}(t) \\ &= M\sum_{s}\left(\frac{1}{2}m_s v_{s,c}^2(t) + \frac{1}{2}I_{s,c}\omega_s^2(t)\right), \end{aligned} \tag{2}$$

where E_{trans} and E_{rot} denote translational and rotational energy, respectively, $v_{s,c}$ the speed of the center of mass of body segment s, $I_{s,c}$ the moment of inertia of body segment s, and ω_s the angular velocity of body segment s.

To estimate the instantaneous velocities of the joints and the angular velocities of the body segments needed for the calculation of kinetic energy, we utilized numerical differentiation to the joint location data. To this end, we used 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.

The inertial constants of the body were obtained from the body segment model proposed by Dempster, Gabel, and Felts (1959; see also Robertson, Caldwell, Hamill, Kamen, & Whittlesley, 2004). The model specifies a number of parameters for each of the body segments. These are the mass of the segment in relation to the total body mass, the distance of the center of mass from the proximal joint in relation to the segment length, and the radius of gyration in relation to segment length with respect to the center of mass, the proximal joint, and the distal joint.

To obtain an overview of the relationship of the total instantaneous kinetic and potential energies and the metrical structure of the musical stimulus, we decomposed the estimated energies into segments, each of which had a length equal to eight beats of the respective musical stimulus. Subsequently, the instantaneous energies were averaged across participants and across eight-beat segments in the stimuli. The results are shown in Figure 3. As can be seen, the average potential energy displays a clear periodicity at the one-beat level, the beat location closely matching the point of maximal decrease in potential energy and thus maximal downward velocity of the body. Average kinetic energy, on the other hand, displays a clear superposition of half-beat and two-beat periods.

Subsequently, we performed a periodicity analysis of the two components of mechanical energy. To this end, we first decomposed each motion capture recording into eight-beat sections with 50% overlap between neighboring sections. Subsequently, we estimated the period of potential and kinetic energy in each eight-beat segment using autocorrelation, which is a standard method for

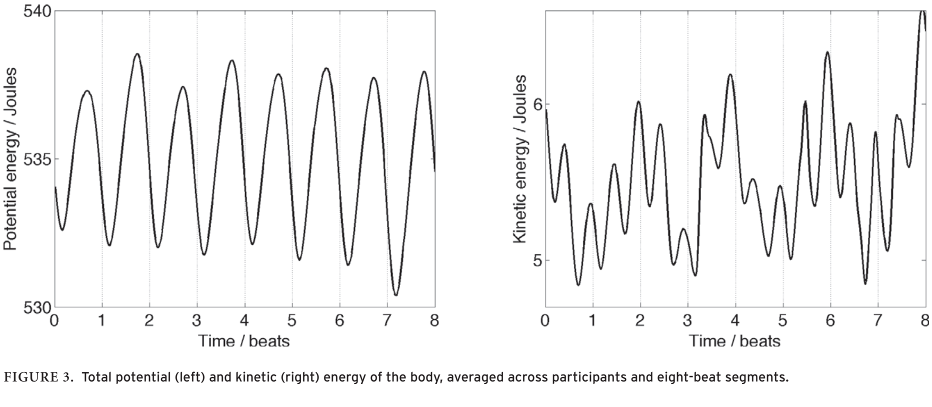


FIGURE 3. Total potential (left) and kinetic (right) energy of the body, averaged across participants and eight-beat segments.

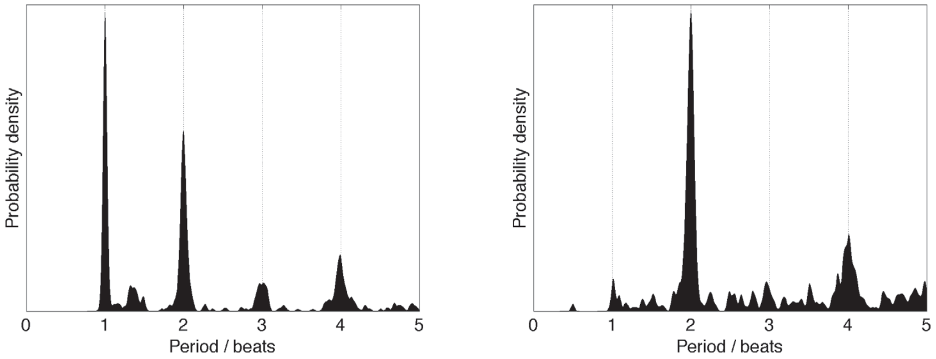


FIGURE 4. Kernel density estimate of distribution of periodicities for (a) potential energy and (b) kinetic energy.

periodicity estimation (Rabiner & Schafer, 1978). In the autocorrelation analysis, the period was defined to be equal to the lag corresponding to the first maximum in the autocorrelation function. Subsequently, we estimated the distribution of the thus obtained periods using Kernel Density Estimation (KDE; Silverman, 1986). KDE is a nonparametric method for estimating the probability distribution of a variable from a sample, and is based on summation of kernel function (such as narrow Gaussian curves) centered at each value in the sample. The thus estimated distributions of the periods of potential and kinetic energy are displayed in Figure 4.

As is evident from Figure 4, both forms of mechanical energy display periodicities at the three main metrical levels corresponding to periods of one, two, and four beats. For potential energy (Figure 4a), the most common periodicity for each tempo corresponds to the length of one beat, with a two-beat period being the second most prominent. For kinetic energy (Figure 4b), the most common period is two beats, while four-beat periods are also relatively common.

It is noteworthy that the potential energy (Figure 4a) shows a relatively prominent peak at the period of two beats, while this period is not clearly visible in Figure 3a. Similarly, for kinetic energy, the peak in Figure 4b at the period of half a beat is relatively small, while this periodicity is clearly visible in Figure 3b. These differences are due to the different methods used to produce these graphs. In Figure 3 the graphs have been obtained by averaging the mechanical energies across the eight-beat segments, while those in Figure 4 have been obtained by picking the most prominent periodicity in each eight-beat segment and estimating the distribution of these values. Thus, it may happen, for instance, that although the half-beat period is often present in the kinetic energy, it is mostly less strong than the two-beat period and is thus seldom picked as the strongest periodicity.

The graphs of Figure 4 suggest that for most of the eight-beat segments, the strongest periodicity lies in the vicinity of the period of one metrical level (one, two, or four beats). Therefore, in terms of mechanical energy, the movements were synchronized to some metrical level for most of the time.

Eigenmovements

As the next step in the analysis, we extracted typical movement patterns from the data and investigated their periodicities. To this end, we used Principal components analysis (PCA). PCA is a method that transforms a large group of variables into a reduced group of uncorrelated variables called principal components (PCs), which are linear combinations of the original variables. The first PC accounts for as much of the variance in the data as possible and the successive PCs each in turn account for as much of the remaining variance as possible. As a result of PCA, each observation in the data set can be expressed as values of the reduced set of variables, called the PC projections.

Again, the data were decomposed into eight-beat sections with 50% overlap between neighboring sections. This decomposition was carried out because many of the participants changed their movement patterns during the presentation of the stimuli; applying PCA to the whole duration of a stimulus would not have produced meaningful results. The decomposition resulted in 44 sections for each participant (11 sections per each tempo) and thus a total of 792 segments. For each segment, the motion capture data were rotated around the vertical axis so that the orientation of the frontal plane of the body, defined by the hip markers, was parallel to the first axis of the coordinate system. Subsequently, to reduce the dimensionality of the data, a PCA was carried out separately for each participant and each section. The data subjected to PCA thus consisted of a series of 60-component vectors. Each vector consisted of the three Euclidean coordinates of each of the 20 joints and there was one vector for each motion capture frame. The PCs obtained from this analysis were thus 60-component vectors, whose components indicated the amount to which each joint moved in each of the three directions in the respective movement pattern. The PC projections, on the other hand, were time series indicating how the movement pattern represented by the respective PC evolves in time.

Periodicity of Eigenmovements

For each segment, the first five PCs were included in the analysis. They accounted on average for 96.7% of the variance in each segment (*SD* = 2.3%). The PC projections from each PCA were subjected to a periodicity analysis using autocorrelation. Again, we estimated the distribution of the periods of the Principal Component projections using kernel density estimation. The result is shown in Figure 5. As can be seen, the most common periods correspond to the three main metrical levels (i.e., periods of one, two, and four tactus beats), with the two-beat period being the most prevalent, followed by four beats and one beat, in this order.

Next, we investigated the prevalence of eigenmovements embodying each metrical level. To this end, we compared the estimated periods of the eigenmovements with the period of one, two, and four beats of the respective musical stimulus. We found that 12% of the segments contained a PC that was synchronized to the one-beat level, in the sense that its period differed from that of one musical beat by less than 10%. For the two-beat and four-beat levels, these proportions were 47% and 34%, respectively. On average, the synchronized eigenmovements comprised a total of 29% of the total kinetic energy of the movements. For 65% of the segments, only one of the eigenmovements coincided with the period of any of the three metrical levels. According to this criterion, the participants thus embodied only one metrical level at a time for most of the segments. The eigenmovements synchronized to the three main metrical levels will be subsequently referred to as *eigenmodes.*

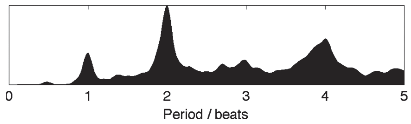


FIGURE 5. Distribution of periods of eigenmovements.

Structure of Eigenmodes

As a next step in the analysis, we investigated whether the eigenmodes embodying each of the three metrical levels differed from each other. More specifically, the aim of this analysis was to investigate whether there were differences between the metrical levels in terms of the movement of each joint along each of the three directions (mediolateral, anteroposterior, vertical).

To investigate these differences, we first performed a one-way ANOVA on each of the 60 components of the PCs. For the purpose of the analysis, for each participant and each metrical level, the PCs whose projections displayed a periodicity that was within ±10% of the length of the respective period of the metrical level where chosen. Subsequently, we calculated for each of the 60 components the average value across the selected PCs. By doing this we obtained one averaged PC for each of the 18 participants and for each of the three metric levels.

For the calculation of the PC averages, absolute values of the components were used. This was because the PCs are only defined up to a multiplicative constant (Strang, 1976). For instance, if **p** is a PC obtained from a data set, so is -**p**. Therefore, if absolute values were not used, the averaging may have cancelled out some PCs having opposite signs. The ANOVAs thus regarded the amount of movement of each joint along each of the three directions without any regard to the direction of movement.

Due to the large number of comparisons (60), we used Bonferroni correction. The level of significance was thus set to .05/60 = .00083. Table 1 displays the joints and movementdirections that showed a significant effect of metrical level in the amount of movement.

The results of the posthoc tests for each comparison and each direction are displayed graphically in Figure 6. In this figure, the joints displaying significant differences are shown as black circles. To summarize the findings displayed in this figure, the main significant difference between the one-beat and two-beat eigenmodes (panels a and b) is that one-beat eigenmodes displayed more movement of shoulders in the vertical direction. Posthoc tests between one-beat and four-beat eigenmodes (panels c and d) reveal that one-beat eigenmodes are

TABLE 1. Results from ANOVA Testing Differences in the Amount of Movement Between the Three Metrical Levels. Only Significant Effects Are Shown.

Direction	Joint	Joint symbol	$F(2, 51)$	p
Vertical	Root	A	10.07	.00026
	Left hip	B	10.05	.00027
	Right hip	F	12.57	.00005
	Manubrium	K	10.59	.00019
	Head	L	13.57	.00003
	Right shoulder	Q	11.57	.00010
	Right elbow	R	9.17	.00050
	Right wrist	S	9.32	.00045
	Left shoulder	M	12.93	.00004
	Left elbow	N	9.69	.00035
	Left knee	C	8.82	.00063
Mediolateral	Root	A	10.84	.00016
	Manubrium	K	16.08	.00001
	Head	L	10.97	.00015
	Right shoulder	Q	10.21	.00024
	Right elbow	R	9.63	.00036
	Left shoulder	M	20.42	.00000
	Left hip	B	12.08	.00007
	Right elbow	R	9.86	.00031
Anteroposterior	Right wrist	S	9.28	.00046
	Left wrist	O	11.76	.00009

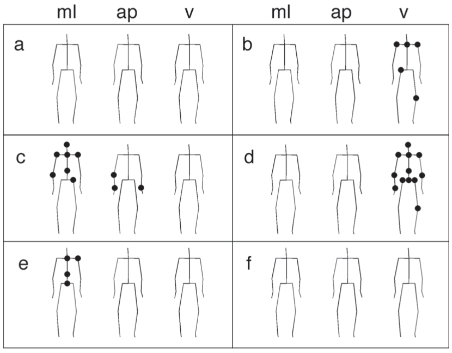


FIGURE 6. Significant differences in the amount of movement of each joint along each direction between eigenmodes (ml = mediolateral direction; ap = anteroposterior direction; v = vertical direction). For each comparison and each direction, joints having significant differences are displayed with black circles; (a) less movement in one-beat eigenmodes than in two-beat eigenmodes; (b) more movement in one-beat eigenmodes than in two-beat eigenmodes; (c) less movement in one-beat eigenmodes than in four-beat eigenmodes; (d) more movement in one-beat eigenmodes than in four-beat eigenmodes; (e) less movement in two-beat eigenmodes than in four-beat eigenmodes; (f) more movement in two-beat eigenmodes than in four-beat eigenmodes.

associated with less movement of the upper torso in the mediolateral direction and hands in the anteroposterior direction than four-beat eigenmodes, while one-beat eigenmodes are associated with more movement of torso and arms in the vertical direction than four-beat eigenmodes. Finally, two-beat eigenmodes showed less movement of the torso in the mediolateral direction than four-beat eigenmodes (panels e and f).

Between-Subjects PCA

To investigate more closely the nature of different movement patterns at the different metrical levels, a second (between-subjects) PCA was carried out on the one-beat, two-beat, and four-beat eigenmodes. In other words, the PCs obtained from the PCA described above, whose projections where synchronized to one of the three metrical levels (allowing a tolerance of 10% of the length of musical beat), were subjected to the between-subjects PCA.

For subsequent analysis we decided to include PCs that contained a minimum of 75% of the variance. Using this criterion, we ended up with six PCs that contained 76.5% of the variance. The PCs obtained from this second-order PCA subsequently are referred to as the second-order PCs. To assess the degree to which each of the second-order PCs represents movement patterns synchronized to each of the metrical levels, the proportion of variance contained in the second-order PC component scores within each of the three metrical levels was quantified. The rationale behind this procedure was that if the second-order component scores for a given metric level display a high amount of variance, the respective second-order PC has a high degree of similarity to the first-order PCs representing this metric level. This follows from the fact that the PC component score is obtained by calculating the dot product between the first-order PC and the second-order PC, and since the PCs are normalized to have unit length, the dot product is equal to the cosine of the angle between first-order PCs and the second-order PC. Thus, such a second-order PC can be regarded to represent a typical movement pattern for the metric level.

The proportion of variance contained in the projections onto the first PCs are shown in Figure 7. As is evident from this figure, the metrical levels differ in terms of the variance contained in each of the PC projections. In particular, the eigenmovements synchronized to the one-beat level retain the highest proportion of variance when projected onto PCs 3 and 5. Similarly, the eigenmovements synchronized to the two-beat and four-beat levels retain the highest proportion of

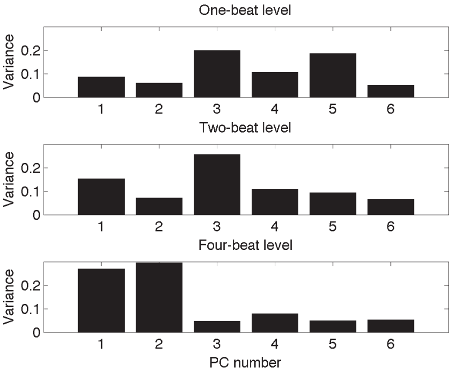


FIGURE 7. Proportion of variance contained in the projections of the primary PCs onto the secondary PCs for each metrical level.

variance when projected onto PCs 3 and 1, and 2 and 1, respectively.

The six secondary PCs are graphically depicted in Figure 8. In what follows, PCs 1, 2, 3, and 5, which show the highest variances in Figure 7, are analyzed in detail. PC 1 contains high values in the anteroposterior components of the shoulders, elbows, wrists, and fingers (gray bars at joints M-T). Moreover, the components corresponding to the two sides of the body (M-P vs. Q-T) have opposite signs. PC 1 thus corresponds to the rotation of the upper torso. PC 2 displays large values in the mediolateral components of most joints (black bars). Moreover, the components have the same sign. Therefore, PC 2 corresponds to lateral swaying of the body. PC 3 displays largest values in the mediolateral components of the elbows, wrists, and fingers (black bars at joints N-P and R-T). Again, the components corresponding to the two sides of the body (N-P vs. R-T) have opposite signs. Therefore, PC 3 corresponds to antiphase mediolateral arm movement. Finally, PC 5 shows largest values in the vertical components of elbows, wrists, and fingers (joints N-P and R-T). Therefore, PC 5 corresponds to vertical movement of the arms.

Based on this analysis, we can identify the following typical movements for each of the metrical levels.

- One-beat level: (1) mediolateral arm movements (PC 3); (2) vertical arm movements (PC 5);
- Two-beat level: (1) mediolateral arm movements (PC 3); (2) rotation of the upper torso (PC 1);
- Four-beat level: (1) lateral swaying of the body (PC 2); (2) rotation of the upper torso (PC 1).

Discussion

We investigated music-induced movement, focusing particularly on the relationship between movement patterns and metrical levels of music. A kinetic analysis of peaks in mechanical energy (potential energy and kinetic energy) revealed that participants embodied the musical stimulus on several metrical levels. A subsequent kinematic analysis of the periodicity structure of the movement revealed that participants synchronized with periods of one, two, and four beats, supporting the findings from the kinetic domain. The analysis also showed that several metrical levels can be embodied simultaneously in the movement, although participants mostly tended to embody only one metrical level at a time. Regarding direction of movement at different metrical levels, we found that the tactus (one-beat level) tended to be embodied mostly as vertical movement, while the four-beat level tended to be embodied as mediolateral movement. This observation is consistent with the kinetic finding that the average potential energy of the body displays a period of one beat.

A more detailed kinematic analysis revealed that the tactus level often was associated with vertical hand and torso movements as well as mediolateral arm movements, the two-beat level with mediolateral arm movements and rotation of the upper torso, and the four-beat level with lateral flexion of the torso and rotation of the upper torso. This observation is in line with our hypothesis that faster metric levels are embodied in the extremities, and slower ones in the central parts of the body. Furthermore, this finding is in tune with the kinetic observation that the most prevalent period of kinetic energy is two beats. The reasons for this are as follows. First, the torso has a significant mass, and thus its movement contributes significantly to the body's total kinetic energy. Second, because in periodic motion kinetic energy obtains two maxima and two minima during one cycle, a two-beat period in kinetic energy corresponds to a four-beat period of movement.

An interesting observation was that, compared to two-beat and four-beat periods, one-beat period was relatively rare in the eigenmovements. There are two possible explanations for this finding. First, the method of Principal components analysis extracts movement patterns that display the highest variance. It is thus possible that there exist movement patterns with one-beat period, but they may possess a low degree of variance and are thus not discovered by this method. This finding also can be explained by the type of variable that was used in the analysis. In particular, we used marker locations in the analysis. Using some other kinematic variable, such as marker speed, would have resulted in different periodicity

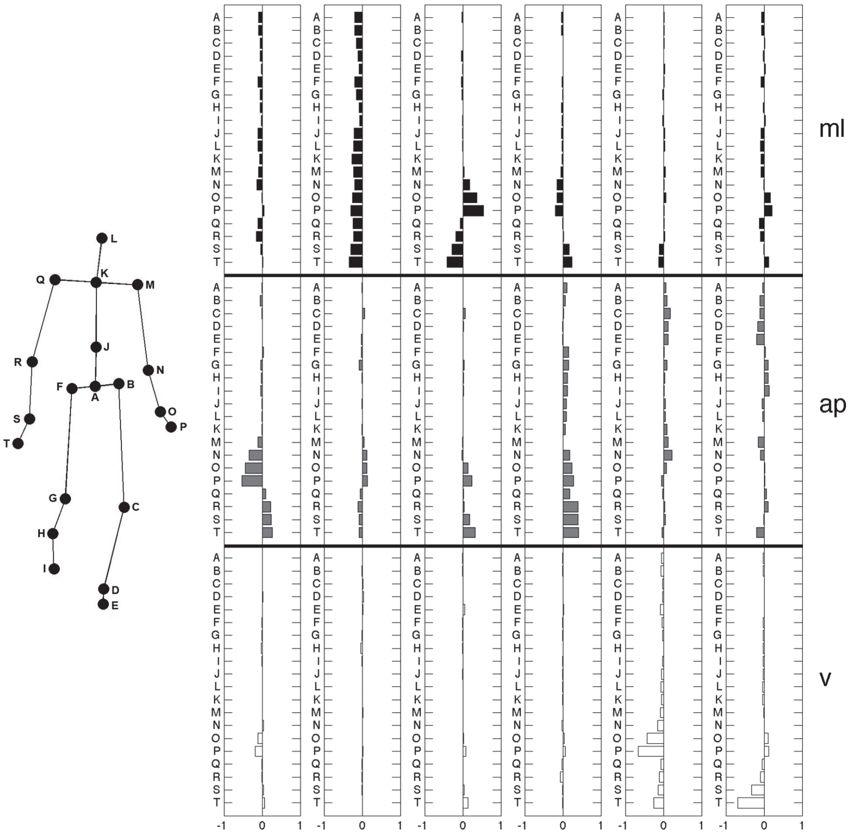


FIGURE 8. First six second-order PCs obtained from the analysis. The first PC is in the left and the sixth in the right. On the left side of each graph, the letters refer to the joints as depicted in the left of the figure. The components are shown separately for mediolateral (ml), anteroposterior (ap), and vertical (v) dimensions.

distributions. For instance, for a mediolateral swaying of the body with a two-beat period, the speed of the torso shows a period of one beat. Also, to produce such movement, overall muscular effort has a period of one beat (muscular effort is maximal at the points where the tilt of the body is maximal). Therefore, one can say that even if this kind of movement pattern shows a period of two beats in terms of marker locations, the one-beat level is embodied in the muscular activity.

A possible limitation of the present study is that it used only one musical stimulus, albeit presented with a range of different tempi. Consequently, it may be that the observed movement patterns represent some musical features characteristic to this stimulus and thus are not generalizable to other musical stimuli. In subsequent work, we plan to include stimuli representing other various musical genres and several time signatures.

One could assume that the tempo at which the stimulus was presented to the participant could have had an effect on the way they moved to the music. For instance, given our preference to move at frequencies close to 2 Hz (McDougall & Moore, 2005), one could expect that stimuli whose tempi are in the vicinity of 120 bpm would elicit larger movements than other stimuli. We compared the average kinetic energies of the participants across the different tempi, but failed to observe any significant differences. More research is needed to clarify this issue.

A further potential limitation of the present study is that the kinematic analysis focused entirely on the location of the markers, while ignoring other kinematic variables such as velocity, acceleration, and jerk. Analysis of these kinematic variables could potentially reveal patterns that remain hidden in the location-based analysis.

The use of principal components analysis for dimensionality reduction makes certain assumptions about the analyzed data that are not necessarily fulfilled by the present data. First, the Principal Components are noncorrelated, or orthogonal, and thus associating the eigenmodes with them makes the assumption that the movement patterns synchronized to the different metrical levels are orthogonal as well. This, however, may not be the case. Therefore, using nonorthogonal dimensionality reduction methods such as Independent Component Analysis (Hyvärinen, Karhunen, & Oja, 2001) or Denoising Source Separation (Särelä & Valpola, 2005) for eigenmode extraction could provide new insights into the problem.

A second assumption of principal components analysis is that the analyzed data are stationary within the analysis window. This may hold true when analyzing simple movement patterns such as walking, but with music-induced movement the patterns may change over time and the stationarity criterion is not met. This problem could be overcome by using eigenmode extraction methods that do not assume stationarity, such as Empirical Mode Decomposition (Huang et al., 1998).

The aforementioned limitations notwithstanding, our study is the first to tackle the question of music listeners' movement and embodiments of musical meter using quantitative methods. We have shown that several levels of metrical hierarchy simultaneously can be embodied in the movements. Furthermore, we have identified the most typical movement types of different metric levels. In the future, we expect to gain a deeper insight into this phenomenon using more varied stimulus sets and improved methods of analysis, as described above.

Author Note

We wish to thank Marc Leman and two anonymous reviewers for helpful comments, and Aki Tulikari for his help in collecting the motion capture data. This study was supported by the Academy of Finland (projects 118616, 125710, and 119959).

Correspondence concerning this article should be addressed to Petri Toiviainen, Department of Music, PO Box 35(M), 40014 University of Jyväskylä, Finland. E-MAIL: petri.toiviainen@jyu.fi

References

AROM, S. (1991). *African polyphony and polyrhythm*. Cambridge, UK: Cambridge University Press.

CROSS, I. (2003). Music, cognition, culture and evolution. In I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 42–56). New York: Oxford University Press.

DEMPSTER, W. T., GABEL, W. C., & FELTS, W. J. L. (1959). The anthropometry of manual work space for the seated subject. *American Journal of Physical Anthropology, 17*, 289–317.

DRAKE, C., PENEL, A., & BIGAND, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18*, 1–23.

EEROLA, T., LUCK, G., & TOIVIAINEN, P. (2006, August). *An investigation of pre-schoolers' corporeal synchronization with music.* Paper presented at the 9th International Conference on Music Perception and Cognition, Bologna, Italy.

FRAISSE, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 148–180). New York: Academic Press.

HUANG, N. E., SHEN, Z., LONG, S. R., WU, M. C., SHIH, H. H., ZHENG, Q., ET AL. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proceedings of the Royal Society of London. Series A, 454*, 903–995.

HYVÄRINEN, A., KARHUNEN, J., & OJA, E. (2001). *Independent component analysis*. New York: Wiley.

LARGE, E. W., FINK, P., & KELSO, J. A. (2002). Tracking simple and complex sequences. *Psychological Research, 66*, 3–17.

LEMAN, M. (2008). *Embodied music cognition and music mediation technology*. Cambridge, MA: MIT Press.

LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

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.

LIBERMAN, A. M., & MATTINGLY, I. G. (1985). The motor theory of speech-perception revised. *Cognition, 21*, 1–36.

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

NAVEDA, L., & LEMAN, M. (2010). The spatiotemporal representation of dance and music gestures using topological gesture analysis (TGA). *Music Perception, 28*, 93–111.

PALMER, C., & KRUMHANSL, C. L. (1990). Mental representations of musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 728–741.

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

PATEL, A. D., IVERSEN, J. R., CHEN, Y., & REPP, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research, 163*, 226–238.

PHILLIPS-SILVER, J., & TRAINOR, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308*, 1430.

PHILLIPS-SILVER, J., & TRAINOR, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition, 105*, 533–546.

PHILLIPS-SILVER, J., & TRAINOR, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition, 67*, 94–102.

RABINER, L. R., & SCHAFER, R. W. (1978). *Digital processing of speech signals*. Englewood Cliffs, NJ: Prentice-Hall.

REPP, B. H. (2005a). Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 1. Qualitative observations. *Music Perception, 22*, 479–496.

REPP, B. H. (2005b). Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 2. The roles of different kinds of accent. *Music Perception, 23*, 165–187.

REPP, B. H., & PENEL, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research-Psychologische Forschung, 68*, 252–270.

ROBERTSON, D. G. E., CALDWELL, G. E., HAMILL, J., KAMEN, G., & WHITTLESLEY, S. N. (2004). *Research methods in biomechanics*. Champaign, IL: Human Kinetics.

SÄRELÄ, J., & VALPOLA, H. (2005). Denoising source separation. *Journal of Machine Learning Research, 6*, 233–272.

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

SETHARES, W. A., & STALEY, T. W. (2001). Meter and periodicity in musical performance. *Journal of New Music Research, 30*, 149–158.

SILVERMAN, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman Hall.

SNYDER, J., & KRUMHANSL, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception, 18*, 455–489.

STRANG, G. (1976). *Linear algebra and its applications.* New York: Academic Press.

STYNS, F., VAN NOORDEN, L., MOELANTS, D., & LEMAN, M. (2007). Walking on music. *Human Movement Science, 26*, 769–785.

TODD, N. P. M., COUSINS, R., & LEE, C. S. (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Musicology Review, 2*, 1–13.

TODD, N. P. M., O'BOYLE, D. J., & LEE, C. S. (1999). A sensory-motor theory of rhythm, time perception and beat induction. *Journal of New Music Research, 28*, 5–28.

TOIVIAINEN, P., & SNYDER, J. S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21*, 43–80.

TRAINOR, L. J., GAO, X., LEI, J.-J., LEHTOVAARA, K., & HARRIS, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex, 45*, 35–43.

WALLIN, N. L., MERKER, B., & BROWN, S. (2000). *The origins of music*. Cambridge, MA: MIT Press.

VAN NOORDEN, L., & MOELANTS, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research, 28*, 43–66.

IV

EFFECTS OF THE BIG FIVE AND MUSICAL GENRE ON MUSIC-INDUCED MOVEMENT

by

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

Journal of Research in Personality vol. 44(6), 714-720.

Journal of Research in Personality 44 (2010) 714–720

Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

Effects of the Big Five and musical genre on music-induced movement

G. Luck *, S. Saarikallio, B. Burger, M.R. Thompson, P. Toiviainen

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

ARTICLE INFO

Article history:
Available online 30 October 2010

Keywords:
Personality
Musical style
Motion-capture
PCA
MANOVA

ABSTRACT

Nine-hundred-and-fifty-two individuals completed the Big Five Inventory, and 60 extreme scorers were presented with 30 music excerpts from six popular genres. Music-induced movement was recorded by an optical motion-capture system, the data from which 55 postural, kinematic, and kinetic movement features were computed. These features were subsequently reduced to five principal components of movement representing Local Movement, Global Movement, Hand Flux, Head Speed, and Hand Distance. Multivariate Analyses revealed significant effects on these components of both personality and genre, as well as several interactions between the two. Each personality dimension was associated with a different pattern of movement characteristics, with Extraversion and Neuroticism eliciting the clearest relationships. Latin, Techno, and Rock music, meanwhile, most clearly elicited different movement characteristics.

1. Introduction

The idea that much of human cognition is based on the corporeal (e.g., Lakoff & Johnson, 1980) has gained increasing momentum in recent decades. The embodied view of human cognition (e.g., Varela, Thompson, & Rosch, 1991), and more recently the notion of embodied music cognition (e.g., Leman, 2007), both stress the importance of an organism's sensorimotor capacities, body, and environment in cognitive processes. Such embodiment can be seen, for example, in the way music listeners employ foot-tapping and body-swaying movements to parse musical structure (e.g., Keller & Rieger, 2009). Music is an effective means of eliciting expressive movement, too, and in many cultures music and movement are inseparable from each other. Many factors will likely affect the precise characteristics of these various music-induced movements. Here, we examine the effect of two such factors: the listener's personality, and the genre of the music.

Personality is integrally related to movement and expression in general, and expressive behaviour is considered to be a key element in understanding personality and individual differences (e.g., Gross, 1999). For instance, people use body motions as reliable indicators of others' personality type (Ball & Breese, 2000), and even the movements of robots have been shown to elicit attributions of "personality" by observers (Heeyoung, Kwak, & Myungsuk, 2008). Moreover, relationships between speakers' movement patterns and perceptions of their personality have been identified (Koppensteiner & Grammer, 2010), and personality disorders have also been found to be reflected in bodily movement (Kluft, Poteat, & Kluft, 1986).

In terms of defining personality, the Big Five personality traits (Costa & McCrae, 1992) (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) are widely considered to describe personality at the highest level of organization (Goldberg, 1993). Relationships between these five factors and a variety of music-related phenomena, including musical preference (e.g., Rentfrow & Gosling, 2003), and perception and experience of emotion in music (e.g., Vuoskoski & Eerola, 2010), are well-documented. For example, Rentfrow and Gosling (2003) found that Openness was positively related to a preference for "reflective and complex" as well as "intense and rebellious" music, whereas Extraversion was related to both "upbeat and conventional" and "Energetic and Rhythmic" music. Vuoskoski and Eerola (2010), meanwhile, found that, in line with the trait-congruency hypothesis (e.g., Rusting, 1998), Neuroticism was positively related, but Extraversion negatively related, to perceived sadness in film music.

Given connections between personality and movement, and personality and various music-related phenomena, it would seem reasonable to suppose that there exist relationships between personality and music-related movement. Since music is an effective means of eliciting body movement, music-induced movement in particular might be an effective means of studying such relationships.

In a recent study, several of the current authors identified, albeit tentatively, such connections between the Big Five and music-induced movement (Luck, Saarikallio, & Toiviainen, 2009). Neuroticism was found to be related to jerky and accelerated movement, whereas both Openness to experience and Agreeableness were related to smoother movement. Extraversion and Conscientiousness,

---

* Corresponding author. Address: Department of Music, P.O. Box 35(M) 40014, University of Jyväskylä, Finland. Fax: +358 14 260 1331.
E-mail address: geoff.luck@jyu.fi (G. Luck).

0092-6566/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jrp.2010.10.001

meanwhile, were related to higher speed of movement. However, although the results of this study demonstrated the possibility of links between long-term stable personality traits and the way people moved to music, the study itself had two fundamental problems. First, the number of participants was relatively small, thus reducing the reliability of the results. Second, participants were selected randomly rather than on the basis of the distribution of their scores along the Big Five personality dimensions. Differences between high- and low-scoring participants would be easier to identify, for example, if the selection process was based on prior knowledge of participants' personality scores.

Another shortcoming of our previous work was that participants were presented with only a single piece of music, an instrumental twelve-bar blues progression written especially for the study. There is reason to believe, however, that listeners' music-induced movements would be influenced by the genre of the music they are listening to.

Certain movements may be more representative of particular genres (e.g., Godøy & Jensenius, 2009), such as the way listeners tend to nod their head or tap their foot when listening to jazz music (e.g., Schmidt, 2002), or indeed remain still when listening to classical music. Stereotypical and culturally-driven movements associated with particular genres, such as swaying of the hips to Latin music, and ''air guitar'' or ''head banging'' motions associated with Rock music, will likely also play a part. Thus, it would seem pertinent to include music from different genres in any investigation of music-induced movement. Yet, to date, there appears to be no empirical work examining the effect of musical genre on listeners' bodily movement.

In the present study, we address the shortcomings of our previous work in relation to both personality and musical genre. As regards personality, we increased our sample size threefold, selecting extreme scorers from a large pool of individuals whose personality had been assessed a priori. As regards genre, musical stimuli were drawn from six different popular music genres, namely Jazz, Latin, Techno, Funk, Pop, and Rock. Given the lack of previous research, specific relationships between personality, genre, and music-induced movement were hard to predict. Our aim instead was to build on our earlier work by exploring the effects of these factors on music-induced movement in a more sophisticated study. In so doing, we hoped to increase our understanding of the effects of both personality and music on human behaviour.

2. Method

2.1. Participants

Participants were selected from a pool of 952 individuals who had previously completed an online version of the Big Five Inventory (BFI) (John, Naumann, & Soto, 2008), a 44-item instrument measuring the five primary personality dimensions (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). Our initial aim was to recruit the six highest- and lowest-scoring individuals possible on each of the five dimensions. Based upon responses from the individuals contacted, we eventually recruited a total of 64 participants. Four participants were excluded from further analyses, however, due to being seated or leaving the capture space during recording. Thus, 60 participants were retained (43 female, 17 male; mean age = 24, SD = 3.3). In order to increase the amount of data at our disposal for each analysis, these participants were ultimately divided into low and high groups by selecting those scoring below the 33rd percentile and those scoring above the 67th percentile, respectively, on each personality dimension. This resulted in 16–19 participants being include in each low/high group. All participants were students from different faculties of the University of Jyväskylä, Finland, and all but two were of Finnish nationality. Six participants had received formal music education, and four participants had received formal dance tuition. Participants were rewarded for their time with a movie ticket.

2.2. Apparatus, stimuli, and procedure

Participants were presented with 30 randomly-ordered musical stimuli, each 30 s in length. Five stimuli were selected from each of the following popular music genres: Jazz, Latin, Techno, Funk, Pop, and Rock. All stimuli were non-vocal and in 4/4 metre, but differed in their rhythmic complexity, pulse clarity, and tempo (82–199 beats per minute, mean = 130 beats per minute). Twenty-eight reflective markers (see Fig. 1 a) were attached to each participant, and their three-dimensional position tracked by an eight-camera optical motion-capture system (Qualisys ProReflex) at 120 frames per second. 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 2.5 m. The microphone input, the direct audio signal of the playback, and the TTL[1] 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 afterwards. Four Sony video cameras were used to record the sessions for reference purposes. Participants were recorded individually, and were asked to move in a way that felt natural with regards to the stimuli presented. Additionally, participants were encouraged to dance if they wanted to, but were requested to try to remain in the centre of the capture space indicated by a 115 × 200 cm carpet.

2.3. Movement feature extraction

To facilitate extraction of kinematic features using the MATLAB Motion Capture (MoCap) Toolbox (Toiviainen & Burger, 2010), 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 Fig. 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, and the locations of the remaining joints were obtained by averaging the locations of two or more markers. Subsequently, the data were transformed to a local coordinate system, in which the location of each joint was expressed in relation to the location of the midpoint of the hips (Joints B and F).

Subsequently, kinematic variables relating to instantaneous velocity, acceleration, and jerk of various body parts were estimated from the joint location data. Jerk refers to the time derivative of acceleration, and thus indicated the rate of change of acceleration of the respective body part (essentially measuring smoothness/jerkiness of the movement). These kinematic variables were estimated using 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. Additionally, instantaneous kinetic energy was estimated using body-segment modelling (see Toiviainen, Luck, & Thompson, 2010). Subsequently, the instantaneous values of each variable were averaged for each stimulus presentation. This yielded a total of 55 postural, kinematic, or kinetic features for each of the 60 participants and 30 stimuli.

Next, the dimensionality of the data was reduced by employing Principal Component Analysis (PCA). Using the Kaiser crite-

---

[1] Transistor transistor Logic. In the case of the Qualisys cameras, the TTL pulse is a timing signal transmitted at the onset of every frame captured. This signal allows us to synchronize the movement data with the audio.

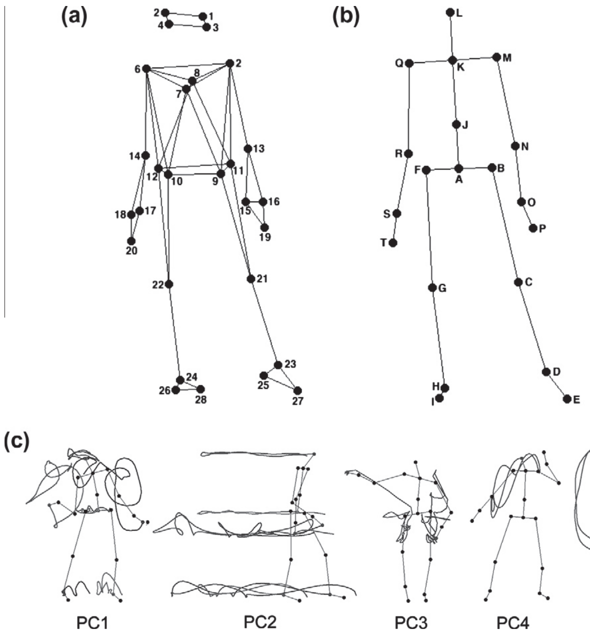


Fig. 1. (a) Anterior view of the location of the markers attached to participants' bodies; (b) anterior view of the locations of the secondary markers/joints used in the analysis; (c) three-second movement traces for the performances that received the highest score for each of the five principal components. The following joints were taken to visualize the traces: L (head), A (centre of mass), P/T (fingers), and I/E (feet). For clarity, only traces for body parts most relevant to each principal component are shown.

rion, five principal components (PCs), which accounted for 90.3% of the variance in the original features, were retained (see Fig. 1c). Although the initial solution was readily interpretable, virtually all the variance was explained by the first component. Thus, the initial solution was rotated using the varimax procedure so as to more evenly distribute the explained variance across the five PCs. Interpretation of the PCs was not unduly affected by the rotation.

PC 1 explained 38.2% of the variance, and had high loadings for features relating to jerky and accelerated movement of the centre of mass, head, feet, and hands, particularly in the vertical direction. PC 1 was therefore labelled (amount of) Local Movement. PC 2 explained a further 25.8% of the variance, and had high loadings for features relating to speed of the feet, bounding rectangle (the smallest rectangle that contains the projection of trajectory of the centre of mass on the horizontal plane, i.e., the floor), and kinetic energy, both anterioposterior and mediolateral. PC 2 was consequently labelled (amount of) Global Movement. PC 3 explained an additional 11.4% of the variance, and had high loadings for features related to vertical, anterioposterior, mediolateral, and overall jerk of hand movement. PC 3 was thus labelled Hand Flux. PC 4 explained a further 7.6% of the variance, and had high loadings for features related to anterioposterior, mediolateral and vertical speed of the head, and torso tilt. PC 4 was therefore labelled Head Speed. PC 5 explained an additional 7.3% of the variance, and had high loadings for features related to size, as well as anterioposterior, mediolateral, and overall speed, of hand movement. PC 5 was consequently labelled Hand Distance.

3. Results

The effect of personality and genre on the five components of movement was subsequently investigated by running a series of 2 (personality dimension: low-scorers versus high-scorers) × 6 (genre) Multivariate Analyses of Variance (MANOVA), one for each of the five personality dimensions. The main effect of personality on participants' movement is shown in Fig. 2, and the main effect of genre on participants' movement is shown in Fig. 3. For both of these figures, non-significant effects are faded-out. Also note that in Fig. 3, and in the description of results below, the main effect of Genre is reported separately for each personality dimension since different participants were included in each analysis. Finally, significant personality × genre interaction effects are shown in Fig. 4.

3.1. Openness

Multivariate tests (using Pillai's Trace) were significant for both personality, $V = .020$, $p = .002$ and genre, $V = .115$, $p = .000$, but not for the personality × genre interaction, $V = .028$, $p = .416$. Specifically, there was a significant main effect of Openness on Local Movement only, $F(1, 933) = 12.869$, $p = .000$, with high scorers exhibiting higher levels of Local Movement (see Fig. 2, first row). In addition, there was a significant main effect of Genre on Local Movement, $F(5, 933) = 8.577$, $p = .000$, Global Movement, $F(5, 933) = 3.066$, $p = .009$, and Head Speed, $F(5, 933) = 7.643$, $p = .000$ (see Fig. 3, first row). Post Hoc comparisons with Bonferroni correction revealed that Local Movement was significantly higher for Techno ($M = .446$,

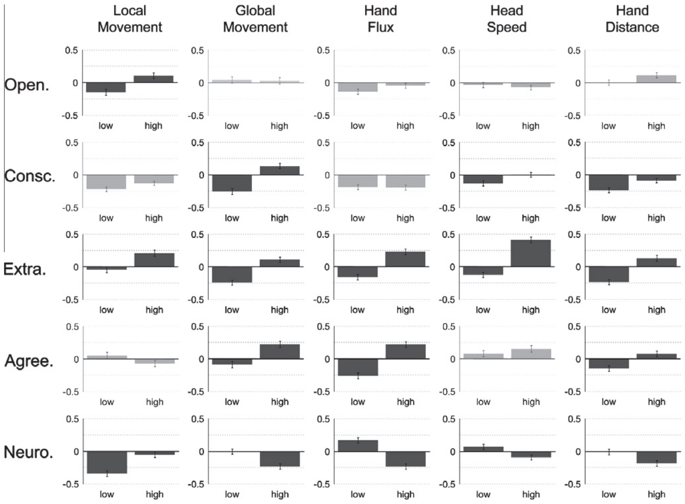


Fig. 2. Results of Multivariate Analyses of Variance (MANOVAs) showing the effect (±*SE*) of the Big Five personality traits on the five principal components of movement. The vertical axis scale refers to the size of the component score. Non-significant effects are faded-out.

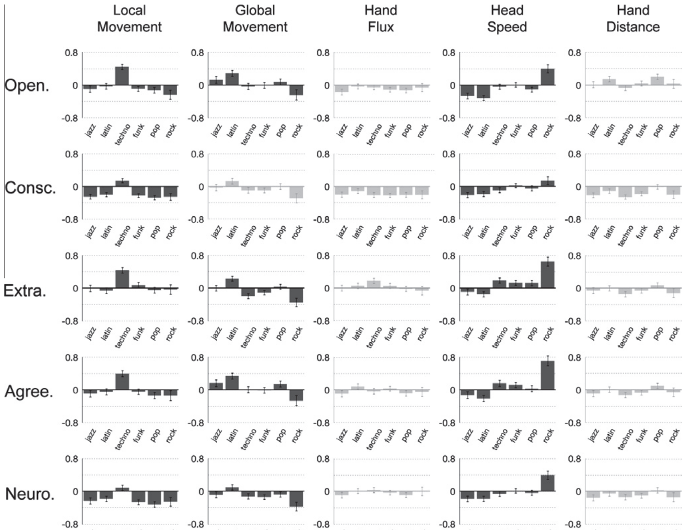


Fig. 3. Main effect of the six genres on the five principal components of movement. Main effects are reported separately for each personality dimension due to different participants being included in each analysis. The vertical axis scale refers to the size of the component score. Non-significant effects are faded-out.

SE = .077) than for all other genres, that Global Movement was significantly higher for Latin (*M* = .289, *SE* = .083) compared to Rock (*M* = −.242, *SE* = .132), that Head Speed was significantly higher for Rock (*M* = .401, *SE* = .110) than for all other genres, and that Head Speed was significantly lower for Latin (*M* = −.314, *SE* = .070) compared to Funk (*M* = .008, *SE* = .063).

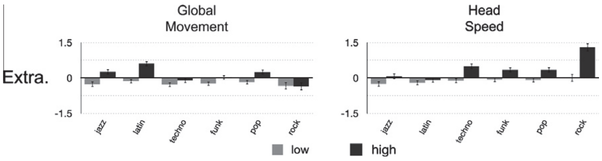


Fig. 4. Significant personality × genre interaction effects. The vertical axis scale refers to the size of the component score.

3.2. Conscientiousness

Multivariate tests were significant for both personality, $V = .046$, $p = .000$, and genre, $V = .076$, $p = .000$, but not for the personality × genre interaction, $V = .028$, $p = .426$. Specifically, there was a significant main effect of Conscientiousness on Global Movement, $F(1, 906) = 31.234$, $p = .000$, Head Speed, $F(1, 906) = 4.707$, $p = .030$, and Hand Distance, $F(1, 906) = 6.449$, $p = .011$, with high scorers exhibiting higher levels of all three of these components (see Fig. 2, second row). In addition, there was a significant main effect of Genre on Local Movement, $F(5, 906) = 6.762$, $p = .000$, and Head Speed, $F(5, 906) = 2.634$, $p = .022$ (see Fig. 3, second row). Post Hoc comparisons with Bonferroni correction revealed that Local Movement was significantly higher for Techno ($M = .145$, $SE = .060$) than for all other genres.

3.3. Extraversion

Multivariate tests were significant for both personality, $V = .158$, $p = .000$, and genre, $V = .107$, $p = .000$, as well as for the personality × genre interaction, $V = .045$, $p = .013$. Specifically, there was a significant main effect of Extraversion on Local Movement, $F(1, 960) = 12.026$, $p = .001$, Global Movement, $F(1, 960) = 30.439$, $p = .000$, Hand Flux, $F(1, 960) = 34.559$, $p = .000$, Head Speed, $F(1, 960) = 60.584$, $p = .000$, and Hand Distance, $F(1, 960) = 28.110$, $p = .000$, with high scorers exhibiting higher levels of all five components (see Fig. 2, third row). In addition, there was a significant main effect of Genre on Local Movement, $F(5, 960) = 4.941$, $p = .000$, Global Movement, $F(5, 960) = 6.232$, $p = .000$, and Head Speed, $F(5, 960) = 8.043$, $p = .000$ (see Fig. 3, third row). Post Hoc comparisons with Bonferroni correction revealed that Local Movement was significantly higher for Techno ($M = .437$, $SE = .079$) compared to all other genres, that Global Movement was significantly higher for Latin ($M = .234$, $SE = .069$) compared to Techno ($M = -.194$, $SE = .069$), Funk ($M = -.109$, $SE = .063$), and Rock ($M = -.349$, $SE = .110$), that Head Speed was significantly higher for Rock ($M = .651$, $SE = .118$) compared to all other genres, and that Head Speed was significantly lower for Latin ($M = -.148$, $SE = .074$) than for Techno ($M = .190$, $SE = .074$. Moreover, there were significant Extraversion × genre interactions for Global Movement, $F(5, 960) = 2.846$, $p = .015$, and Head Speed, $F(5, 960) = 3.914$, $p = .002$ (see Fig. 4). As regards Global Movement, high Extraverts had higher values compared to low Extraverts for Jazz (mean difference = .528), Latin (mean difference = .741), and Pop (mean difference = .424). For Head Speed, meanwhile, high Extraverts had considerably higher values for Rock (mean difference = 1.290), and moderately higher values for Techno (mean difference = .611), Funk (mean difference = .417), and Pop (mean difference = .428).

3.4. Agreeableness

Multivariate tests were significant for both personality, $V = .077$, $p = .000$, and genre, $V = .104$, $p = .000$, but not for the personality × genre interaction, $V = .016$, $p = .948$. Specifically, there was a significant main effect of Agreeableness on Global Movement, $F(1, 906) = 16.384$, $p = .000$, Hand Flux, $F(1, 906) = 47.369$, $p = .000$, and Hand Distance, $F(1, 960) = 11.275$, $p = .001$, with high scorers exhibiting higher levels of all three of these components (see Fig. 2, fourth row). In addition, there was a significant main effect of Genre on Local Movement, $F(5, 906) = 6.029$, $p = .000$, Global Movement, $F(5, 906) = 4.033$, $p = .001$, and Head Speed, $F(5, 906) = 8.554$, $p = .000$ (see Fig. 3, fourth row). Post Hoc comparisons with Bonferroni correction revealed that Local Movement was significantly higher for Techno ($M = .402$, $SE = .081$) compared to all other genres, that Global Movement was significantly higher for Latin ($M = .342$, $SE = .083$) compared to Funk ($M = -.010$, $SE = .076$) and Rock ($M = -.262$, $SE = .132$), that Head Speed was significantly higher for Rock ($M = .710$, $SE = .129$) compared to all other genres, and that Head Speed for Latin ($M = -.208$, $SE = .082$) was significantly lower than for both Techno ($M = .164$, $SE = .082$) and Funk ($M = .123$, $SE = .075$).

3.5. Neuroticism

Multivariate tests were significant for both personality, $V = .082$, $p = .000$, and genre, $V = .077$, $p = .000$, but not for the personality × genre interaction, $V = .006$, $p = 1.000$. Specifically, there was a significant main effect of Neuroticism on Local Movement, $F(1, 852) = 18.137$, $p = .000$, Global Movement, $F(1, 852) = 12.172$, $p = .001$, Hand Flux, $F(1, 852) = 40.706$, $p = .000$, Head Speed, $F(1, 852) = 6.253$, $p = .013$, and Hand Distance, $F(1, 852) = 6.580$, $p = .010$ (see Fig. 2, fifth row). High scorers exhibited higher levels of Local Movement, but lower levels of all other components. In addition, there was a significant main effect of Genre on Local Movement, $F(5, 852) = 3.707$, $p = .003$, Global Movement, $F(5, 852) = 2.806$, $p = .016$, and Head Speed, $F(5, 852) = 4.705$, $p = .000$ (see Fig. 3, fifth row). Post Hoc comparisons with Bonferroni correction revealed that Local Movement was significantly higher for Techno ($M = .084$, $SE = .074$) compared to Funk ($M = -.266$, $SE = .068$) and Pop ($M = -.321$, $SE = .074$), that Global Movement was significantly higher for Latin ($M = .097$, $SE = .071$) compared to Rock ($M = -.374$, $SE = .113$), and that Head Speed was significantly higher for Rock ($M = .395$, $SE = .112$) compared to all other genres.

4. Discussion

As regards personality, we found that Openness was positively related to Local Movement only, whereas Conscientiousness was positively related to Global Movement, Head Speed, and Hand Distance. Extraversion, meanwhile, was positively related to all five movement components, whereas Agreeableness was positively related to Global Movement, Hand Flux, and Hand Distance. Finally, Neuroticism was positively related to Local Movement, but negatively related to Global Movement, Hand Flux, Head Speed, and Hand Distance.

Several points are worth noting about the overall pattern of results for personality, as well as about similarities and differences in terms of movement features between the Big Five. First, each personality dimension was associated with a different pattern of movement characteristics in terms of: (i) significant movement components, (ii) the overall value for each significant component, and (iii) the magnitude of the differences between high and low-scorers for each dimension. Second, Extraversion and Neuroticism were the only dimensions which had a significant effect on all five components of movement. Moreover, although Extraversion was positively related to all five components, Neuroticism was negatively related to four of the five, the exception being Local Movement. Furthermore, Neuroticism was the only dimension to have a negative effect on any movement components. Finally, Openness appeared to have the weakest effect on movement characteristics, with only Local Movement being influenced by this dimension. Some similarities can be seen between these results and the findings of our previous study, as well as with typical characteristics associated with the Big Five dimensions, perhaps most clearly for Extraversion and Neuroticism.

In our previous work, Extraversion was associated with higher speed of movement of the head, hands, and centre of mass, as well as elevated levels of kinetic energy. In the present study, high Extroverts had the highest levels of Head Speed, as well as elevated levels of Local Movement (upon which centre of mass loaded) and Global Movement (upon which kinetic energy loaded). Moreover, this pattern of results seems consistent with typical behaviour exhibited by individuals scoring highly on Extraversion, who tend to be energetic, expressive of positive emotions, and looking for stimulation.

As regards Neuroticism, this dimension was associated in our previous work with jerky and accelerated movement, particularly of the head, hands, feet, and centre of mass. In the present study, all four of these features loaded onto Local Movement, the only component for which high Neurotics exhibited elevated levels. One might imagine that the combination of jerky movement of the head, hands, feet, and centre of mass along with the reduced levels of all other movement components was indicative of an individual expressing typical characteristics of high Neuroticism, i.e., anxious but depressed mood, especially given the added stress of the laboratory situation.

We might speculate that the effects of personality on music-induced movement were so clear for Extraversion and Neuroticism because these two dimensions are strongly connected to emotional expressivity (e.g., Gross & John, 1995). Music is known to evoke strong emotions in people (e.g., Juslin & Sloboda, 2001), and emotion can be expressed through bodily movement (e.g., Pollick, Paterson, Bruderlin, & Sanford, 2001), including music-related body movement (e.g., Luck, Toiviainen, & Thompson, 2010). It seems logical, therefore, that individuals scoring highly on these two particularly affective dimensions would more clearly express, via their body, typical emotional characteristics associated with them. Further research could help shed light on this matter.

In addition to personality, we found that several genres stood out as being particularly influential on the movement components. Rock, for example, was associated with the highest levels of Head Speed, particularly for those scoring highly on Extraversion, and the lowest levels of Global Movement. Techno, on the other hand, was associated with the highest levels of local movement, as well as elevated levels of Head Speed for individuals scoring highly on Extraversion. Latin, meanwhile, elicited the highest levels of Global Movement, especially for those scoring highly on Extraversion, but, along with Jazz, the lowest levels of Head Speed,

In real terms, this suggests that Rock music elicited stereotypical "head banging" movements, especially from individuals scoring highly on Extraversion, but little movement around the room. Techno music, meanwhile, seemed to elicit increased levels of rhythmic limb movement, as well as head movement for high Extraverts. Finally, Latin music made people move around the room the most, while keeping their head relatively still.

The musical excerpts used in this study were drawn from Western styles of music, and, although they differed in rhythmic complexity, all had a strong, periodic, rhythmic element. This was important since we wanted the music to induce movement in our listeners. We were also interested in bucking the trend in music cognition research of employing classical music, which most people among the general population do not listen to, as stimuli. Nonetheless, future work in this area might examine how other genres of music, such as classical or world music, influences listeners' spontaneous movements. Such music may not elicit the same kind of rhythmic dancing movements observed in the present study, but would help us better understand the effects of music on bodily movement.

Future work might also, for example, explore alternatives to the Big Five taxonomy of personality traits, as well as consider other individual difference factors, such as musical preference, music/dance training, or other movement-related experience such as sports training. Given that personality and individuals differences so fundamentally influence people's everyday cognitions, motivations, and behaviours, and that music is such a pervasive aspect of people's daily lives, the identification of relationships between these two factors and the way in which people move offers huge potential for further research.

Acknowledgments

We thank Jonna Vuoskoski for collecting the personality data, and Mikko Leimu for help and technical assistance during collection of movement data. This study was funded by the Academy of Finland (projects 118616, 125710).

References

Ball, G., & Breese, J. (2000). Relating personality and behaviour: Posture and gestures. In A. Paiva (Ed.), *Affective interactions: Lecture notes in computer science* (pp. 196–203). Berlin: Springer.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R. Professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Godøy, R. I., & Jensenius, A. R. (2009). Body movement in music information retrieval. In K. Hirata, G. Tzanetakis, & K. Yoshii (Eds.), *Procedings of ISMIR 2009*, pp. 45–50.

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

Gross, J. J. (1999). Emotion and emotion regulation. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality. Theory and research* (2nd ed., pp. 525–552). New York: Guilford.

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

Heeyoung, K., Kwak, S. S., & Myungsuk, K. (2008). Personality design of sociable robots by control of gesture design factors. In *Proceedings of the 17th international symposium on robot and human interactive communication* (pp. 494–499).

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.

Juslin, P. N., & Sloboda, J. A. (2001). *Music and emotion: Theory and research*. London: Oxford University Press.

Keller, P. E., & Rieger, M. (2009). Editorial: Special issue – musical movement and synchronization. *Music Perception, 26*, 397–400.

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

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

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

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

Luck, G., Saarikallio, S., & Toiviainen, P. (2009). Personality Traits Correlate With Characteristics of Music-Induced Movement. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, P.-S. Eerola (Eds.), *Proceedings of the 7th triennial conference of the European society for the cognitive sciences of music* (pp. 276–279). Jyväskylä, Finland: University of Jyväskylä.

Luck, G., Toiviainen, P., & Thompson, M. R. (2010). Perception of expression in conductors' gestures: A continuous response study. *Music Perception, 28*, 47–57.

Pollick, F., Paterson, H., Bruderlin, A., & Sanford, A. (2001). Perceiving affect from arm movement. *Cognition, 82*, 51–61.

Rentfrow, P. J., & Gosling, S. D. (2003). The Do Re Mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*, 1236–1256.

Rusting, C. L. (1998). Personality, mood, and cognitive processing of emotional information: Three conceptual frameworks. *Psychological Bulletin, 124*(2), 165–196.

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

Schmidt, E. (2002). Tension between invention and convention in jazz performance: Some effects on the listener. In Paper presented at *7th international conference on music perception and cognition*, Sydney, Australia.

Toiviainen, P., & Burger, B. (2010). *MoCap toolbox manual*. University of Jyväskylä: Jyväskylä, Finland. <www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual>.

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

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

Vuoskoski, J.K., & Eerola, T. (2010). *The role of mood and personality in the perception of emotions represented by music*, submitted for publication.

V

LEARNING AND SYNCHRONIZING DANCE MOVEMENTS IN SOUTH AFRICAN SONGS —CROSS-CULTURAL MOTION-CAPTURE STUDY

by

Tommi Himberg & Marc R. Thompson

Dance Research vol. 29(2), 305-328.

Learning and Synchronising Dance Movements in South African Songs – Cross-cultural Motion-capture Study

TOMMI HIMBERG AND MARC R. THOMPSON

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

Contact for corresponding author (TH): tommi.himberg@gmail.com

Music and dance are human universals. Understanding the communicative nature and the interpersonal dynamics of making music and dancing has a wide area of applications from academic to artistic, educational and therapeutic uses. Cross-cultural and embodied cognitive approaches are important, as they ensure a view across a spectrum of cultural practices and allow us to explore which aspects of cognitive performance are learned and how.

In this study, our aims were to use a case study to investigate possible cross-cultural differences in movement, especially corporeal representation of beat and metre; to study group entrainment and factors contributing to synchronisation accuracy. From earlier studies in various fields of behavioural and brain imaging research (perception and attention, music performance, action observation network in the brain etc.) we expected that experts would be more coherent and better entrained, or mutually synchronised to each other, but we were interested in the temporal dynamics of entrainment in a group and the details of these differences.

In our study, a choir from South Africa and a group of Finnish choir singers were brought together for a two-day workshop. Songs with choreographed dance movements from various cultures in southern Africa (e.g. Zulu, Sotho and Xhosa) were taught to Finnish participants, and a simple dance choreography was made for a Finnish song that was taught by the Finnish participants. Video, audio and movement data were recorded over a number of performances and practice sessions. Several participants were interviewed informally during the course of the workshop. In this study we analyse two recordings of performances, one of the African and the Finnish song-and-dance.

Dance Research 29.2 (2011): 305–328
Edinburgh University Press
DOI: 10.3366/drs.2011.0022

www.eupjournals.com/drs

As expected, the analysis showed differences in embodiment of rhythm and synchronisation between the novices and experts. The novices were very focused on footsteps and their whole body was entrained to just the beat-level of the metrical hierarchy. The experts, however, demonstrated entrainment to multiple metrical levels, with different parts of their bodies. Analysis of the temporal dynamics of the interpersonal relationships within each group revealed a process of continuous, mutual adaptation to achieve accurate entrainment. Investigation of group entrainment and individual deviations from the mean phase revealed roles of leaders and followers and illustrated differences in beat-by-beat synchrony and coordination of larger structures, including bars, phrases and the whole song.

The findings demonstrate how making music and dancing in a group is an enlightening example of joint action and that dynamics of interpersonal coordination can be studied in a relatively naturalistic setting. Mimicry, mirroring, shared intentions and intersubjectivity, with all their emotional consequences can be experienced, observed and to some degree manipulated and studied in these settings. Our rich data set of natural musical behaviour are intended to help to direct the stricter, experimental research designs on this matter, as well as inform especially multi-cultural educationalists about learning patterns of rhythmic dance movements.

1 INTRODUCTION

Choir singing is one of the most common musical activities in the world and popular across cultures, social groups and ages (Nettl, 2000). It is a social activity that promotes social cohesion, and has even been shown to have health benefits (Beck et al., 1999, Kreutz et al., 2004). Singing in a choir, and performing a dance while singing, recruits numerous cognitive resources just like any music or dance performance (Gabrielsson, 2003, Palmer & Meyer, 2000). It involves memory, voice control, production of the melody and rhythm of the music and the dance movements, perception of the music and movements of the others and coordinating and synchronising all movements and vocal outputs with the other performers according to strict aesthetic criteria (Janata & Grafton, 2003, Grafton, 2008, Sawyer, 2006, Rasch, 1988). There are high communication demands for such joint activities; not only in terms of the emotional character or other meaning that is to be conveyed by the performance, but also in terms of timing of movements and music, or the pitch and timbre of the music. These demands also include having to monitor one's own output as well as those of the others, and adapting them smoothly and in a co-operative manner (Repp & Keller, 2008), taking into account that the other performers are also aware of the errors in timing or tuning and are also making adjustments.

In this study, we investigate song and dance performances of a South African choir with novice Finnish participants. In this cross-cultural study, our focus is on investigating the embodiment of rhythm and metre and analysing entrainment, or the mutual synchronisation of the performers' body movements with the use of optical motion capture methodology (explained in detail below). We wanted to see how familiarity with the musical and movement material is reflected in synchronisation accuracy, and in the temporal dynamics of entrainment. To illustrate these aspects of the performance, we present motion capture data from two songs, one where both the music and dance are from the experts' culture, and another where the music is from the novices' culture, and movement from that of the experts.

In synchronisation studies, a distinction is made between rhythm and metre. Rhythm refers to the temporal sequence of notes, while meter is the cognitive temporal grid or hierarchy according to which the beats are organised, for example strong and weak ones (endnote/footnote 1). Synchronisation has typically been studied using a finger-tapping paradigm with individual participants in a lab (see Repp, 2005 for a review), but it has also been investigated in actual performances in music ensembles. For example, Rasch (1988) found that performers in a chamber music ensemble would synchronise their note onsets to within 30–50 milliseconds, and that the leading instrument would lead by an average of 7ms. Individual finger-tappers will reach similar levels of accuracy, although they often anticipate the metronome onsets they are synchronising with (Aschersleben & Prinz, 2005). These studies have provided ample information concerning the details of synchronisation, but relatively little understanding as to the purpose of this automatic and apparently (although debatably) uniquely human skill (Bispham, 2006, Patel et al., 2009). The studies have also neglected to investigate synchronization in real-world settings such as music and dance performances. Such settings involve complex interactions whereby two or more subjects must synchronize with one other. For such interactions, we prefer the term *entrainment* (Clayton, Sager & Will, 2004, Himberg, forthcoming).

Entrainment is an automatic (and hard to resist) process that is interesting not only in terms of its mechanisms and the minutiae of how the error correction processes work, but also in what it allows, what it facilitates. Entrainment, together with mechanisms of joint attention and shared intentionality, form the basis of complex social interactions and joint action in all communication domains, music and dance as well as language (Marsh et al., 2009, Knoblich & Sebanz, 2008, Overy 2003). The alignment of mental states, which is a result of the coordination of individual intentions (Tomasello & Carpenter, 2007) leads to and calls for shared representations, such as having the same, correctly aligned metrical and tonal schemata of the music being performed or movement pattern representations of the dance in question. This means that as performers have the same metrical schemata, they have the same idea of where the strong and weak beats should occur, for instance, and therefore are capable of a very synchronous and concerted performance.

Furthermore, our attention is temporally entrained to the periodicities in our environment (Jones & Boltz, 1989, Large & Jones, 1999), and in ensemble performance, directing the limited attentional resources is a complex affair (Keller, 1999), and a skill where learning, expertise and familiarity play a role. Recent studies have also discovered a link between entrainment and positive affect and affiliation (Hove & Risen, 2009). This suggests that entrainment links social and communication processes and emotional rewards. Thus, studying entrainment in social contexts such as in music and dance has a lot of potential from the point of view of cognitive sciences and neuropsychology.

The advantages of taking an embodied, rather than a disembodied approach to cognition when studying rhythm and meter should be self-evident (Iyer, 2002). As Grafton (2009) notes in his excellent review, there are many definitions of embodied cognition, and of the one he adopts he says that it refers to 'the existence of a memory system that encodes knowledge of a person's physical competencies and [how] a person is capable of interacting with the physical world' (Grafton, 2009, p. 97). Concretely, from a disembodied perspective, knowledge and understanding of different rhythms and different metres (duple metre as in a march, or a triple metre as in a waltz) is a representation somewhere in the mind. The embodied approach complements this view by saying that this representation also comprises the movement patterns these different metres afford. For example, in a previous motion capture study on embodied metre, Toiviainen, Luck & Thompson (2010) found that different levels of metrical hierarchy (see London, 2004, Lerdahl & Jackendoff, 1983) were represented in different types of movements. For example, they found that medio-lateral (side-to-side) movements were synchronized with the tactus level (the most salient beat in the music), while anterior-posterior (front-to-back) movements were synchronized to slower metrical levels (periods of two and four beats).

A key concept in embodied cognition is simulation. Our understanding of the goals and intentions behind other peoples' actions is at least to some degree based on simulating the observed actions in one's mind. This close connection between action planning, action execution and action observation has been a hot topic for cognitive neuroscience since the lucky discovery of mirror neurons in macaque monkey's brains in the late 1980's and early 1990's (di Pellegrino et al., 1992, Gallese et al., 1996, Rizzolatti et al., 1996). These premotor cortex neurons were active both when the monkeys observed an action as well as when they themselves performed it. Subsequent research has mapped this mirror neuron system (or action observation network) also in humans, and the close ties of action and observation have inspired, among other things, a host of music and dance-related studies in cognitive neuroscience (Calvo-Merino et al., 2005 and 2006, Cross, Hamilton & Grafton, 2006.)

As dancing in scanners is difficult to do (but not impossible, see Brown, Martinez and Parsons, 2006), this similarity of activation patterns and the network of brain regions that are involved in action and observation allows scientists to learn about the neural basis of the movement execution by scanning the brains of participants while they watch these movements being performed. Beatriz Calvo-Merino and her associates (Calvo-Merino et al., 2005, 2006) studied the action observation network by comparing brain activation when participants were watching performances of dance styles that they either had trained to perform or not. They compared expert dancers of capoeira and ballet (2005) and male and female ballet dancers (2006), and were able to show that the brain's response to seeing an action is influenced by the motor skills acquired by the observer. This central claim by the proponents of the embodied approach to cognition was corroborated by a study by Cross et al. (2006), where they taught participants sets of new movements, and showed that while observing these movement patterns, activation levels of the action observation network increased in accordance with levels of proficiency in executing these movements.

These increases in observed activation levels could be the result of developed motor programmes. The theory of motor programming states that with repetitive training, action sequences become automated and retrievable when stored in our memory. Such action sequences are hierarchical in nature and several of the sequence's parameters can be altered at will. In the case of music performance, this facility not only enables musicians to perform fast musical sequences, but also the ability to control the timing and dynamics of the music. The individual choices made by performers regarding these musical parameters are largely responsible for what can be perceived as *musical expressivity*. Motor programmes must also be adjusted when performing with other musicians. In the context of dance and the current study, adjustable motor programmes were needed to entrain with the steps of a partner in order to keep the performance 'unified' (Davidson & Correia, 2002, 244).

The recent introduction of motion capture systems to music and dance research has started to provide accounts of music-related body movement that are more objective and accurate than video observations. Optical motion capture is a technique whereby a subject's movements are recorded at a high temporal resolution using a system of calibrated image sensors. The set-up we used is a system of eight infrared cameras. The cameras emit fast infrared pulses in order to track the motions of retro-reflective passive markers attached to a subject. The cameras' overlapping projections are used to produce a three-dimensional (3D) model of the subject's position. For the current experiment, we collected motion capture positional data of the choir performing a South African as well as a Finnish song.

Although video can be analyzed computationally (Jensenius, Godøy & Wanderley, 2005; Castellano, Mortillaro, Camurri, Volpe & Scherer, 2008), motion capture data is three-dimensional, and has higher spatial and temporal precision (Wanderley, Vines, Middleton, McKay & Hatch, 2005; Eerola, Luck & Toiviainen, 2006; Luck & Toiviainen, 2006). Motion capture has its own limitations, for instance, analysing the vast quantity of data produced is complicated and requires specialist software, and the systems are still expensive and bulky, which is why video analysis is still the first choice for ethnomusicologists, for instance (Moran, forthcoming). But motion capture makes it possible to study dance- and music-related movement and embodied cognition in settings that are more natural than in the traditional lab experiments.

Based on studies from multiple sources, from attentional resources to embodied cognition, behavioural studies to brain imaging, the assumption is that novices would fare less well than the experts, would show weaker inter-group entrainment and larger deviations from the mean. Our challenge was to look at these differences in more detail, with the additional help of the song unfamiliar to our experts. The cross-cultural setting, to our knowledge unique for a motion capture study on music and dance, provides an interesting field for comparisons, as the expertise required for the performances that we study is different from the prior knowledge of the novices. Complex rhythms (at least compared to stereotypical 'Western' music) characterise many African music genres (Temperley, 2000; Agawu, 2003), and a previous experiment has indicated differences in rhythm perception between Finnish and South African participants (Toiviainen & Eerola, 2003).

This is a cross-cultural study, in that we contrast and compare two culturally different groups and their performance in tasks. There are dangers in categorising people with very general cultural/national labels and this gets especially dangerous if the authors or the readers assume these labels are the source of any cross-group differences. The members of the choir are a diverse group and represent a wide range of South African peoples, and different ethnic and linguistic backgrounds (while sharing a religious background in Lutheran Christianity and a part of their musical background in the choir). As a touring choir, they are however a very coherent group from the point of view of their music and dance performances. The Finnish participants are linguistically and culturally the more homogenous of the two groups. While we hypothesise that some differences that we found between these groups' performances are attributable to their respective cultural backgrounds, we are convinced that these differences, for example relating to the embodiment of metre and rhythm are learned, not innate differences. Thus we have decided to use the terms 'experts' and 'novices' in this study to highlight the likeliest source of the group-level differences in performance.

2 METHODS

2.1 Participants and materials

A cross-cultural choir workshop was organised at the University of Jyväskylä, in Finland. The South African participants were members of Emmanuel Lutheran Church Choir from Johannesburg. The Finnish participants were recruited from the extended community of the Music department of the University of Jyväskylä.

The Emmanuel Choir performs traditional African music, originating mainly from the various southern African cultures (e.g. Zulu, Sotho and Xhosa). Coordinated movement choreography is part and parcel of these songs, therefore making them especially suitable for a music and movement study. Most of the South African participants were actively performing professional and semi-professional musicians, familiar with South African folk music as well as Western popular and classical music, but not with Finnish or other European folk music. The Finnish participants also had musical experience and musical training, mostly in Western music and choir singing. Some of the Finnish participants had limited experience of African music or general knowledge of the musical style, but none was familiar with any of the African songs performed in the workshop, nor choir choreographies in general. Throughout this paper, in all the figures and analyses, we use the term 'experts' for the South African participants and the term 'novices' for the Finnish participants, regardless of the song or individual skill level. It could be argued that the Finns are experts in the Finnish song, but as the choreography and movement style in that song was South African, this is not so clear from the point of view of movement analysis. Also, we did not want to use the terms Finnish and South African, as we felt they might be loaded and too broad. Also, we hypothesised that any differences between the groups are likely to be due to different levels of expertise rather than nationality. Furthermore, we believe that the details of these differences are interesting, not the differences themselves.

The workshop was a two-day event, and consisted of half-a-dozen African songs and one Finnish song being learned and performed by the participants. The workshop was conducted in English.

Two songs were chosen for this study. The South African song was Fiela Ngwanana, or Sweeping Girl, a Sotho song (video example 1). This was a song from the performance repertoire of the choir, and had an accompanying dance. The Finnish song chosen was Kukkuva kello, or Cuckoo Clock, a folk song that is often sung as a canon (video example 2). The South African participants prepared a very simple choreography for the Finnish song, so that both songs would have music and movement. Although the Finnish folk music style was not familiar for the South African participants, the movement style used in the choreography of the Finnish song was, and thus this setting was not perfectly balanced, from the point of view of cross-cultural experimentation (Berry, Poortinga & Pandey, 1997).

Video 1. Performance of Fiela Ngwanana, the Sotho song. Video and animation rendered from motion capture data.

Video 2. Performance of Kukkuva Kello, the Finnish song. Video and animation rendered from motion capture data.

2.2 Apparatus and procedure

Video and audio recordings of the workshop were made. Movement data from selected sessions of the workshop were captured using an 8-camera Qualisys Motion Capture system. Reflective markers were attached to the participants, and the cameras recorded the trajectory of these markers at a rate of 120 fps. For the African song, 8 participants (4 expert, 4 novice) wore markers (head + feet) and were positioned at the centre of the room in a 2 by 4 –formation. In the Finnish song, there were 10 participants, 3 of whom were novices, and on this occasion they wore only head markers. During both performances, the rest of the workshop participants also contributed to the music making and dancing but stayed outside the capture volume. The discrepancy between marker placement and number of participants in each performance is mainly a result of the experiment's ecological setting. The Finnish song's performance featured more participants. With more participants in the capture volume, we were concerned that markers on the feet would not be visible to the motion capture cameras, resulting in poor motion capture data. It was decided not to include feet markers for this performance. However, we are confident that the difference in marker placements between both conditions does not affect the results we are reporting.

The songs were learned in the group, aurally and by imitation, with the lyrics either pinned to the wall (African song) or printed on the sheet (Finnish song). No notation was used, and the choreography of the dance was also taught solely by imitation. A simple step pattern in the style of the choir's other repertoire was created for the Finnish song. After a few practices, a performance of the song was recorded using the motion capture system.

2.3 Analysis

The movement data acquired from the motion capture system was imported into MATLAB as time series data representing the location of each marker on the three dimensions of the Cartesian coordinate system. The computation of kinematic variables from the movement data (velocity, acceleration etc.), statistical analysis and visualisations were performed using the MoCap Toolbox (Toiviainen & Burger, 2010) and other MATLAB functions.

We analysed the data in the African song in more detail than the data in the Finnish song – partly because the marker setup in the African song (head and feet markers) allowed for more analyses than the one in the Finnish song (head markers only), and partly because our main interest was in the situation where true novices would learn a musical and movement style they had not encountered before. As our focus was in timing, metre, and entrainment, we performed a number of different synchronisation analyses. The aim of these analyses was to measure the period and phase of the movements of the individuals and to visualise and quantify the coherence of pairs and groups, and to see what kinds of differences there were in experts and novices, and whether the two songs were different.

First, we compared the two groups by means of pairwise cross-correlation analyses. This allowed us to look at the coherence within and between groups and between all participants. These cross-correlation analyses were carried out using the vertical acceleration data of the feet markers in the African song, and the vertical and horizontal acceleration of the head marker in the Finnish song.

In order to investigate the temporal development of synchronisation, we conducted windowed cross-correlation analyses. We focused on the vertical movement of the head markers, as the head markers have been found to represent the overall periodicity of body movement very well (Toiviainen, Luck & Thompson, 2010).

In addition to the pairwise analyses, we conducted analyses of synchronicity in the whole group. To this end, the 'raw' data representing the position of the head markers needed to be filtered and then transformed to phase angles. A model was then applied to the phase data to calculate indicators of sychronisation accuracy and coherence for the whole group.

The first step in this process is to band-pass filter the movement data to filter out less important frequencies from the data and focus the analysis to the range of periodicities that were considered most crucial. A Gaussian filter was used, with a center frequency at the period of the tactus of the accompanying music. The tactus represents the most salient metrical level of the music, the beat that a listener is most likely to lock on when tapping a foot along to that piece of music. This selection decision was supported by the choreographies, as well as the music of these two pieces. Both choreographies used step patterns that were locked on to these periods, or in other words, they had steps to be performed at these rates.

This filtering allows the immediate phase of the vertical movement to be calculated unambiguously, using a procedure called Hilbert transformation (Khvedelidze, 2002). Once we had calculated the phase of the vertical movements, a model for measuring synchronisation in oscillators developed by K. Kuramoto (Acebron et al., 2005) was used, and an indicator of coherence of the oscillators called the 'order parameter' was calculated. This order parameter is an index measure that ranges from 0 to 1. The maximum score of 1 refers to perfect synchrony, while 0 indicates unordered, random relationships between the oscillators (or in our case, dancers). Based on this analysis, we can also calculate how each individual deviates from the group phase, and can plot the evolution of either the group order parameter or the individual deviations. This set of analyses can, in our opinion, offer a much fuller picture of dyadic and group synchronisation than previous methods. One specific benefit of this analysis is that visualising the temporal evolution of synchrony might help in identifying leaders and followers or other roles from within the groups.

3 RESULTS

3.1 Pairwise analyses – synchronisation in groups

Cross-correlations were calculated on the vertical acceleration of the foot markers in the South African song and on the vertical and horizontal acceleration of the head markers in the Finnish song. In cross-correlation analysis, the data time-series from one performer is correlated with the time-series of another performer, to see how well they match. To investigate whether one time series follows the other with a delay (rather than changing in perfect synchrony), these correlations are also calculated after shifting one of the time series by a small delay or lag. We calculated these cross-correlations up to a lag of 0.5 seconds in both directions. We were interested in these small, sub-beat deviations;

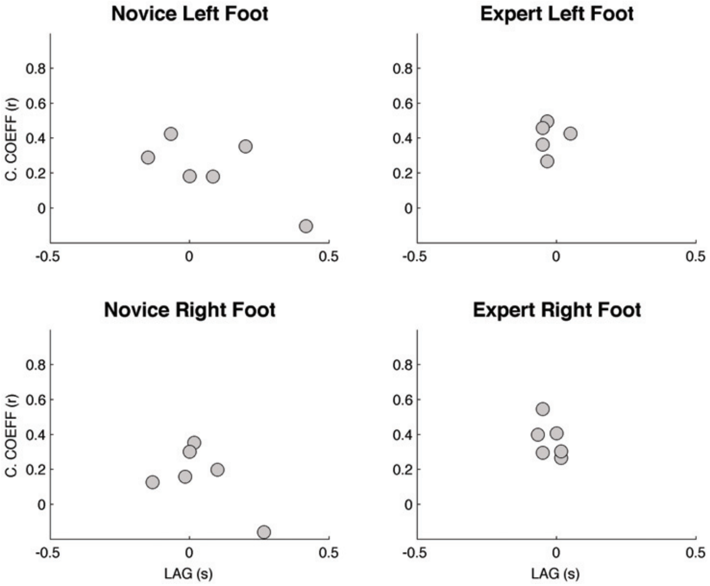


Fig. 1. Pair-wise cross-correlations for foot acceleration in the vertical dimension, displayed by group. The circles represent the maximum correlation coefficients (*r*) between the six possible pairs within each group. The x-axis shows the lag or synchronization error between pairs and the y-axis shows the strength of the correlation between pairs.

cross-correlations at lags longer than the beat interval tend to repeat, as for instance the peaks of acceleration of two different beats line up. Figures 1, 2 and 3 indicate the maximum coefficient (*r)* from the cross-correlation functions and the lags at which the coefficient occurs. Group differences are presented in Figure 1 while individual differences are presented in Figures 2 and 3. The more similar the two movement patterns are, the higher the coefficient—perfect similarity would yield a correlation coefficient of 1. The lag indicates the average temporal relationship between performers, their synchronisation error. For example, a larger lag at which the maximum coefficient occurs indicates a larger phase error between the two participants.

For Figure 1, the circles in each of the subplots represent the maximum correlation coefficient (*r)* between the six possible pairs within each group (the South African song was performed by four novices and four experts). Note that we are not considering which pair in particular was more synchronized; we are rather concerned with the synchronization as a group in general. For this reason, the individual circles in Figure 1 are not identified.

Regarding the South African song, the first issue to notice in Figure 1 is that the correlation coefficients (on the y-axes) are relatively low. This is mainly due to using unfiltered data in the analysis, and the fact that although the choreography dictates the steps that the dancers should take, there is considerable freedom in actually implementing those steps, as can be seen in the video example 1.

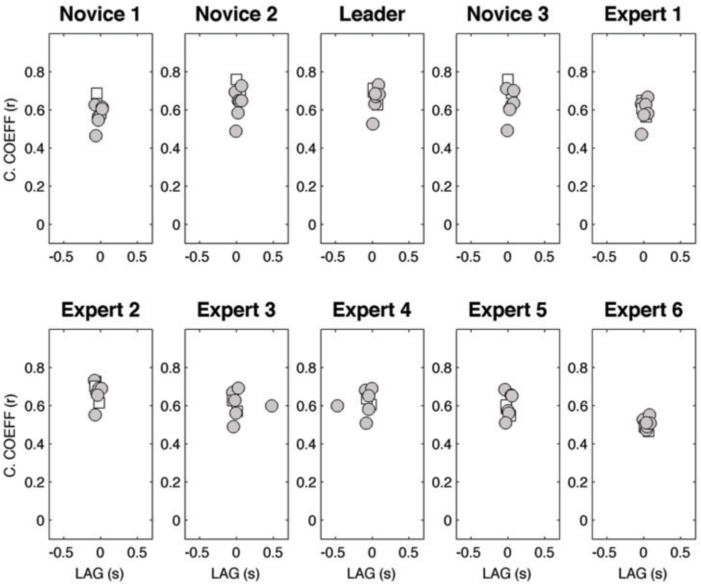


Fig. 2. Pair-wise cross-correlations for head acceleration in the vertical dimension. Each subplot shows the maximum correlation coefficient between an individual and each of the other performers. Shaded circles indicate correlations with experts and white squares indicate correlations with novices. The x-axis shows the lag or synchronization error between pairs and the y-axis shows the strength of the correlation between pairs. The subplot entitled 'Leader' is for the expert participant acting as leader of the performance.

It can clearly be seen that the experts are more coherent than the novices. For both feet, the lags at which the maximum pairwise cross-correlations occur are closer to zero (on the x-axes). Also the coefficients are slightly higher than for the novices.

Figures 2 and 3 show the data for cross-correlations between individuals for the Finnish song. The subplots in these figures show the maximum correlation coefficient plotted against its lag, for each individual participant. For example, the top left subplot in Figure 2 shows coefficients between Novice 1 and all other performers. The shaded circles represent cross-correlations with expert performers while the white squares represent cross-correlations with other novice performers. Figure 2 shows the maximum coefficients for vertical head acceleration while Figure 3 shows the maximum coefficients for horizontal (left-to-right) head acceleration. Overall, the correlation coefficients are much higher in the vertical dimension. This reflects the choreography that was based on side-to-side movement, which caused a greater chance of the participants to be out of phase horizontally. The vertical dimension captures the beat-level movement of the steps, which was very similar across individuals.

Plotting data from both vertical and horizontal dimensions exposes movement properties which are not visible from the video recording. One interesting point is Novice 1, who according to the video example, is constantly

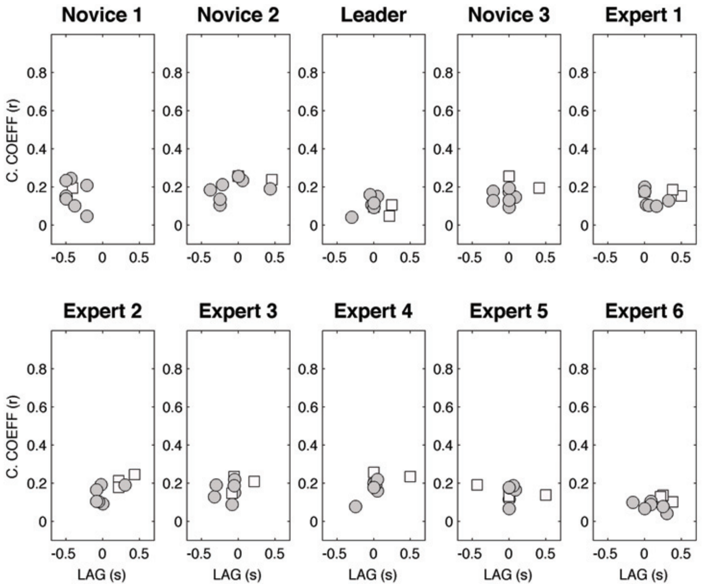


Fig. 3. Pair-wise cross-correlations for head acceleration in the horizontal dimension. Each subplot shows the maximum correlation coefficient between an individual and each of the other performers. Shaded circles indicate correlations with experts and white squares indicate correlations with novices. The x-axis shows the lag or synchronization error between pairs and the y-axis shows the strength of the correlation between pairs. The subplot entitled Leader is for the expert participant acting as leader of the performance.

behind the others in her movements. Looking at the vertical dimension (Figure 2, up-down movement of the head marker), she is actually in time with the others, not standing out at all, while some expert dancers at the back row are more out of sync than her. However, the horizontal dimension (Figure 3, left-to-right movement of the head marker) reveals that in her movements along this dimension, she is clearly late compared to everyone else. And conversely, the experts in the back row are in time with the others, in terms of their sideways movement.

3.2 Windowed analysis – examples of pairs

The relationship between people in the group or in any given pair is not fixed or permanently locked. The previous analysis looked at a section of the trial as a whole, and gives a score based on the global average performance. In order to investigate the temporal development of these interpersonal relationships, we conducted windowed cross-correlation analyses of one novice-expert pair in both songs. This method visualises the development of the relationship over the course of the trial, and might uncover structures that the global score does not show. The first expert-novice pair from the front row (closest to the video camera) was chosen for the analysis.

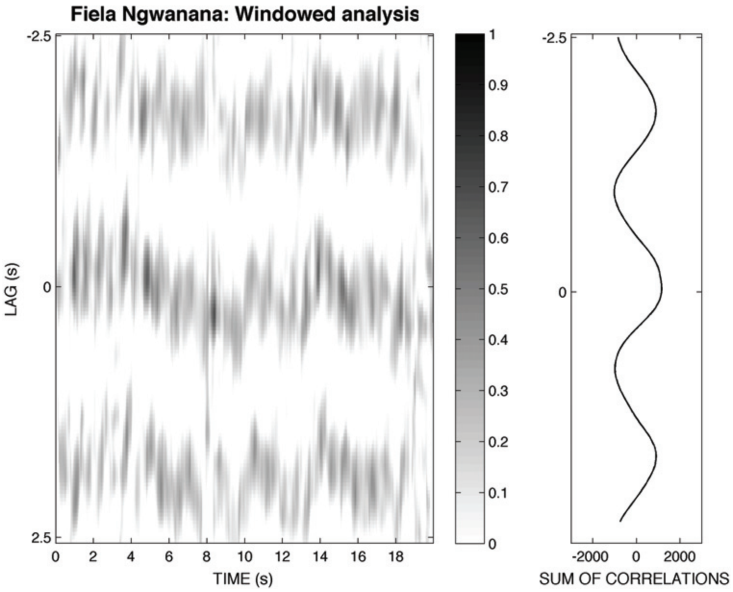


Fig. 4. Cross-correlogram and sums of correlation values for each lag, vertical position of head for one pair of expert and novice performers in the African song. Darker shades indicate higher correlations.

In these cross-correlograms (Figures 4 & 5), the shading represents correlation coefficients, with darker shades indicating higher correlations. The vertical axis represents lag, and the horizontal axis time. Tracking the darker shades and the lags at which they occur, gives an idea of how the relationship evolves, which participant introduces changes (or small mistakes) and how the other one follows. Whenever the darker shade occurs below the zero line, it indicates the expert is leading, and when it occurs above the zero-line, the novice is leading.

In both songs (Figures 4 & 5), the sum curves show a peak very close to lag zero, which means the two participants are entrained, but the windowing reveals that the peak correlation oscillates around lag zero, from one participant to another, and there does not seem to be a clear, consistent leader. In addition to the zero lag correlations, there are 'bands' of correlations, 'harmonics' at plus and minus half a beat. This is a reflection of the strongly periodic nature of the movement. In the Finnish song the rhythm is un-syncopated, with the rhythmic and metrical accents aligning perfectly, amplifying this effect.

3.3 Phase analysis – group synchronisation

By analysing the phase of vertical movements of the head markers, a more detailed analysis of group synchronisation can be conducted. Plotting the order parameter, a unitless phase coherence score between 0 (unordered) and 1 (perfect simultaneity), shows how group coherence develops over time.

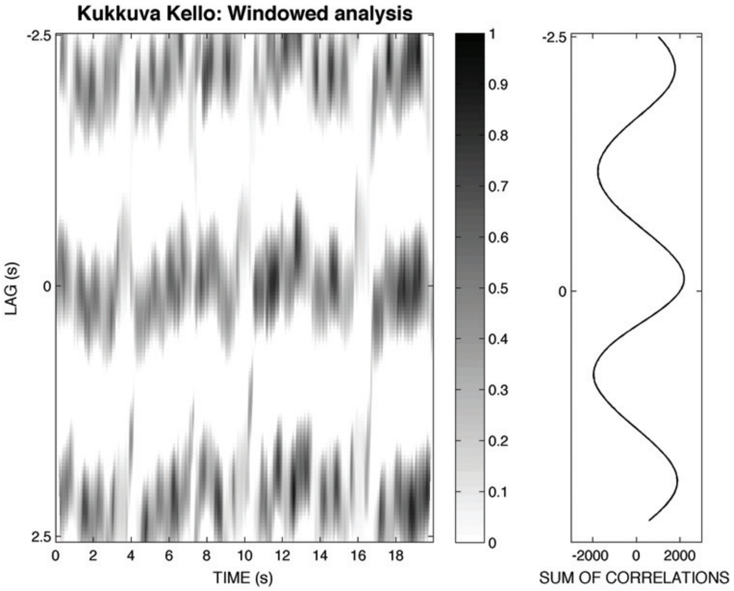


Fig. 5. Cross-correlogram and sums of correlation values for each lag, vertical position of head for one pair of expert and novice performers in the Finnish song. Darker shades indicate higher correlations.

In the Finnish song (Figure 6) the order parameter is very low and erratic at the beginning and in the end, where the choir is standing still. There is minimum amount of movement at this stage, mainly just fidgeting and some postural sway, and no choreography. The order parameter rises to a steady level once the movement starts. As the data is band-pass filtered using the tactus of the music as the center frequency, this order parameter refers to the synchronisation accuracy at around this periodicity. For consistency, the Finnish group is labelled 'novices' in this plot, although they know the music of this song/dance better than the other group. Figure 6 shows that the novices seem to be the better-coordinated group, apart from occasional dips. This is not only because it is the smaller group and therefore the more easily coordinated (there are three novices and seven experts), but also the novice group was more coherent even when controlling for the group size and comparing the three novices against three randomly picked experts.

Rather, here the Finnish participants are experts of the song, and their novice status in the movements was perhaps less evident due to the simplicity of the choreography. Meanwhile the South African experts needed a lot more cognitive effort to remember the melody, and especially remembering or reading the lyrics and pronouncing them. This would have left fewer attentional resources to spare for coordination or musical expression.

In addition to the beat level (as depicted in Figure 6), the body movements can be coordinated in other periodicities, as well. For instance, the movement

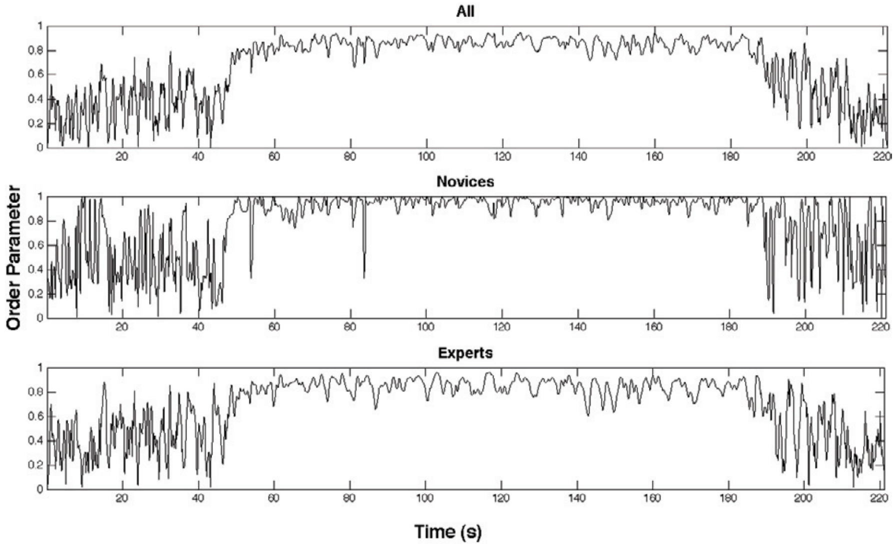


Fig. 6. Order parameter for the whole choir, the novices and the experts, for the Finnish song.

pattern is a repeating side-to-side movement, which has a periodicity of about 3.5 seconds. Looking at the horizontal dimension in the direction of the movement after filtering the data using a band-pass filter centered at this periodicity, we obtain a very similar curve.

In the South African song (Figure 7), the order parameter seems to mark the boundaries of different sections in the choreography. The order parameter dips to a low point when the movement pattern changes. Looking at the novice and expert groups separately, we can see that the expert group is more consistent across these movement pattern changes, and it is the novices who contribute the dips in the overall order parameter. This probably reflects the fact that experts do not have problems in remembering what the next pattern is, while at least some of the novices might pause or hesitate before moving to the next pattern. It can also be that greater expertise involves knowledge about how to perform these transitions smoothly, not just what the sequence of different patterns is. The novices had very little practice and, based on our observations, the focus of the learning process was on learning the repeating patterns in the correct order, not on the transitions.

This overall coordination is of course the contribution of the individuals. The group phase is the mean of the individuals' phases, and the deviation of the individuals from the group mean can be calculated by a simple subtraction. This average deviation of the individual from the group mean can be used to assess their roles in the group entrainment, as in Figure 8.

In the South African song, the eight participants are very close to each other (the scale on the y-axis is in radians) apart from one novice participant in the

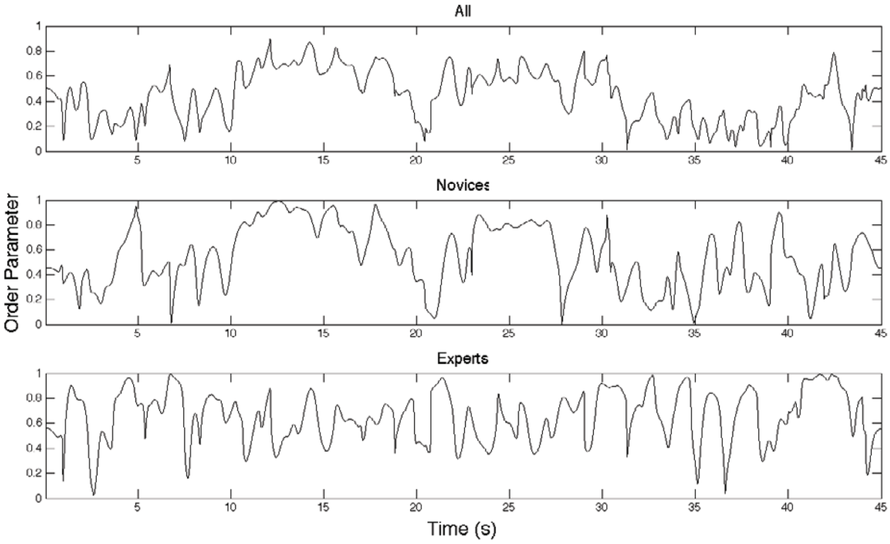


Fig. 7. Order parameter for the whole choir, the novices and the experts, for the African song.

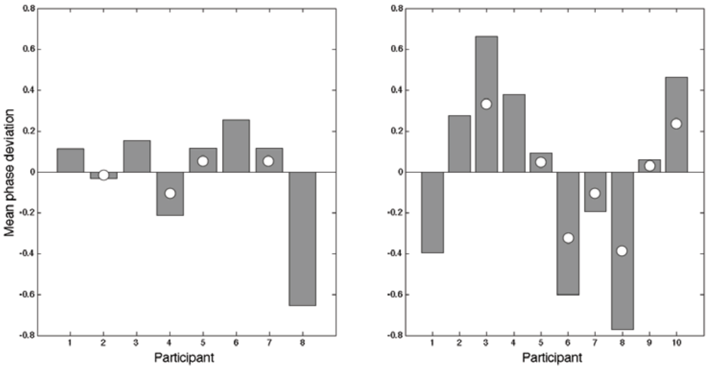


Fig. 8. Deviations from the mean phase for each of the participants. African song in the left panel, Finnish song in the right panel. White dots indicate experts.

second row who is late in relation to everyone else. In the Finnish song, the deviations are larger in general. It is noteworthy that participant 3 in this trial had the role of a leader. She was standing in the centre of the front row, and was given the responsibility of leading the dance. According to the average phase deviation measure, she is also ahead of the others.

3.4 Metrical levels

In a study of spontaneous movement to music, Toiviainen, Luck and Thompson (2010) found that different parts of the body were used to exhibit different levels

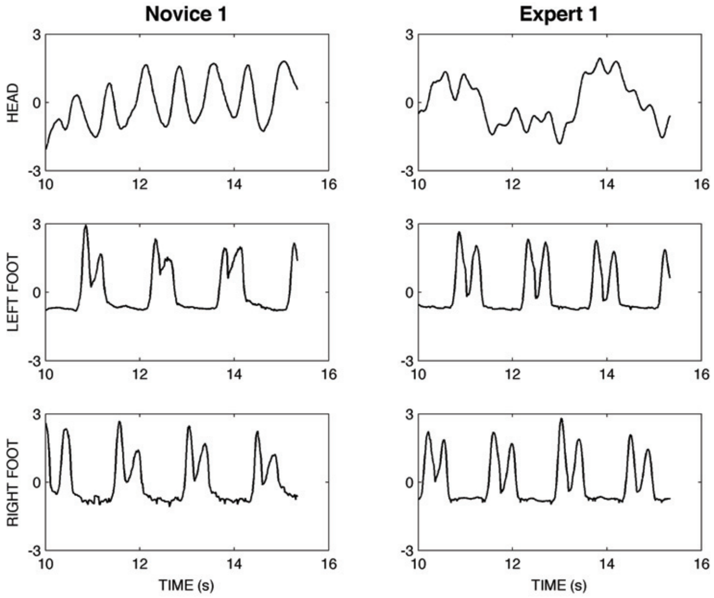


Fig. 9. Vertical head and foot acceleration, comparison of the expert and novice performers (the pair in the right of front row) in one section of the African song. For the purposes of presentation, the data was z-scored and the y-axis shows standardized values.

of metrical hierarchy. We found supporting evidence for this. Figure 9 depicts the vertical foot and head trajectories of a five-second section of the South African song. For the novice in this example, the foot movements are immediately reflected in the vertical movement of the head. For the expert, however, the trajectories of the feet and the head look much more different from each other. The head trajectory seems to capture not only the step-level periodicity, but also a slower movement, a vertical pattern that has a period that is four times longer than the step-level pulse. The vertical head trajectory seems like a summation of two sine curves, one that captures the step-to-step movement that is locked to the beat of the music, and another that embodies the slower, bar-level rhythmicity of the music.

Thus the head marker of the expert seems able to capture multiple metrical hierarchies at once. This could be a reflection of the richer metrical, embodied schemata of the experts, compared to the simpler beat-based embodiment of the novices. Here higher expertise might allow more cognitive resources (attentional and otherwise) to be spent on artistic expression, where focusing on the core of the body as the 'source' of the beat is important. Novices have to focus on getting the steps right and thus their imagined origin of movement is in the feet, as well, with the rest of the body echoing that metrical interpretation in a relatively stiff manner.

4 DISCUSSION

We recorded movement data in a choir workshop, where members of a South African choir taught a group of Finnish novices songs and dances, and to reciprocate, the Finnish participants taught one song, which the South African members of the choir choreographed in their style. We wanted to study group entrainment and embodiment of rhythm and metre in the choir, and make comparisons between experts and novices.

Our data indicated that the novices exhibited mainly just one metrical level in their movement – the tactus or the most salient beat. The experts, in contrast, had multiple levels of the metrical hierarchy at once. This was illustrated most clearly in comparing the foot and head marker vertical position data in the South African song. The novices were very focused on their foot movements, getting the 'steps' of the dance right. This results in a very different, stiffer style of movement than the flexible and soft movement of the experts. Many of the experts described that in the movement it is important to feel the beat, rather than stomp your feet on it. In order to help achieve this, they imagined the source of their movement in the core of the body, not in the feet. This helps in decoupling the movements of the feet and the head, and allows for a richer spectrum of metrical aspects to be embodied. Learning this should be investigated in a further study. Anecdotal evidence from this project suggests that with practice, novices were able to improve in this regard. Similar results have been found in another study of African dance (Himberg & Thompson, 2010), but studying the learning processes involved in embodying metrical information, a dedicated experiment is needed. For example, this requires a full-body marker setup, as the multiple metrical hierarchies are often represented in the movement of different body parts (Toiviainen, Luck & Thompson, 2010).

We showed that there were differences in group entrainment between novices and experts, the experts being the more coherent group in their own song. The effect of familiarity seems strong as this advantage disappeared or even reversed when the song was not familiar to the experts. This was a predictable result, as theories about attentional resources (Keller, 1999) as well as the ones about motor programs (Davidson & Correia, 2002) and performance studies (Palmer, 1989, 1997; Wanderley et al., 2005) (and might we add, common sense), would suggest this to be the case. Also, the evidence from brain imaging studies of dancers (Cross et al., 2006; Calvo-Merino et al., 2005, 2006) suggests that activation in the action observation network increases as the observer gets more physically familiar with the observed movements. This increased activation reflects the deepening understanding of the intentions behind those movements and the strengthening motor programs used to execute the motor sequences (Palmer, 1989; Wanderley et al., 2005).

The details are less obvious. First, we showed with visualisations that synchrony in the trial is a result of continuing small adaptations that all participants make. This result is in line with previous results from tapping studies (Repp & Keller, 2008; Pecenka & Keller, 2010; Konvalinka et al., 2009; Himberg, forthcoming).

A more detailed phase analysis revealed that the novices were actually better coordinated in the Finnish song than the experts, when we look at vertical movement of their heads. This movement corresponds to the beat-to-beat and step-to-step synchrony. But, looking at the horizontal movement of the head, this no longer is the case, as one of the novices had trouble synchronising with others in the more global, larger pattern of movement. Her steps fall on the beats, but she is out of phase in relation to the rest of the performers in the side-to-side movement that repeats slower than the beat of the music and corresponds to the bars of music (sets of 4 beats).

Even shared intention to perform a piece together does not guarantee success in said performance, as different members of the group, coming from different cultures, might have different schemata, or representations of metre, movement programmes and so on. In this study, the South African experts would have developed better-fitting schemata, through exposure and training, having learned the cultural conventions of that style. This shared knowledge would benefit their in-group cooperation. Conversely, the novices would have schemata for their own music and dance, but these might not entirely match the South African traditional styles. This potential lack of fit could have an effect on how the movements are performed, as well as the accuracy of their timing. One reason for the lack of congruence might be the differences between rhythmic and metrical concepts as between the Finnish and South African participants, which were demonstrated in a rhythm perception and production experiment by Toiviainen and Eerola (2003). In this study, it was shown that when tapping along to European and African folk melodies, the two groups often agreed upon the period, but had differences in the phase of their tapping, so that the Finns tapped on the note onsets, while the South Africans tapped slightly after, suggesting that their understanding of the same melodies were much more syncopated.

In the current study, the role of familiarity, but also the patterns of learning, were investigated. The novices tried to learn the South African styles, their songs and dances. In addition to the links between observation and doing, Grafton & Cross (2008) showed that there are strong links between learning and doing and learning and observing. They found that the same areas of the action observation network activate when learning by observing and when learning physically. These results not only lend support to ideas that early exposure to dance will help in learning dancing later, but are interesting in relation to our study, where the participants were learning the movement sequences by a combination of observing and reproducing the movements. This has been shown to be the quickest way to learn new dance movement sequences (Grafton & Cross, 2008, p. 62 and their references). Learning in a group might have facilitated things further, as observing someone else learning a complex skill has been shown to facilitate learning that skill oneself (Mattar & Gribble, 2005).

This experience of learning-by-doing and learning-by-imitation was relatively unfamiliar to some of the more classically-trained Finnish participants, as in their own genre, notation serves as a mediator between the musical ideas and the learner. Also, in their own culture music and dance are considered separate arts and skills; and having to combine the two, in an unfamiliar genre made this task even more challenging, and the feat of learning to perform half a dozen songs in just two days even more impressive.

Identifying the aspects of performance that were experienced as the most difficult might help in developing pedagogical approaches that can be used in multi-cultural music education (Volk, 1993). One of these aspects was the ‘source of movement’ in the core rather than in the feet that was discussed above. Another weak spot seems to be the transitions from one movement pattern, or one musical section to another. This is of course partly simply a memory issue, but the seamless performance of the experts is in contrast with the jagged, sectioned performance of the novices. The experts were better able to direct their attention to the overall structure of the piece and the unified character of the whole, while the novices were understandably struggling to manage from one movement pattern and musical sequence to the next.

Dancing and singing together is joint action par excellence – one participant truly is dependent on the other’s actions. This can be visualised in the cross-correlogram – entrainment is achieved by the way of constant, mutual adaptation. This is psychologically different from synchronisation, where an individual synchronises their movements with a non-responsive metronome, for example. The inter-dependence of entrainment also has important emotional consequences, as it has been shown that entrainment increases affiliation (Hove & Risen, 2009) and as such might be an important factor in the benefits of music and dance.

Joint action is often studied using artificial, meaningless tasks and strictly controlled laboratory environments (e.g. Sebanz, Knoblich & Prinz, 2003). Using actual music and dance performances as a starting point for these studies would improve their ecological validity. The richness of the schemata and representations, musical and movement structure and a virtually endless range of genres and styles, provide a field of study that would encourage the generation of suitably organised experiments to address joint action, imitation, action simulation, observation and the related neural processes. Motion capture data can be visualised and animated (as in our video examples), and the data itself can be manipulated in many ways. Using this kind of data as experimental stimulus in brain imaging studies would allow the features of the action observation network to be studied in more detail than before.

This study presented data from just one case, which of course needs to be taken into account when evaluating the results. This two-day workshop of singing and dancing provided us with the opportunity to collect some data from a fascinating cross-cultural event, but it was not a pre-planned, controlled experiment. The two songs analysed here had for example slightly different marker setups, different numbers of performers, and a different placement of lyrics sheets (on the wall / in hand). These differences make comparing performances across the two songs more difficult.

Comparing the groups is not perfectly easy, either. While the Finnish participants did not seem to benefit much from their advantage of having some general knowledge of African music (while the South African group had no prior knowledge of Finnish music), in this study the concept of 'familiarity' is confounded with cultural background. Breaking 'familiarity' down to its component parts would be an interesting aim for a further study. The concepts of rhythm and metre differ somewhat between these cultures, as mentioned above. Also, dancing while singing is the usual state of affairs in many African cultures, but these two skills are rather separate for instance in Finland. Separating the contribution of these factors (and potentially others) might be achieved in a controlled experiment. Carrying out cross-cultural brain research using designs similar to those of Calvo-Merino et al. (2005, 2006) and Cross et al. (2006) might shed light to which neural networks and brain systems work similarly and where the possible differences lie.

However, even with all the limitations of the current data set, we are confident that the data collection and analysis methodology that we presented would be very useful if applied to more controlled studies. These methods could be used to study factors that we could not address here. These would include studying the effects of personality on group synchronisation, investigating leader/follower roles within these groups or, for example, to address issues discussed in joint action studies, and investigate differences in situations where the intention is to collaborate or to compete with each other.

Author Notes: We thank professor Jukka Louhivuori for organising this workshop, professor Petri Toiviainen for guidance and help with the data analysis, and all our Finnish and South African participants for their time, flexibility and insights. We would also like to thank the editors and two anonymous reviewers of this manuscript for their excellent suggestions and corrections that made this a much better paper. This work was supported by the Academy of Finland (Project Number 131356). Finally, both authors contributed equally to performing the study and the preparation of the manuscript.

NOTES

1. Justin London defines the difference as follows: 'Rhythm involves patterns of duration that are phenomenally present in the music, and these patterns are often referred to as *rhythmic groups*. [–] By contrast, metre involves our initial perception as well as subsequent anticipation of a series of beats that we abstract from the rhythmic surface of the music as it unfolds in time. In psychological terms, rhythm involves the structure of the temporal stimulus, while meter involves our perception and cognition of such stimuli.' (London, 2004, p. 4) Vijay Iyer argues that from an embodied perspective, rhythm is related to human motion, and metre to the regularity of human motion (Iyer, 2002, p. 394).

REFERENCES

Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F. & Spigler, R. (2005). Kuramoto model: a simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77, 137–185.

Agawu, K. (2003). *Rhythmic topoi in West Africa*. Wort Lectures, Faculty of Music, University of Cambridge.

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: the role of sensory information. *Perception & Psychophysics*, 57(3), 305–317.

Beck, R. J., Cesario, T. C., Yousefi, A., & Enamoto, H. (1999). Choral singing, performance perception, and immune system changes in salivary immunoglobulin A and cortisol. *Music Perception*, 18, 87–106.

Berry, J.W., Poortinga, Y.H. and Pandey, J. (1997). *Handbook of Cross-cultural Psychology: Theory and method*. Boston, MA: Allyn & Bacon.

Bispham, J. (2006). Rhythm in music: What is it? who has it? and why? *Music Perception*, 24(2), 125–134.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). The neural basis of human dance. *Cerebral Cortex*, 16, 1157–1167.

Calvo-Merino, B., Glaser, D.E., Grezes, J., Passingham, R.E., & Haggard, P., (2005). Action observation and acquired motor skills: an fmri study with expert dancers. *Cerebral Cortex* 15 (8), 1243–1249.

Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R.E. and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, 16, 1905–1910.

Castellano, G., Mortillaro, M., Camurri, A., Volpe, G. & Scherer, K. (2008). Automated analysis of body movement in emotionally expressive piano performances. *Music Perception* 26(2), 103–120.

Clayton, M., Sager, R., & Will, U. (2004). In Time With The Music: The concept of entrainment and its significance for ethnomusicology. *ESEM CounterPoint*, 1, 1–45.

Cross, E. S., Hamilton, A. F. D. C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage*, 31(3), 1257–1267.

Davidson, J. W. & Correia, J. S. (2002). Body Movement. In Parncutt, R. & McPherson, G. (eds.). The Science and Psychology of Music Performance. New York, NY: Oxford University Press, 237–252.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental brain research*, 91(1), 176–180.

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

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music*, 31(3), 221–272.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2), 593.

Grafton, S. T. (2009). Embodied Cognition and the Simulation of Action to Understand Others. *Annals Of The New York Academy Of Sciences*, 1156, 97–117.

Grafton, S., & Cross, M. (2008). Dance and the brain. In C. Asbury & B. Rich (Eds.), *Learning, the arts, and the brain: The Dana Consortium Arts and Cognition report* (pp. 61–70). New York: Dana Foundation.

Himberg, T. (forthcoming). *Cognitive Foundations of Interaction in Musical Time*. PhD Thesis. Cambridge, UK: University of Cambridge.

Himberg, T. & Thompson, M. R. (2010). Dyadic entrainment and interaction in African dance. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th International Conference on Music Perception and Cognition, (ICMPC11)*. Seattle, WA: University of Washington.

Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, 27(6), 949–960.

Iyer, V. (2002). Embodied Mind, Situated Cognition, and Expressive Microtiming in African-American Music. *Music Perception*, 19(3), 387–414.

Janata, P. & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience* 6(7), 682–687.

Jensenius, A. R., Godøy, R.I. & Wanderley, M.M., (2005). Developing tools for studying musical gestures within the Max/MSP/Jitter environment. In *Proceedings of the International Music Computer Conference*. Barcelona: ICMC.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological review*, 96(3), 459–491.

Keller, P. E. (1999). Attending in Complex Musical Interactions: The adaptive dual role of meter. *Australian Journal of Psychology*, 51(3), 66–175.

Khvedelidze, B. V. (2002). Hilbert Transform. In M. Hazewinkel (Ed.), *Encyclopaedia of Mathematics*. Springer Verlag. Retrieved from http://eom.springer.de/H/h047430.htm

Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philosophical Transactions fo the Royal Society B: Biological Sciences*, 363, 2021–2031.

Konvalinka, I., Vuust, P., Roepstorff, A. & Frith, C.D. (2009). A Coupled Oscillator Model of Interactive Tapping. In In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P.-S. Eerola (Ed.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)*. University of Jyväskylä. Finland.

Kreutz, G., Bongard, S., Rohrmann, S., Hodapp, V., & Grebe, D. (2004). Effects of Choir Singing or Listening on Secretory Immunoglobulin A, Cortisol, and Emotional State. *Journal of Behavioral Medicine*, 27(6), 623–635.

Large, E. W., & Jones, M. R. (1999). The Dynamics of Attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

London, J. (2004). *Hearing in Time: Psychological aspects of musical meter*. Oxford: Oxford University Press.

Luck, G., Toiviainen, P., (2006). Ensemble Musicians' Synchronization with Conductor's Gestures: An Automated Feature-Extraction Analysis, *Music Perception*, 24, 189–199.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social Connection Through Joint Action and Interpersonal Coordination. *Topics in Cognitive Science*, 1(2), 320–339.

Mattar, A. A., & Gribble, P. L. (2005). Motor learning by observing. *Neuron*, 46(1), 153–160.

Moran, N. (forthcoming). Social co-regulation and communication in North Indian duo performances. In M. Clayton, B. Dueck & L. Leante (Eds.), *Experience and Meaning in Music Performance*.

Nettl, B. (2001). An Ethnomusicologist Contemplates Universals in Musical Sound and Musical Culture. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins Of Music*, (pp. 463–472). Cambridge, MA: MIT Press.

Overy, K. (2003). Dyslexia and music. From timing deficits to musical intervention. *Annals of the New York Academy of Sciences*, 999, 497.

Palmer, C. (1989). Mapping Musical Thought to Musical Performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15(12), 331–346.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, 48.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science*, 11(1), 63–68.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current biology*, 19(10), 827–830.

Pecenka, N. & Keller, P. (2010). The Role of Temporal Prediction Abilities in Interpersonal Sensorimotor Synchronization. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th International Conference on Music Perception and Cognition, (ICMPC11)*. Seattle, WA: University of Washington.

Rasch, R. A. (1988). Timing and Synchronization in Ensemble Performance. In J. A. Sloboda (Ed.), *The Generative Processes in Music* (pp. 70–90). Oxford: University Press.

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

Repp, B. H., & Keller, P. E. (2008). Sensorimotor Synchronization with Adaptively Timed Sequences. *Human Move*, 27(3), 423–456.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3(2), 131–141.

Sawyer, K. (2006). Group Creativity: Musical performance and collaboration. *Psychology of Music*, 34(2), 148–165.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own. *Cognition*, 88(3), B11–B21.

Temperley, D. (2000). Meter and Grouping in African Music: A View from Music Theory. *Ethnomusicology*, 44(1), 65–96.

Toiviainen, P., & Eerola, T. (2003). Where is the Beat?: Comparison of Finnish and South African Listeners. In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *5th Triennial ESCOM Conference* (pp. 501–504). Hanover, Germany: Hanover University of Music and Drama.

Toiviainen, P. & Burger, B. (2010). *MoCap Toolbox Manual 2.0*. University of Jyväskylä. Jyväskylä, Finland. Available at http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual

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

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

Volk, T. (1993). The History and Development of Multicultural Music Education as Evidenced in the Music Educators Journal, 1967–1992. *Journal of Research in Music Education*, 41(2), 137–155.

Wanderley, M. M., Vines, B., 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.

VI

EXPLORING RELATIONSHIPS BETWEEN PIANISTS' BODY MOVEMENTS, THEIR EXPRESSIVE INTENTIONS, AND STRUCTURAL ELEMENTS OF THE MUSIC

by

Marc R. Thompson & Geoff Luck

Musicae Scientiae, DOI: 10.1177/1029864911423457.

Article

Musicae Scientiae
1–22
© The Author(s) 2011
Reprints and permission: sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1029864911423457
msx.sagepub.com
SAGE

Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music

Marc R. Thompson and Geoff Luck
University of Jyväskylä, Finland

Abstract

Body movements during music performance have been found to be indicative of the performer's musical intentionality, and contribute to an observer's perception of expressive playing. This study investigates the effect of structural elements of the score, and the playing of different levels of expression on body movements during a piano performance. Pianists were required to play the same piece in four different performance conditions. Their movements were tracked by an optical motion capture system, and the comparisons that were made between specific parts of the body used, performance condition, and musical score locations were subsequently statistically examined. We found that the head and shoulders exhibited more movement per measure, as well as larger differences between each condition, than the fingers, wrists and lower back. Differences between performance conditions were observed primarily at structurally significant portions of the score, and biomechanical factors also played a role. Moreover, our data supports the view that performers equate playing without expression to playing without nonessential movements.

Keywords

gestures, expressive body movement, motion capture, performance conditions

Introduction

Expressive body movements are an integral part of performing music. This attribute, which is shared by all musical cultures (Blacking, 1973), may have evolved because of music's ancient association to dance, ritual, and ceremony. Like in dance, music-related movements can act as an indicator of expressive and emotional states. The communication of such states between individuals can function as a catalyst towards social bonding and shared experiences. Levitin (2008) notes that synchronized movements and coordinated song helped to create strong bonds among proto-humans, which ultimately allowed them to live in larger groups, not to mention augmenting their cognitive and social flexibilities (Cross, 2003).

Corresponding author:
Marc R. Thompson, Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, University of Jyväskylä, P.O. Box 35(M), 40014, Finland
Email: marc.thompson@jyu.fi

Before the onset of recording technology, music could not be perceived without accompanying exposure to expressive body movements. To this day, given that the recordings of any modern musician are accessible with the click of a button, live music performances remain arguably the quintessential way in which music can be perceived. One reason concerts are so engrossing is likely related to seeing the musicians' expressive gestures as they perform. While music is auditory, concerts present music as a multimodal experience, in which both auditory and visual modalities share an equal footing.

Musicians have proven successful in their ability to convey expression and/or emotions through dynamics and timing using acoustic cues (Juslin, 2000). However, a growing number of studies from the past 2 decades have focused on musicians' ability to convey expression via body movements (Davidson, 1993; Clarke & Davidson, 1998; Wanderley, Vines, Middleton, McKay, & Hatch, 2005; see also Schutz, 2008). An early finding within this time period showed that nonmusicians predominantly used the visual channel over the auditory channel to detect varying levels of expression in violin and piano performances (Davidson, 1993). More recently, it has been found that body movements can influence one's interest in the performance (Broughton & Stevens, 2009), provide cues to discriminate between specific emotional intentions (Dahl & Friberg, 2007) and affect the perceived duration of notes performed by a percussionist (Schutz, 2007). Meanwhile, a greater access to technology has allowed researchers to study body movements using a more computational approach. For instance, Wanderley (2002) and Wanderley et al. (2005) used a motion tracking system to quantify expressive body movement in clarinet performances.

A number of these studies adopted a framework in which musicians were requested to perform the same piece of music using different performance conditions based either on expressive intentions (Davidson, 1993, 1994), emotional intentions (Dahl & Friberg, 2007) or quantity of body movement (Wanderley, 2002; Wanderley et al., 2005).

Davidson (1993, 1994) asked musicians to play in performance manners denoted as deadpan, projected, and exaggerated while Wanderley (2002) used the terms immobile, standard, and expressive. The two sets of performance conditions could be seen as analogous to one another as Davidson's manners emphasized a gradient of expression and Wanderley's a gradient in amount of movement. However, the conditions themselves may overlap. For instance, projected and standard both refer to performing normally, as if for a public audience. On the other hand, the conditions deadpan and immobile differ, as the immobile condition specifically requires musicians to perform with as little movements as possible while deadpan requires musicians to perform without expression. Interestingly, Davidson (1993, 2007) observed that musicians associate playing without expression with using smaller or fewer movements. Therefore, an interesting relationship between both conditions may reveal itself if musicians were asked to perform in both the deadpan and immobile conditions.

Consequently, for the current study, we synthesized the performance conditions used by Davidson (1993, 1994), Wanderley (2002) and Wanderley et al. (2005) and tested the statistical differences that occur in body movements when each manner of performance is executed. We asked pianists to perform a short prelude (E minor, Op. 28, No. 4) by Chopin in conditions denoted as deadpan, normal, exaggerated, and immobile. The first three conditions refer to the amount of expression to be employed in the performance while the last one refers to playing normally, but with as little movement as possible.

Musical expression is a multidimensional construct and its definition can often change depending on the context (Juslin, 2003). For our purposes, we have espoused the view that musical expression refers to the micro and macro variations in timing and dynamics that occur in music performance. As Palmer (1997) put it, it is what differentiates one performance from another.

Some attributes of musical expression are immediately recognizable when listening to a piece of music. For instance, Gabrielsson and Juslin (1996) found that listeners could easily perceive changes in tempo, dynamics, and timing when musicians altered their expressive intentions. These parameters, particularly expressive timing, have been widely studied perceptually (Clarke, 1989) and successfully modelled (Sundberg, Friberg, & Bresin, 2003; Todd, 1985).

The parameter of musical expression we are interested in, expressive movement, is perceived through vision. Just as music's audible properties require a high-resolution medium such as high-quality audio or MIDI to be analyzed, the proper tracking of movements also commands high-fidelity technology. Therefore, we have used a motion capture system to collect data on pianists' expressive movements. The system uses multiple infrared cameras to triangulate the displacement of markers attached to a subject within a Cartesian positioning system. Using such systems, movements can be tracked in three dimensions and analyzed as time series data, from which kinematic features can be extracted.

Our main goal was to perform statistical analyses on the differences in body movements that arise when pianists perform the same piece in performance conditions based on those used by Davidson and Wanderley, and to discover which parts of the body changed most between conditions. We also looked at changes in expressive timing and dynamics, which led to interesting comparisons between audio and kinematic features (Caramiaux, Bevilacqua, & Schnell, 2010; Palmer, Koopmans, Loehr & Carter, 2009; Thompson & Luck, 2008).

We argue that expressive movements during performances are related to factors ranging from the instrument's physical shape, the music's structure and the musician's interpretive choices. While our main focus was to conduct quantitative research on motion capture data, we collected interview data in which participants self-reported their strategies for performing each condition. This qualitative element allowed us to parse our analytical results through the lenses of the pianists.

An embodied approach

The study of music-related body movements has become a far-reaching interdisciplinary field bridging areas such as music theory, biomechanics, cognitive science, and artificial intelligence. These disparate fields are all merged within the framework of embodied music cognition. The embodied paradigm is driven by the idea that the physiological makeup of the human body governs the nature of the mind (Varela, Thompson, & Rosch, 1991). The body acts as a mediator of the physical world in which experience, intentionality, and engagement are parsed to and from a mental level (Leman, 2008). This results in a perpetual action–perception relationship involving the mind, body, and environment. At the crux of embodied music cognition, there is the gesture, which is produced in the physical world but is deemed meaningful through mental activation (Godøy, 2008; Jensenius, Wanderley, Godøy, & Leman, 2010).

Some gestures are not directly involved in the production of sound, but are purposeful for musical expression and interpretation (Cadoz & Wandereley, 2000; Jensenius et al., 2010). These so-called ancillary gestures (Wanderley, 2002) include foot tapping, head bobbing, or facial expressions.

Wanderley (2002) studied ancillary gestures in clarinettists' performances by requesting they perform in the standard, expressive, and immobile conditions. Focusing on the vertical motion of a sensor attached to the clarinet's bell, it was found that the standard and expressive renditions had similar motion while the immobile rendition showed some of the same movements but at a smaller order of magnitude. Similar results occurred in a follow-up study by Wanderley et al. (2005). In the immobile condition, clarinettists tended to attenuate their body sway and also the motion of the clarinet bell. However, small movements following similar trajectories as the movements in the standard and expressive conditions remained. Furthermore, immobile conditions were usually played at a faster tempo than standard and expressive conditions. Wanderley and colleagues posited that sway contributed to a clarinettist's sense of global timing and was integrated into a motor program[1] conceived over the course of hours of practice. They argued that it was the lack of continual motion in the immobile condition that contributed to an attenuated sense of timing and phrasing, resulting in rushing through the piece.

Davidson (2007) made similar observations about movement attenuation when studying a video of a pianist performing a Beethoven bagatelle in the deadpan, projected, and exaggerated expressive manners. She identified expressive locations in which the same body movements in the head and hands were repeated throughout every intention. Although the movements were similar in shape throughout all performance manners, the movements were smaller in the deadpan intention and largest in the exaggerated intention. Not only does this show a clear relation between expressive intention and amplitude of movement but suggests that the choice of a movement at an expressive location is rooted in functionality. When expression is removed from a performance, the movements persists but at a smaller amplitude. As the level of intended expression grows, the movement increases in size yet remains similar in shape.

The embodied approach demonstrates how musicians use gestures to communicate expressive intentions and provide visual cues as to the music's temporal organization and structure. Consequently, gestures are an overt display of one's engagement with the music being performed.

Musical gestures, structure, and expression

The gestures effectuated in music performances are said to be analogous to the para-linguistic gestures found in everyday speech (Wanderley et al., 2005). McNeil (2005) stressed that gestures are engrained within language and, due to their spontaneity, play a pivotal role in shaping speech. Such naturally occurring gestures can be grouped into categories outlining different ways in which gestures enhance speech (Ekman & Friesen, 1969; McNeil, 1992). In much the same way, gestures used for communicating musical ideas both support and clarify musical intentionality. The content of such gestures rely on a series of key factors related to the instrument being played and the musical work being performed that are situational and context-dependent. A framework established by Delalande (1988) and reported by Wanderley (2002) attempts to outline these factors as being either material/physiological, structural, or interpretative.

Material and physiological factors

Material and physiological factors refer to the manipulation and ergonomics of an instrument in order to make sound. The size and locality of a movement is largely based on the instrument being performed on. Sound propagation techniques (bowing, blowing, or pressing) will likewise result in different motion content. Clarinettists performing a solo are usually standing. This gives them more freedom to sway, bend the knees and produce circular motion with the clarinet bell. Instruments that require a more stationary playing position such as piano will result in expressive movements localized to the upper body. Likewise, pitched percussion instruments such as the marimba require a standing position and additionally produce a limited timbre with fast decaying notes. Given the limitations on these instruments, the musician's movements are of most importance for conveying expressive intentionality (Broughton & Stevens, 2009).

Structural factors

Clarke and Davidson (1998) and Palmer (1989) have studied relationships between musical structure and expressive body movement. Musical timing and phrase structure are among the parameters of music that musicians use to communicate structural interpretation (Clarke, 1989). It has been showed that structure is reflected in expressive movement. Conductors parse their interpretation of music's temporal structure to syncrohonize a choir (Luck & Toiviainen, 2006) and phrasal structure can be perceived through movement without sound (Vines, Wanderley, Krumhansl, Nuzzo & Levitin, 2004). For piano, the temporal organization of music can often be reflected in body sway, although a clear periodic motion referencing the music's tempo is rare (Clarke & Davidson, 1998).

Interpretive factors

Finally, expressive movement has a causal relationship to interpretive factors. While standard musical scores used in Western music provide information regarding pitch and duration, they are more ambiguous when it comes to dynamics and larger structures such as form and phrase boundaries. These larger structures afford musicians interpretative choices according to their musical intentionality (Palmer, 1997). These interpretive choices can be driven to extremes such as in the frenetic gestures of Glenn Gould (Delalande, 1988), or more recently, the pianist Lang Lang.

Current study

We chose to study pianists because unlike smaller wind and string instruments, which may be performed while standing, pianists are restricted in the amount of expressive movement they can execute during performances. The rationale behind studying piano performances is similar to that of Broughton and Stevens' (2009), who were interested in the body movements of marimba players. Both instruments are limited in expressiveness (e.g., absence of vibrato effect) and have a relatively short decay time (Schutz, 2008). More importantly, the piano like the marimba is performed in a manner in which the head is allocated a larger degree of freedom than other parts of the body. It is likely, therefore, that the musician will use the head and upper torso to express musical ideas (Castellano, Mortillaro, Camurri, Volpe, & Scherer, 2008; Davidson, 2007). Meanwhile, the fingers, wrists, and elbows are ergonomically required to play the notes on the piano, effectively restricting movements to the physical requirements of the musical score.

We report on pianists performing the Chopin's Prelude in E minor Op. 28, No. 4. The decision to use this particular piece was influenced by a study by Clarke and Davidson (1998), which also used this piece. Whereas their study relied on manual video frame tracking for positional information, we have used an optical infrared motion capture system to track the movements of pianists in three dimensions and at a much higher temporal resolution than regular video can capture.

We conducted several analyses that focused on differences in movement, timing, and dynamics across the four conditions. We also collected pianists' strategies for performing each condition. The variety in expressive timing throughout each measure and across the conditions was quantified by conducting separate analyses of variance (ANOVA) for each measure. We also investigated differences in dynamic range between performance conditions by calculating the root mean square (RMS) of the audio signal on each measure in the piece. Following this, the total amount of movement per measure for selected parts of the body was averaged across participants and ANOVAs were conducted on this data for each individual measure. Finally, Pearson correlations were conducted on velocity and acceleration data extracted from the position data. The goal was to compare performance conditions for each condition statistically and give credence to the idea that expressive movement gestures are formed as a result of the physiological, structural, and interpretative factors discussed previously. For example, expressive ancillary head movement may increase at sections of the score that are climactic and/or structurally significant because pianists are physiologically able to use head movements to demonstrate increased musical expressivity.

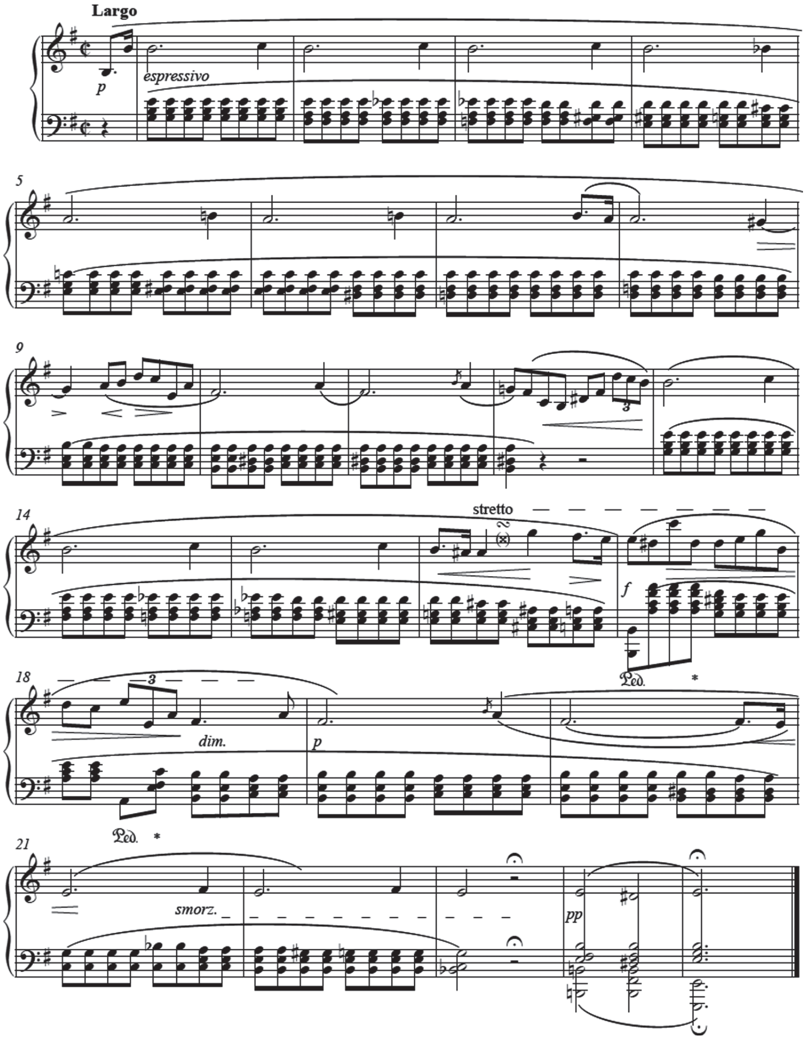



Figure 1. Chopin: Piano Prelude in E minor, Op. 28, No. 4.

Materials and method

Performers

Eight pianists performed the Chopin's Prelude in E minor, Op. 28, No. 4 (see Figure 1). A theoretical description and analysis of the piece's form can be found in Clarke and Davidson (1998). The participants were associated to either the University of Jyväskylä or to the Jyväskylä

University of Applied Sciences, Finland, and were recruited via university student body mailing lists. The participants, consisting of five females and three males, each had between 10 and 20 years of piano-playing experience and an average age of 24.6 years. Their nationalities were the following: three Finns, one Vietnamese, two Taiwanese, one Hungarian, and one German. The pianists were asked to identify themselves as professional musicians (two participants), music students (three participants), or musical hobbyists (three participants).[2]Upon showing interest in the study, the participants were provided with a score of the music and requested to practice the piece at home before the date of the experiment. They were compensated for their time and effort with a €30 gift voucher for Sokos department store.

Apparatus and procedure

The pianists performed on a Roland digital piano keyboard (HP-series) (Roland Corporation, Helsinki, Finland) with weighted keys and were captured at 120 frames per second using an 8-camera Qualisys ProReflex Motion Capture System (Qualisys Motion Capture Systems, Gothenburg, Sweden). A total of 26 markers were placed on the pianists' upper body while two reference markers were placed at each end of the keyboard. Audio was recorded using the ProTools 8 software (Pro Tools, 2010). The sources for the audio recordings were two omni-directional condenser microphones positioned above the keyboard 2.5 meters from the ground, as well as the direct line signal from the keyboard. A synchronization pulse signal, generated by the motion capture cameras was additionally routed into ProTools where it could be viewed as an audio signal on its own track. The pulse signal showed the start and stop time of the motion capture recordings. This provided an accurate visual cue for the lapse of time between the start of a motion capture recording and the onset of the performance, and enabled movement and audio data to be properly synchronized during postprocessing.

While the participants were requested to practice the piece at home prior to attending the recording session, they were not aware of the exact task. When they arrived, they were instructed to perform the piece a total of 12 times (three cycles of four expressive intentions). Playing through the cycle three times provided a measure of consistency for each pianist's interpretation of the instructions. The players read the instructions from a sheet of paper, which was kept by the side of the keyboard for reference. Verbal instructions were restricted unless a participant specifically requested clarification. The instructions were as follows:

> *Play the piece in the following performance types.*
>
> *Repeat the cycle 3x.*
>
> *Play at the tempo you think is appropriate for the performance type.*
>
> ***Normal Performance:*** *Play the piece with the normal level of expressive intensity that you would use if you were performing the piece in a recital, for a friend or for a teacher.*
>
> ***Deadpan Performance:*** *Play the piece at a reduced level of expressive intensity. Ignore dynamic markings in the score and play in a deadpan fashion.*
>
> ***Exaggerated Performance:*** *Play the piece with a maximum level of expressive intensity. Exaggerate the dynamic markings and add your own.*
>
> ***Immobile Performance:*** *Play the piece with as little movement as possible. Restrict your movement to what is absolutely necessary while trying to play the piece with a regular level of expression.*

Table 1. Areas of interest and the original marker locations.

#	Area of interest	Original markers
1	Hip centre	Centre of four hip markers
2	Mid torso	Midpoint sternum and middle spine markers
3	Neck	Midpoint of clavicle and upper spine markers
4	Head centre	Centre of four head markers
5	Left shoulder	Left shoulder marker
6	Left elbow	Left elbow marker
7	Left wrist	Left wrist marker
8	Left middle finger	Left middle finger marker
9	Right shoulder	Right shoulder marker
10	Right elbow	Right elbow marker
11	Right wrist	Right wrist marker
12	Right middle finger	Right middle finger marker

Table 2. Average duration for each performance type and their standard deviations.

Performance condition	Average duration (sec)	Standard deviation (sec)
Normal	116.5	13.8
Deadpan	110	15.5
Exaggerated	119.3	13.8
Immobile	115.6	13.6

takes as input the motion capture data, the values representing the duration of each segment and a timing model. It uses spline interpolation, which is a polynomial-based curve fitting technique, to map the performances' segments to a timing model by stretching or compressing the data. The technique lined up the onset of each measure despite the performances' disparities in individual timing. A time differentiation algorithm was then applied to the time-warped data to calculate the velocity and acceleration.

Results

Performance duration and timing

The total durations of each of the pianists' performances were averaged together to examine how the performance conditions differed in terms of general timing. Table 2 shows that the deadpan condition generally had the shortest duration (110 seconds) while the exaggerated condition had the longest duration (119.3). Figure 2 plots the average duration of each measure and presents a more detailed view of the differences between performance conditions. From Table 2 and Figure 2 it can be seen that increased expression resulted in performances of longer duration, both at the level of each individual measure and overall. The exception is the immobile condition, the duration of which was similar to the normal condition. Deadpan performances contained the least amount of timing variation while the exaggerated performances contained the most. Immobile and normal performances shared about an equal amount of

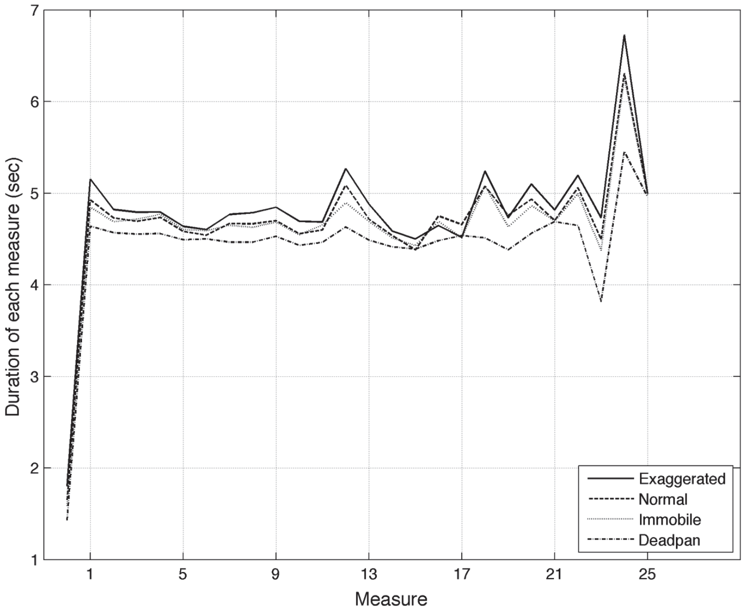


Figure 2. Duration comparison for each measure across each performance condition.

timing variation. Tempo was regular for most of the first half of the piece, with increased deviation occurring around Measure 12, and continuing variation in the second half. Deviations increased near the climax of the piece (Measures 22–24) to determine whether the duration was significantly different under each performance condition, the durations of all 96 performances (eight pianists, four performance conditions, each played three times) were organized into a 24 x 4 matrix with the columns representing the performance conditions. A one-way analysis of variance (ANOVA) indicated that overall timing differences between performance conditions were not significant, $F(3, 92) = 1.77$, $p = 0.158$.

Still, Figure 2 suggests timing differences between performance conditions occurred at individual measures. For this reason, one-way ANOVAs were conducted for each measure individually. While most of these analyses produced nonsignificant results, timing differences were found for Measure 12, $F(3, 92) = 4.68$, $p < 0.01$, Measure 18, $F(3, 92) = 3$, $p < 0.05$, Measure 22, $F(3, 92) = 3.38$, $p < 0.05$, Measure 23, $F(3, 92) = 7.79$, $p < 0.001$, and Measure 24, $F(3, 92) = 4.84$, $p < 0.01$. As a follow-up, post hoc tests with Tukey-Kramer correction were performed. The results from these tests are shown in Table 3. They indicate that it was mostly discrepancies in duration between the deadpan and exaggerated performance types that accounted for the ANOVAs being statistically significant. However, for Measure 23, other comparisons were found to also have statistically significant different mean durations (normal and deadpan; deadpan and immobile). This is to be expected as Measure 23 contains a fermata, which instructs pianists to hold for an indeterminate duration of time.

Table 3. The results from pair-wise comparisons for the measures that were found to have statistically significant different durations.

	Mean difference	95% Confidence interval	Sig.
Measure 12			
Normal/Deadpan	0.455	[−0.0105 0.92]	*ns*
Normal/Exaggerated	0.181	[−0.646 0.284]	*ns*
Normal/Immobile	0.19	[−0.275 0.656]	*ns*
Deadpan/Exaggerated	0.636	[−1.1 −0.171]	$p < 0.05$
Deadpan/Immobile	0.264	[−0.729 0.201]	*ns*
Exaggerated/Immobile	0.372	[−0.0936 0.837]	*ns*
Measure 18			
Normal/Deadpan	0.563	[−0.119 1.25]	*ns*
Normal/Exaggerated	−0.167	[−0.849 0.516]	*ns*
Normal/Immobile	−0.000416	[−0.683 0.682]	*ns*
Deadpan/Exaggerated	−0.73	[−1.41 −0.0476]	$p < 0.05$
Deadpan/Immobile	−0.564	[−1.25 0.118]	*ns*
Exaggerated/Immobile	0.166	[−0.516 0.848]	*ns*
Measure 22			
Normal/Deadpan	0.412	[−0.0565 0.88]	*ns*
Normal/Exaggerated	−0.135	[−0.603 0.334]	*ns*
Normal/Immobile	0.0666	[−0.402 0.535]	*ns*
Deadpan/Exaggerated	−0.546	[−1.01 −0.078]	$p < 0.05$
Deadpan/Immobile	−0.345	[−0.813 0.123]	*ns*
Exaggerated/Immobile	0.201	[−0.267 0.669]	*ns*
Measure 23			
Normal/Deadpan	0.679	[0.165 1.19]	$p < 0.05$
Normal/Exaggerated	−0.234	[−0.748 0.28]	*ns*
Normal/Immobile	0.12	[−0.394 0.634]	*ns*
Deadpan/Exaggerated	−0.914	[−1.43 – 0.4]	$p < 0.05$
Deadpan/Immobile	−0.559	[−1.07 −0.0449]	$p < 0.05$
Exaggerated/Immobile	0.355	[−0.159 0.869]	*ns*
Measure 24			
Normal/Deadpan	0.852	[−0.0434 1.75]	*ns*
Normal/Exaggerated	−0.423	[−1.32 0.473]	*ns*
Normal/Immobile	0.0476	[−0.848 0.943]	*ns*
Deadpan/Exaggerated	−1.27	[−2.17 −0.379]	$p < 0.05$
Deadpan/Immobile	−0.805	[−1.7 0.091]	*ns*
Exaggerated/Immobile	0.47	[−0.425 1.37]	*ns*

Evaluating amount of movement per measure across performance conditions

Using the time-warping technique described above, the entire dataset was temporally aligned according to the onsets of each measure. To create evenly spaced measures, the timing model used was a linearly spaced vector with a range from 0 to 1 in 26 increments. The 26 steps accounted for the durations of each measure (25) plus the duration of the pickup to Measure 1.

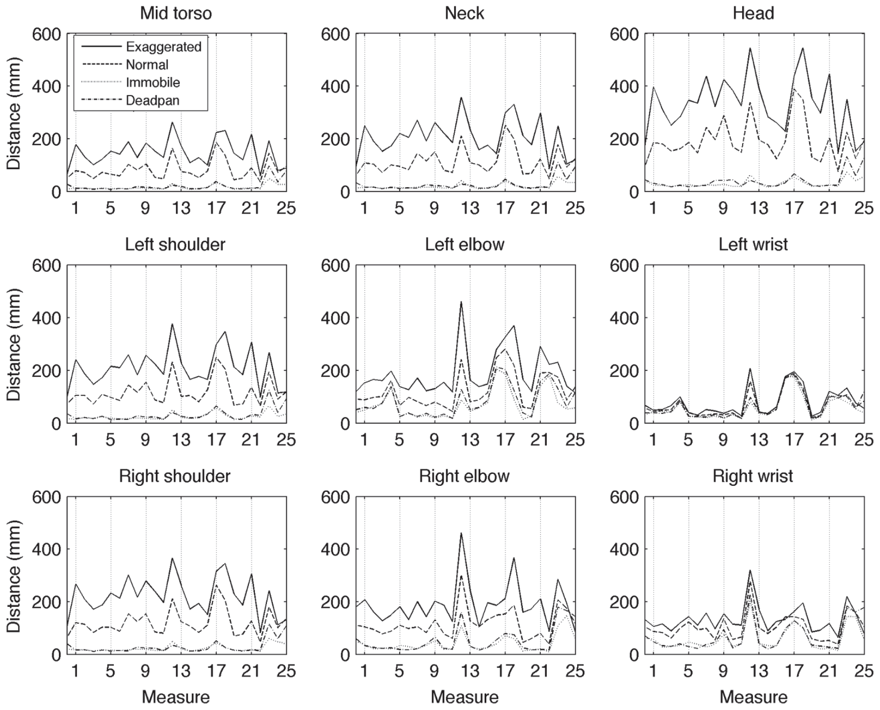
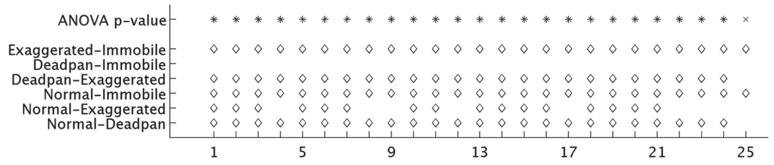


Figure 3. The total distance per measure for nine areas of interest. Parts of the body further from the keyboard show more movement as expression was increased.

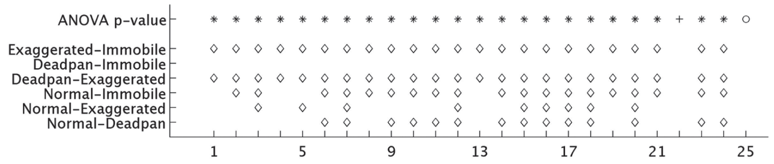
From the time-warped data, the total distance per measure for each area of interest (see Table 1) was calculated and averaged across participants for all performance types. Figure 3 shows the progression of amount of movement throughout the piece for nine areas of interest. From the figure it can be seen that the wrists, particularly the right wrist, differed very little in amount of movement between the performance conditions. However, for areas such as the head, torso, shoulders, and elbows, there appear to have been changes in the amount of movement as the players performed with more expression. Additionally, the amount of movement between deadpan and immobile conditions was very similar for most parts of the body, demonstrating that the pianists equated deadpan performances with immobile performances in terms of the amount of movement.

Figure 3 presents evidence that differences in amount of movement occurred as a result of increased expression and that these differences were greater in the parts of the body further away from the keyboard and not involved in sound production. A series of one-way ANOVAs were conducted to test whether the differences in total amount of movement between performance conditions were indeed greater in the torso and head than in the arms and wrists. We predicted that differences between performance conditions should be statistically significant for the head and torso region while differences in the arms and wrists would not be statistically significant. In total, 300 ANOVAs were conducted (12 areas of interest x 25 measures). The

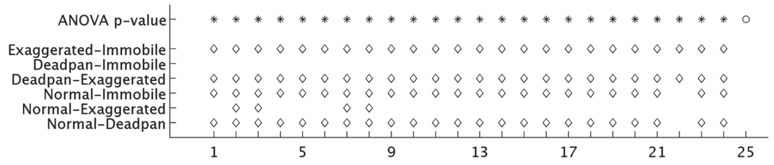
a) Head



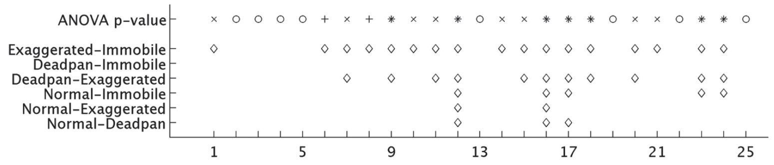
b) Left elbow



c) Right elbow



d) Left wrist



e) Right wrist

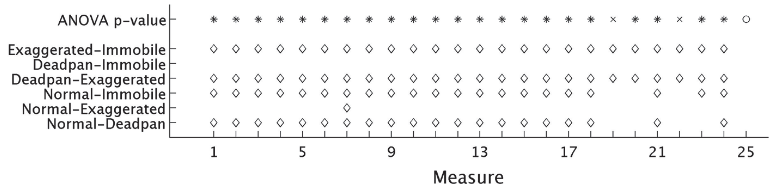


Figure 4. Results of one-way ANOVAs on total distance travelled of different body parts during each measure of the piece. For each section, the top row indicates the overall significance level of each ANOVA, while the following six rows indicate pair-wise comparisons (Tukey-Kramer) significant at $p < .05$.
Note: $^*p < 0.001$; $^\times p < 0.01$; $^+p < 0.05$; $^\circ p = ns$ (pair-wise comparisons with significantly different means).

results from this analysis are summarized in Figure 4. Upon performing the analysis, it was noted that the hip center, mid torso, neck, and both shoulders all had very similar results to the head while the elbows and wrists had the most contrasting results. For this reason, only the results for the head, elbows, and wrists are summarized in Figure 4.

To clarify how to interpret Figure 4, a couple of examples will be given. The first column of Figure 4a, for instance, indicates the results of the ANOVA for head distance travelled during Measure 1. The uppermost symbol shows that the overall ANOVA for this measure was significant at $p < .001$, while the symbols below indicate that there were significant differences in amount of head movement between all conditions except between deadpan and immobile. The seventh column of Figure 4d, meanwhile, shows the results of the ANOVA for the left wrist distance travelled during Measure 7. The uppermost symbol reveals that the overall ANOVA for this measure was significant at $p < .01$, while the symbols below show that there were significant differences in amount of left-wrist movement between the exaggerated and immobile and deadpan and exaggerated conditions only.

The information presented in Figure 4 exposes two key points. First, pair-wise comparisons yielded no significant differences between the immobile and deadpan performance types. This trend points towards an inherent association between playing with no expression and playing with restrained movements. Second, pair-wise comparisons for areas such as the head (Figure 4a) yielded significant differences throughout the entire piece with the exception of normal and exaggerated performance comparisons. Meanwhile the left wrist was the part of the body that varied the least throughout the four performance conditions with the exception of Measures 12, 16, and 17. The contrast draws attention to the notion that parts of the body chiefly responsible for sound production are less used for demonstrating musical expression through body movement.

Correlations between the velocity and acceleration of each performance type

In order to compare the velocity and acceleration contours across performance conditions, Pearson correlations were performed for each of the areas of interest. First, the position data from each performance were once again time warped according to the onset of each measure. However for this analysis, the timing model was different for each participant. The timing model for each participant used for the time warping was the performance with the median duration of the 12 performances (three performances by four performance conditions). Again, this initial procedure allowed for the data compared to be of equal duration and normalized with respect to the onset of each measure across performances.

Following the time-warping procedure, the components of velocity and acceleration were estimated by calculating the first and second time derivative of the position data. In order to minimize the amount of high-frequency noise inherent in time derivation procedures, a Savitzky-Golay FIR smoothing filter was used (Savitzky & Golay, 1964). This algorithm fitted a second-order polynomial to 13 consecutive frames and computed the derivative of the thus obtained polynomial at the centre frame (Luck & Toiviainen, 2006). Finally, the magnitude of velocity and magnitude of acceleration for each area of interest as listed in Table 1, was obtained by calculating the norm of the vectors containing the components of velocity and acceleration.

The curves representing the magnitude of velocity and magnitude of acceleration were collapsed across the participants' performances for each area of interest. These curves were organized into matrices with each column representing one of the four performance conditions. Pearson correlations were performed on these data. Results from the Pearson correlations are presented in Figure 5. This figure contains two topographic charts in which the shading of each rectangle represents the strength of the correlation between two performance conditions and one of the twelve areas of interest. For example, the first column from

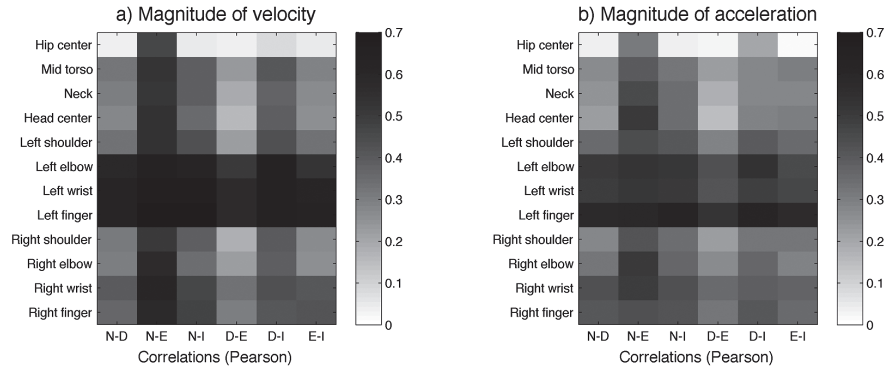


Figure 5. Results of Pearson correlations performed on the speed and acceleration curves on different parts of the body. The shade of each rectangle represents the strength of the correlation between paired performance conditions. Lighter shades represent correlations with low *r*-values while rectangles with darker shades represent correlations with higher *r*-values. The pairing of the performance conditions is indicated on the horizontal axis (N = normal performance, D = deadpan performance, E = exaggerated performance, I = immobile performance). All *r*-values are significant at $p < 0.05$.

the left in Figure 5a represents the *r*-values for correlations between normal and deadpan performances; the second column represents correlations between normal and exaggerated performances; etcetera. From Figure 5a, it can be seen that magnitude of velocity curves was most highly correlated between normal and exaggerated performances while the lowest correlations occurred between the deadpan and exaggerated performances. The effect of the deadpan-expression and exaggerated-expression performances being uncorrelated demonstrates the greatest discrepancy between two performance conditions with respect to expressive movement. The exception to this trend occurs in correlations involving the left elbow, wrist, and finger. These areas of interest were highly correlated across all performance conditions. This result was most probably caused by the monotonous nature of the piece's left-hand part, which plays block chords throughout the piece, leaving little room for movement variation. Figure 5b shows the correlations for magnitude of acceleration. While the trend of the correlations is similar to that of the magnitude of velocity correlations, the chart features *r*-values, which are generally lower than those of velocity. This is to be expected as acceleration curves generally contain more random high-frequency noise than velocity curves. Still, the correlations for the left finger are almost as high as those for the magnitude of velocity, outlining the consistency with which the left-hand part of the score was played despite performing with more or less expression.

Performance dynamics

To explore the changes in dynamics across performance condition, we calculated the RMS value for each measure of every performance individually, converted the values to a percentage and took an average value for each pianist, and then a group average. The conversion to a percentage was necessary because not all RMS values were in the same comparable range. Figure 6 shows the group average RMS values, presented as a percentage of the maximum

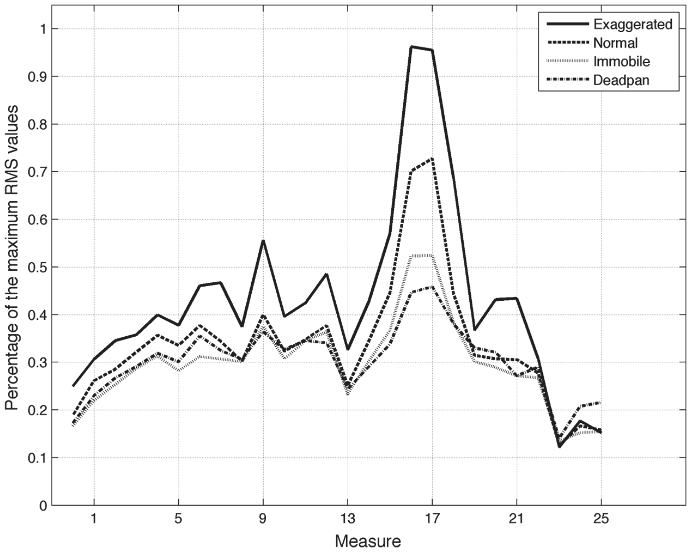


Figure 6. RMS comparison for each measure across each performance condition. The RMS values are presented as a percentage of the maximum RMS value.

RMS value for all performance conditions plotted against measure. The figure generally shows three things. First, the exaggerated performance condition yielded the most dynamically varied performances. Second, the other performance conditions were similar in dynamic range, particularly at the beginning of the piece. Third, the dynamics increase for all performance conditions to some extent around the *stretto* and *forte* markings of Measure 16 in the score (Figure 1).

Interview data

The pianists were asked to write about the strategies they had adopted for playing in each performance condition. The instruction was phrased as follows: Briefly explain what you did to properly convey the different conditions of playing.

For the normal condition, pianists generally wrote that they focused on playing naturally. Three of them mentioned how the normal condition influenced them to imagine they were playing the piece for a friend or for their teacher while others “tried not to think too much” and “forgot where they were.”

The deadpan condition required “surprisingly much concentration [sic]” and was characterized as being “mechanical” and “machine like.” One pianist avoided proper phrasing and purposely “searched for a bad balance between right and left hand,” while others ignored dynamic markings on the score. Only one pianist mentioned movement, pointing out that they purposely suppressed their movements while trying to mimic a MIDI performance. Others imagined they disliked the piece, that it had been assigned as homework and they were just trying to get through it.

For the exaggerated condition, pianists emphasized score dynamics and varied their tempo throughout. Some were apprehensive about this condition as they were unsure what they should exaggerate while others were enthusiastically in the moment. One colourful commentary mentioned "physically leaping for the loud notes and crouching for the quiet ones" while another mentioned greater head movements as their dynamics became more *forte*.

Overwhelmingly, the immobile condition proved to be the most challenging condition. The participants took the instruction very seriously and restricted not only head and torso movements but also tried to reduce hand, leg, and foot movements. Some pointed out how difficult it was to perform in a "normal" way while attenuating all unnecessary movements. One pianist mentioned that "it was difficult to include expressive playing without moving; I learned to concentrate on immobility by slightly reducing my level of expressive interpretation. Sometimes the music carried me away and I quickly had to think of my movements again."

To summarize, the pianists' descriptions of their strategies focused on what they were thinking as opposed to what they physically did. The differences in movements from one condition to the next were guided by mental imagery such as playing the piece for a friend, or imitating a MIDI performance. This use of imagination applied predominantly for the deadpan, normal, and exaggerated conditions. For the immobile condition however, descriptions focused on physical strategies such as how to play expressively without extra movements. The overall consensus was that this condition required much more concentration and alertness.

Discussion

The aims of this study were to investigate changes in body movement, timing, and dynamics across different performance conditions. Eight pianists performed the Chopin's Prelude in E minor, Op. 28, No. 4. The instruction given to the pianists was to cycle through the piece using four performance conditions; normal expression, deadpan expression, exaggerated expression, and immobile. The cycle was repeated three times to provide a measure of consistency for each pianist's performance style.

When comparing performance durations across performance conditions, it was found that deadpan performances were on average the shortest in duration while exaggerated performances were the longest in duration (Table 1). Further, deadpan performances contained the smallest timing variations between each measure (Figure 2). The discrepancy in timing from low expression to high expression highlights a greater use of timing variation as the piece was played more expressively. However, as post hoc pair-wise comparisons revealed, the difference in timing between deadpan and exaggerated conditions was only statistically significant for five out of 25 measures. The differences occurred at areas of the score that contained harmonic tension and/or phrase boundaries. For example in Measure 12, the melody contains an arpeggiated dominant seventh chord, which leads to the return of the opening melody. This is additionally seen in Measure 24 (Figure 2), which contains two held chords. In both cases, performers were influenced by structural factors of the composition to use expressive timing at these points.

While the greatest differences in duration were between deadpan and exaggerated performances, immobile and normal performances were very close to being the same in overall duration (Table 2 and Figure 2). The similarity in overall duration between the immobile and normal performances reflects the instructions given to the participants. The immobile condition was played with a normal level of expression but with the minimum amount of movements needed to perform the piece. Similarities in timing between the normal and immobile renditions indicate that although the pianists were instructed not to move, they still made use of expressive timing when playing the immobile condition. This is related to findings by Wanderley et al. (2005), but with slightly different outcomes. In their study involving clarinet performance, the immobile condition was generally performed faster than the standard performance. The authors hypothesized that the absence of natural movement altered the players' sense of global timing caused by a disruption in their motor programs. The interview data showed no indication that the pianists felt playing without movement affected their timing. However, they did mention that the immobile condition was the hardest to play, requiring the most concentration.

There are several reasons why the current results may differ to those of Wanderley and colleagues. First, the instrument being examined was different, as was the musical style. Wanderley et al. (2005) examined performances of Stravinsky's Second Piece for Solo Clarinet. If one were to compare both pieces, Stravinsky's piece is arguably more technically demanding for clarinettists than Chopin's prelude is for pianists. It could very well be that playing without extraneous movements affected global timing in the case of clarinettists performing the Stravinsky's piece. However, in the current study, there is no evidence of this happening. On the contrary, normal and immobile performances were most similar in performance duration. This highlights the fact that while pianists restrained their amount of movement, they still made use of timing as a parameter of expression. Future work might extend the current study by examining the effect of more complex and technically demanding pieces on the movements of pianists to see how they compare with those of the clarinettists in Wanderley and colleagues' study.

The second analysis dealt with examining the amount of movement throughout the piece across performance conditions. First, the movement data was time warped so that each measure contained the same number of data points. For each measure, the consecutive data points within the measure were differenced and summed. These values were dubbed the total amount of movement per measure. From observing the distance curves in Figure 3, it was found that increasing expression from deadpan, to normal, to exaggerated resulted in more movement in areas of the body further from the keyboard such as the head, neck, and shoulders. Further, most of the variation in amount of movement occurred in the same measures that were shown to have significant differences in duration between the deadpan and exaggerated conditions (see Table 3; Measures 12, 18, 22, 23, and 24). For example, Measure 12 is structurally significant as it prepares a return to the opening melody and Measure 18's content is the culmination of a *stretto*, contains a *forte* marking, and can be seen as the piece's climax. From these results, it may be hypothesized that pianists use the parts of the body with the highest degree of freedom to parse their expressive intentionality. A future direction of this finding may involve observing the pianists as individuals other than investigating the trends within a group. Observing individual differences between pianists (and perhaps other musicians) will help determine if the hypothesis can be generalized to all musicians.

The other finding in this analysis is the similarity in amount of movement between the deadpan and immobile conditions. Figure 3 shows little variation between these conditions for all areas of interest including the head. The observation was confirmed by ANOVAs performed for each measure and area of interest. Figure 4 shows that the amount of movement per measure was never significantly different between immobile and deadpan performances, indicating a clear association between playing without expression and playing without extraneous movements. Clearly as a group, the pianists associated playing in the deadpan condition to playing in the immobile condition. Still, as far as timing is concerned musicians were able to employ expressive timing as it occurred in the normal condition to the immobile condition. The relationship between the immobile and deadpan conditions is interesting and could be investigated further in a perceptual study where observers could be asked to discriminate between immobile and deadpan performances.

The third analysis compared the velocity and acceleration data between each performance condition using Pearson correlations. For both velocity and acceleration, the most highly correlated areas of interest were for the left elbow, wrist, and finger. This underlines the consistency with which these parts of the body moved regardless of the performance condition. This again helps to establish that movements are a result of structural factors found in the score being performed. Regarding pair-wise comparisons, normal and exaggerated performances contained the most highly correlated velocity and acceleration trajectories. An explanation for this would be that velocity and acceleration do not change much after a certain level of expressivity has been reached. They could only be altered up to a certain point. Related to this is the fact that the least correlated performance conditions were deadpan and exaggerated. If we consider that the amount of expression is a continuum, the differences exist when comparing the extremes (deadpan and exaggerated).

Our brief evaluation of the differences in dynamics calculated the RMS value for each measure individually. As an indicator of energy for that time segment, the RMS values were inferred as the amount of amplitude or volume currently being performed. The analysis confirmed that as a group, the pianists showed most dynamic variations in the exaggerated condition, while the other conditions remained relatively equal, except for measures around the prelude's climax.

When pieced together, several interesting observations surrounding the data can be made. First, the study provided further evidence that playing with more expression led to larger movements, reinforcing the findings of Davidson (2007) and others. Further, the increased movement amplitude seem to always take place at the same moments and at structurally significant portions of the score, such as Measures 12, 18, and at the final cadential progression. This finding adds credence to the idea that expressive movement is influenced by musical structure (Clarke & Davidson, 1998; Wanderley, 2002).

Physiological factors also played a part. Movement variation occurred in the parts of the body furthest from the keyboard for these were the areas with the highest degrees of freedom. Areas close to the keyboard, particularly the left wrist and middle finger, did not vary in movement with more expression because of their role in sound production. This contrasts with the expressive motion found in the Wanderley et al.'s (2005) study. In that study, clarinettists used body sway and knee bending as both a means for individual expression and as the authors hypothesized, a way to maintain global timing.

This finding highlights the fact that expressive gestures are dependent on the performer's environment. The actual shape of the instrument dictates what movement possibilities are available for expressive ancillary movements and where their limitations lie. This is relevant for the embodied approach to music cognition, which emphasizes that action–perception couplings are what guide musical experiences (Jensenius, 2007). It could be, for instance, that musical experiences and meaning are reliant on corporal articulations, which are in turn heavily influenced by the environment, or instrument (Leman, 2008). Interview data revealed that participants often used their imagination to situate themselves in different environments so they could more easily adopt a different manner of playing for each condition. One participant performed deadpan by pretending to be a morose student bored with the task, while another dramatized his movements for playing under the exaggerated condition. There was undoubtedly a connection between this forethought and the ensuing musical experiences.

Musical expression can be influenced by the dynamic markings written on the musical score. However, these findings show that musicians' expressive movements are also influenced by factors such as the shape of the instrument, and the mental imagery they experience while performing.

The inclusion of these physiological and individual factors into the study allowed for a more complete investigation of musical expression.

This study also provided evidence that pianists recognize an inherent association between playing without expression and playing without extra movements. This point is outlined by the lack of significant differences in the amount of movement between the immobile and deadpan performances. Related to this, is the similarity in duration of the immobile and normal performances. The latter finding is interesting as it contradicts findings in past research, such as Wanderley et al. (2005). Whereas clarinettists predominantly performed at a quicker tempo when restricting their movements, pianists' timings in this condition equalled the timing in normal performances. The instructions for the immobile condition were to play with a normal amount of expression but with restricted movements. The fact that the timings between these two conditions were so similar point to the fact that (a) the restriction of movement did not affect the performers' sense of timing and (b) they made use of timing as a parameter of expression.

In conclusion, this study offers new insights into the relationship between pianists' body movements, their expressive intentions, and structural elements of the music being played. The results support and extend previous work in this area, and further highlight the pivotal role of bodily movement as an indicator of expression in music performance.

Funding

This work was supported by the Academy of Finland (Project Numbers 7118616 and 131356).

Notes

1. Motor programs are representations of actions stored in memory, which are translated into movement sequences (Palmer, 1997; Shaffer, 1981). While a motor program allows for the repeatability of a sequence of actions, its parameters allow for the sequence to shift temporally and dynamically. This is the mechanism that allows musicians to perform, for example, the same piece with varying levels of expressivity.
2. At first, it was thought the varied group of participants would warrant an analysis focused on differences in pianists of varying musical ability and level. However, early versions of our analysis showed nonsignificant differences between individual pianists. This, and the relatively small pool of participants, persuaded us to focus on the participants as a group despite their disparate backgrounds and musical experience.

References

Blacking, J. (1973). *How musical is man?* Seattle, WA: University of Washington Press.

Broughton, M., & Stevens, C. (2009). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, *37*(2), 137–153.

Cadoz, C., & Wanderley, M. M. (2000). Gesture: Music. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* [CD]. Paris, France: IRCAM/Centre Pompidou.

Caramiaux, B., Bevilacqua, F., & Schnell, N. (2010). Towards a gesture-sound cross-modal analysis. In S. Kopp & I. Wachsmuth (Eds.), *Gesture in embodied communication and human–computer interaction* (pp. 158–170). Berlin, Germany: Springer-Verlag.

Castellano, G., Mortillaro, M., Camurri, A., Volpe, G., & Scherer, K. (2008). Automated analysis of body movement in emotionally expressive piano performances. *Music Perception*, *26*(2), 103–120.

Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, *51*(1), 2–9.

Clarke, E. F., & Davidson, J. W. (1998). The body in performance. In W. Thomas (Ed.), *Composition-performance-reception* (pp. 74–92). Aldershot, UK: Ashgate Press.

Cross, I. (2003). Music and biocultural evolution. In M. Clayton, T. Herbert & R. Middleton (Eds.), *The cultural study of music: A critical introduction* (pp. 19–30). New York, NY: Routledge.

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

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

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies*, *26*(6), 279–301.

Davidson, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music*, *35*, 381–401.

Delalande, F. (1988). Le gestique de Gould [The gestures of Gould]. In G. Guertin (Ed.), *Glenn Gould Pluriel [Glenn Gould: Plural]* (pp. 85–111). Verdun, Canada: Louise Courteau.

Ekman, P., & Friesen, W. (1969). The repertoire of nonverbal behaviour: Categories, origins uage and coding. *Semiotica*, *1*(1), 49–98.

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*(1), 68–91.

Godøy, R. I. (2008). Reflections on chunking in music. In A. Schneider (Ed.), *Systematic and comparative musicology: Concepts, methods, findings* (pp. 117–132). Frankfurt am Main, Germany: Peter Lang.

Jensenius, A. R. (2007). Action–sound: Developing methods and tools to study music-related bodily movement. Unpublished doctoral dissertation, University of Oslo.

Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). Concepts and methods in research on music-related gestures. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement and meaning* (pp. 25–48). New York, NY: Routledge.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1797–1813.

Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perceptive on music performance. *Psychology of Music*, *31*(3), 273–302.

Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.

Levitin, D. J. (2008). *The world in six songs: How the musical brain created human nature*. New York, NY: Dutton.

Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: An automated feature-extraction analysis. *Music Perception*, *24*(2), 189–199.

MATLAB. (2009). MATLAB R2009b (Version 7.9.0) [Computer software]. Natick, MA: The Mathworks.

McNeil, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeil, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(12), 331–346.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115–138.

Palmer, C., Koopmans, E., Loehr, J. D., & Carter, C. (2009). Movement-related feedback and temporal accuracy in clarinet performance. *Music Perception*, *26*(5), 439–450.

Pro Tools. (2010). Pro Tools LE (Version 8.0.4) [Computer software]. Burlington, MA: Avid Technology.

Qualisys. (2011). Qualisys Track Manager (Version 2.6) [Computer software]. Gothenburg, Sweden: Qualisys Motion Capture Systems.

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

Schutz, M. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*, 888–897.

Schutz, M. (2008). Seeing music? What musicians need to know about vision. *Empirical Musicology Review*, *3*(3), 83–108.

Shaffer, L. H. (1981). Performances of Chopin, Bach, and Barktok: Studies in motor programming. *Cognitive Psychology*, *13*(3), 326–376.

Sundberg, J., Friberg, A., & Bresin, R. (2003). Attempts to reproduce a pianist's expressive timing with director musices performance rules. *Journal of New Music Research*, *32*(3), 317–325.

Thompson, M. R., & Luck, G. (2008). Exploring relationships between expressive and structural elements of music and pianists' gestures. In E. Cambouropoulos, R. Parncutt, M. Solomos, D. Stefanou, & C. Tsougras (Eds.), *Proceedings of the 4th Conference on Interdisciplinary Musicology, (CIM08)*. Thessaloniki, Greece: Aristotle University of Thessaloniki.

Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception*, *3*(1), 33–58.

Toiviainen, P., & Burger, B. (2010). *MoCap toolbox manual*. Retrieved from http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual

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

Verron, C. (2005). *Traitement et visualisation de données gesturalles captées par Optotrak* [*Processing and visualizing gesture data captured by Optotrak*]. Unpublished lab report, McGill University, Input Devices and Music Interaction Laboratory.

Vines, B. W., Wanderley, M. M., Krumhansl, C. L., Nuzzo, R. L., & Levitin, D. J. (2004). Performance gestures of musicians: What structural and emotional information do they convey? In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in human–computer interaction* (pp. 468–478). Berlin, Germany: Springer-Verlag.

Wanderley, M. M. (2002). Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and sign language in human–computer interaction* (pp. 241–253). Berlin, Germany: Springer Verlag.

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*(1), 97–113.

JYVÄSKYLÄ STUDIES IN HUMANITIES

1 Kostiainen, Emma, Viestintä ammattiosaamisen ulottuvuutena. - Communication as a dimension of vocational competence. 305 p. Summary 4 p. 2003.

2 Seppälä, Antti, Todellisuutta kuvaamassa – todellisuutta tuottamassa. Työ ja koti television ja vähän radionkin uutisissa. - Describing reality – producing reality. Discourses of work and home in television and on a small scale in radio news. 211 p. Summary 3 p. 2003.

3 Gerlander, Maija, Jännitteet lääkärin ja potilaan välisessä viestintäsuhteessa. - Tensions in the doctor-patient communication and relationship. 228 p. Summary 6 p. 2003.

4 Lehikoinen, Taisto, Religious media theory - Understanding mediated faith and christian applications of modern media. - Uskonnollinen mediateoria: Modernin median kristilliset sovellukset. 341 p. Summary 5 p. 2003.

5 Jarva, Vesa, Venäläisperäisyys ja ekspressiivisyys suomen murteiden sanastossa. - Russian influence and expressivity in the lexicon of Finnish dialects. 215 p. 6 p. 2003.

6 Uskali, Turo, "Älä kirjoita itseäsi ulos" Suomalaisen Moskovan-kirjeenvaihtajuuden alkutaival 1957–1975. - "Do not write yourself out" The beginning of the Finnish Moscow-correspondency in 1957–1975. 484 p. Summary 4 p. 2003.

7 Valkonen, Tarja, Puheviestintätaitojen arviointi. Näkökulmia lukioikäisten esiintymis- ja ryhmätaitoihin. - Assessing speech communication skills. Perspectives on presentation and group communication skills among upper secondary school students. 310 p. Summary 7 p. 2003.

8 Tampere, Kaja, Public relations in a transition society 1989-2002. Using a stakeholder approach in organisational communications and relation analyses. 137 p. 2003.

9 Eerola, Tuomas, The dynamics of musical expectancy. Cross-cultural and statistical approaches to melodic expectations. - Musiikillisten odotusten tarkastelu kulttuurienvälisten vertailujen ja tilastollisten mallien avulla. 84 p. (277 p.) Yhteenveto 2 p. 2003.

10 Paananen, Pirkko, Monta polkua musiikkiin. Tonaalisen musiikin perusrakenteiden kehittyminen musiikin tuottamis- ja improvisaatiotehtävissä ikävuosina 6–11. - Many paths to music. The development of basic structures of tonal music in music production and improvisation at the age of 6–11 years. 235 p. Summary 4 p. 2003.

11 Laaksamo, Jouko, Musiikillisten karakterien metamorfoosi. Transformaatio- ja metamorfoosiprosessit Usko Meriläisen tuotannossa vuosina 1963–86. - "Metamorphosis of musical characters". Transformation and metamorphosis processes in the works of Usko Meriläinen during 1963–86. 307 p. Summary 3 p. 2004.

12 Rautio, Riitta, *Fortspinnungstypus* Revisited. Schemata and prototypical features in J. S. Bach's Minor-Key Cantata Aria Introductions. - Uusi katsaus kehitysmuotoon. Skeemat ja prototyyppiset piirteet J. S. Bachin kantaattien molliaarioiden alkusoitoissa. 238 p. Yhteenveto 3 p. 2004.

13 Mäntylä, Katja, Idioms and language users: the effect of the characteristics of idioms on their recognition and interpretation by native and non-native speakers of English. - Idiomien ominaisuuksien vaikutus englannin idiomien ymmärtämiseen ja tulkintaan syntyperäisten ja suomea äidinkielenään puhuvien näkökulmasta. 239 p. Yhteenveto 3 p. 2004.

14 Mikkonen, Yrjö, On conceptualization of music. Applying systemic approach to musicological concepts, with practical examples of music theory and analysis. - Musiikin käsitteellistämisestä. Systeemisen tarkastelutavan soveltaminen musikologisiin käsitteisiin sekä käytännön esimerkkejä musiikin teoriasta ja analyysistä. 294 p. Yhteenveto 10 p. 2004.

15 Holm, Jan-Markus, Virtual violin in the digital domain. Physical modeling and model-based sound synthesis of violin and its interactive application in virtual environment. - Virtuaalinen viulu digitaalisella alueella. Viulun fysikaalinen mallintaminen ja mallipohjainen äänisynteesi sekä sen vuorovaikutteinen soveltaminen virtuaalitodellisuus ympäristössä. 74 p. (123 p.) Yhteenveto 1 p. 2004.

16 Kemp, Chris, Towards the holistic interpretation of musical genre classification. - Kohti musiikin genreluokituksen kokonaisvaltaista tulkintaa. 302 p. Yhteenveto 1 p. 2004.

17 Leinonen, Kari, Finlandssvenskt sje-, tje- och s-ljud i kontrastiv belysning. 274 p. Yhteenveto 4 p. 2004.

18 Mäkinen, Eeva, Pianisti cembalistina. Cembalotekniikka cembalonsoittoa aloittavan pianistin ongelmana. - The Pianist as cembalist. Adapting to harpsichord technique as a problem for pianists beginning on the harpsichord. 189 p. Summary 4 p. 2004.

19 Kinnunen, Mauri, Herätysliike kahden kulttuurin rajalla. Lestadiolaisuus Karjalassa 1870-1939. - The Conviction on the boundary of two cultures. Laestadianism in Karelia in 1870-1939. 591 p. Summary 9 p. 2004.

20 Лилия Сиберг, "БЕЛЫЕ ЛИЛИИ". ГЕНЕЗИС ФИНСКОГО МИФА В БОЛГАРИИ. РОЛЬ РУССКОГО ФЕННОИЛЬСТВА. ФИНСКО-БОЛГАРСКИЕ КОНТАКТЫ И ПОСРЕДНИКИ С КОНЦА XIX ДО КОНЦА XX ВЕКА. 284 с. - "Belye lilii". Genezis finskogo mifa v Bolgarii. Rol' russkogo fennoil'stva. Finsko-bolgarskie kontakty i posredniki s konca XIX do konca XX veka. 284 p. Yhteenveto 2 p. 2004.

JYVÄSKYLÄ STUDIES IN HUMANITIES

21 Fuchs, Bertold, Phonetische Aspekte einer Didaktik der Finnischen Gebärdensprache als Fremdsprache. - Suomalainen viittomakieli vieraana kielenä. Didaktinen fonetiikka. 476 p. Yhteenveto 14 p. 2004.

22 Jääskeläinen, Petri, Instrumentatiivisuus ja nykysuomen verbinjohto. Semanttinen tutkimus. - Instrumentality and verb derivation in Finnish. A semantic study. 504 p. Summary 5 p. 2004.

23 Mertanen Tomi, Kahdentoista markan kapina? Vuoden 1956 yleislakko Suomessa. - A Rebellion for Twelve Marks? – The General Strike of 1956 in Finland. 399 p. Summary 10 p. 2004.

24 Malherbe, Jean-Yves, L'œuvre de fiction en prose de Marcel Thiry : une lecture d'inaboutissements. 353 p. Yhteenveto 1 p. 2004.

25 Kuhna, Matti, Kahden maailman välissä. Marko Tapion *Arktinen hysteria* Väinö Linnan haastajana. - Between two worlds. Marko Tapio's Arktinen hysteria as a challenger to Väinö Linna. 307p. Summary 2 p. 2004.

26 Valtonen, Heli, Minäkuvat, arvot ja mentaliteetit. Tutkimus 1900-luvun alussa syntyneiden toimihenkilönaisten omaelämäkerroista. - Self-images, values and mentalities. An autobiographical study of white collar women in twentieth century Finland. 272 p. Summary 6 p. 2004.

27 Pusztai, Bertalan, Religious tourists. Constructing authentic experiences in late modern hungarian catholicism. - Uskontoturistit. Autenttisen elämyksen rakentaminen myöhäismodernissa unkarilaisessa katolisuudessa. 256 p. Yhteenveto 9 p. Summary in Hungarian 9 p. 2004.

28 Pääjoki, Tarja, Taide kulttuurisena kohtaamispaikkana taidekavatuksessa. - The arts as a place of cultural encounters in arts education. 125 p. Summary 3 p. 2004.

29 Juppi, Pirita, "Keitä me olemme? Mitä me haluamme?" Eläinoikeusliike määrittelykamppailun, marginalisoinnin ja moraalisen paniikin kohteena suomalaisessa sanomalehdistössä. - "Who are we? What do we want?" The animal rights movement as an object of discursive struggle, marginalization and moral panic in Finnish newspapers. 315 p. Summary 6 p. 2004.

30 Holmberg, Jukka, Etusivun politiikkaa. Yhteiskunnallisten toimijoiden representointi suomalaisissa sanomalehtiuutisissa 1987–2003. - Front page politics. Representation of societal actors in Finnish newspapers' news articles in 1987-2003. 291 p. Summary 2 p. 2004.

31 Lagerblom, Kimmo, Kaukana Kainuussa, valtaväylän varrella. Etnologinen tutkimus Kontiomäen rautatieläisyhteisön elinkaaresta 1950 – 1972. - Far, far away, nearby a main passage. An ethnological study of the life spans of Kontiomäki railtown 1950 – 1972. 407 p. Summary 2 p. 2004.

32 Hakamäki, Leena, Scaffolded assistance provided by an EFL teacher during whole-class interaction. - Vieraan kielen opettajan antama oikea-aikainen tuki luokkahuoneessa. 331 p. Yhteenveto 7 p. 2005.

33 Viergutz, Gudrun, Beiträge zur Geschichte des Musikunterrichts an den Gelehrtenschulen der östlichen Ostseeregion im 16. und 17. Jahrhundert. - Latinankoulujen musiikinopetuksen historiasta itäisen Itämeren rannikkokaupungeissa 1500- ja 1600-luvuilla. 211 p. Yhteenveto 9 p. 2005.

34 Nikula, Kaisu, Zur Umsetzung deutscher Lyrik in finnische Musik am Beispiel Rainer Maria Rilke und Einojuhani Rautavaara. - Saksalainen runous suomalaisessa musiikissa, esimerkkinä Rainer Maria Rilke ja Einojuhani Rautavaara. 304 p. Yhteenveto 6 p. 2005.

35 Syvänen, Kari, Vastatunteiden dynamiikka musiikkiterapiassa. - Counter emotions dynamics in music therapy. 186 p. Summary 4 p. 2005.

36 Eloranta, Jari & Ojala, Jari (eds), East-West trade and the cold war. 235 p. 2005.

37 Hiltunen, Kaisa, Images of time, thought and emotions: Narration and the spectator's experience in Krzysztof Kieslowski 's late fiction films. - Ajan, ajattelun ja tunteiden kuvia. Kerronta ja katsojan kokemus Krzysztof Kieslowskin myöhäisfiktiossa. 203 p. Yhteenveto 5 p. 2005.

38 Ahonen, Kalevi, From sugar triangle to cotton triangle. Trade and shipping between America and Baltic Russia, 1783-1860. 572 p. Yhteenveto 9 p. 2005.

39 Utriainen, Jaana, A gestalt music analysis. Philosophical theory, method, and analysis of Iegor Reznikoff's compositions. - Hahmoperustainen musiikkianalyysi. Hahmofilosofinen teoria, metodi ja musiikkianalyysi Iégor Reznikoffin sävellyksistä. 222 p. Yhteenveto 3 p. 2005.

40 Murtorinne, Annamari, *Tuskan hauskaa*! Tavoitteena tiedostava kirjoittaminen. Kirjoittamisprosessi peruskoulun yhdeksännellä luokalla. - Painfully fun! Towards reflective writing process. 338 p. 2005.

41 Tunturi, Anna-Riitta, Der Pikareske Roman als Katalysator in Geschichtlichen Abläufen. Erzählerische Kommunikationsmodelle in *Das Leben des Lazarillo von Tormes*, bei Thomas Mann und in EinigenFinnischen Romanen. 183 p. 2005.

42 Luoma-aho, Vilma, Faith-holders as Social Capital of Finnish Public Organisations. - Luottojoukot – Suomalaisten julkisten organisaatioiden sosiaalista pääomaa. 368 p. Yhteenveto 8 p. 2005.

JYVÄSKYLÄ STUDIES IN HUMANITIES

43 Penttinen, Esa Martti, Kielioppi virheiden varjossa. Kielitiedon merkitys lukion saksan kieliopin opetuksessa. - Grammar in the shadow of mistakes. The role of linguistic knowledge in general upper secondary school German grammar instruction. 153 p. Summary 2 p. Zusammenfassung 3 p. 2005.

44 Kaivapalu, Annekatrin, Lähdekieli kielenoppimisen apuna. - Contribution of L1 to foreign language acquisition. 348 p. Summary 7 p. 2005.

45 Salavuo, Miikka,Verkkoavusteinen opiskelu yliopiston musiikkikasvatuksen opiskelukulttuurissa - Network-assisted learning in the learning culture of university music education. 317 p. Summary 5 p. 2005.

46 Maijala, Juha, Maaseutuyhteisön kriisi-1930-luvun pula ja pakkohuutokaupat paikallisena ilmiönä Kalajokilaaksossa. - Agricultural society in crisis – the depression of the 1930s and compulsory sales as a local phenomenon in the basin of the Kalajoki-river. 242 p. Summary 4 p. 2005.

47 Jouhki, Jukka, Imagining the Other. Orientalism and occidentalism in Tamil-European relations in South India. -Tulkintoja Toiseudesta. Orientalismi ja oksidentalismi tamileiden ja eurooppalaisten välisissä suhteissa Etelä-Intiassa. 233 p. Yhteenveto 2 p. 2006.

48 Lehto, Keijo, Aatteista arkeen. Suomalaisten seitsenpäiväisten sanomalehtien linjapapereiden synty ja muutos 1971–2005. - From ideologies to everyday life. Editorial principles of Finnish newspapers, 1971–2005. 499 p. Summary 3 p. 2006.

49 Valtonen, Hannu, Tavallisesta kuriositeetiksi. Kahden Keski-Suomen Ilmailumuseon Messerschmitt Bf 109 -lentokoneen museoarvo. - From Commonplace to curiosity – The Museum value of two Messerschmitt Bf 109 -aircraft at the Central Finland Aviation Museum. 104 p. 2006.

50 Kallinen, Kari, Towards a comprehensive theory of musical emotions. A multi-dimensional research approach and some empirical findings. - Kohti kokonaisvaltaista teoriaa musiikillisista emootioista. Moniulotteinen tutkimuslähestymistapa ja empiirisiä havaintoja. 71 p. (200 p.) Yhteenveto 2 p. 2006.

51 Iskanius, Sanna, Venäjänkielisten maahanmuuttajaopiskelijoiden kieli-identiteetti. - Language and identity of Russian-speaking students in Finland. 264 p. Summary 5 p. Реферат 6 с. 2006.

52 Heinänen, Seija, Käsityö – taide – teollisuus. Näkemyksiä käsityöstä taideteollisuuteen 1900-luvun alun ammatti- ja aikakauslehdissä. - Craft – Art – Industry: From craft to industrial art in the views of magazines and trade publications of the early 20th Century. 403 p. Summary 7 p. 2006.

53 Kaivapalu, Annekatrin & Pruuli, Külvi (eds), Lähivertailuja 17. - Close comparisons. 254 p. 2006.

54 Alatalo, Pirjo, Directive functions in intra-corporate cross-border email interaction. - Direktiiviset funktiot monikansallisen yrityksen englanninkielisessä sisäisessä sähköpostiviestinnässä. 471 p. Yhteenveto 3 p. 2006.

55 Kisantal, Tamás, „…egy tömegmészárlásról mi értelmes dolgot lehetne elmondani?" Az ábrázolásmód mint történelemkoncepció a holokauszt-irodalomban. - "...there is nothing intelligent to say about a massacre". The representational method as a conception of history in the holocaust-literature. 203 p. Summary 4 p. 2006.

56 Matikainen, Satu, Great Britain, British Jews, and the international protection of Romanian Jews, 1900-1914: A study of Jewish diplomacy and minority rights. - Britannia, Britannian juutalaiset ja Romanian juutalaisten kansainvälinen suojelu, 1900–1914: Tutkimus juutalaisesta diplomatiasta ja vähemmistöoikeuksista. 237 p. Yhteenveto 7 p. 2006.

57 Hänninen, Kirsi, Visiosta toimintaan. Museoiden ympäristökasvatus sosiokulttuurisena jatkumona, säätelymekanismina ja innovatiivisena viestintänä. - From vision to action. Environmental education in museums as a socio-cultural continuum, regulating mechanism, and as innovative communication 278 p. Summary 6 p. 2006.

58 Joensuu, Sanna, Kaksi kuvaa työntekijästä. Sisäisen viestinnän opit ja postmoderni näkökulma. - Two images of an employee; internal communication doctrines from a postmodern perspective. 225 p. Summary 9 p. 2006.

59 Koskimäki, Jouni, Happiness is… a good transcription - Reconsidering the Beatles sheet music publications. - Onni on… hyvä transkriptio – Beatles-nuottijulkaisut uudelleen arvioituna. 55 p. (320 p. + CD). Yhteenveto 2 p. 2006.

60 Hietaharju, Mikko, Valokuvan voi repiä. Valokuvan rakenne-elementit, käyttöympäristöt sekä valokuvatulkinnan syntyminen. - Tearing a photograph. Compositional elements, contexts and the birth of the interpretation. 255 p. Summary 5 p. 2006.

61 Jämsänen, auli, Matrikkelitaiteilijaksi valikoituminen. Suomen Kuvaamataiteilijat -hakuteoksen (1943) kriteerit. - Prerequisites for being listed in a biographical encyclopedia criteria for the Finnish Artists Encyclopedia of 1943. 285 p. Summary 4 p. 2006.

62 Hokkanen, Markku, Quests for Health in Colonial Society. Scottish missionaries and medical culture in the Northern Malawi region, 1875-1930. 519 p. Yhteenveto 9 p. 2006.

63 Ruuskanen, Esa, Viholliskuviin ja viranomaisiin vetoamalla vaiennetut työväentalot. Kuinka Pohjois-Savon Lapuan liike sai nimismiehet ja maaherran sulkemaan 59 kommunistista työväentaloa Pohjois-Savossa vuosina 1930–1932. - The workers' halls closed by scare-mongering and the use of special powers by the authorities. 248 p. Summary 5 p. 2006.

64 Vardja, Merike, Tegelaskategooriad ja tegelase kujutamise vahendid Väinö Linna romaanis "Tundmatu sõdur". - Character categories and the means of character representation in Väinö Linna's Novel *The Unknown Soldier*. 208 p. Summary 3 p. 2006.

65 Takáts, József, Módszertani berek. Írások az irodalomtörténet-írásról. - The Grove of Methodology. Writings on Literary Historiography. 164 p. Summary 3 p. 2006.

66 Mikkola, Leena, Tuen merkitykset potilaan ja hoitajan vuorovaikutuksessa. - Meanings of social support in patient-nurse interaction. 260 p. Summary 3 p. 2006.

67 Saarikallio, Suvi, Music as mood regulation in adolescence. - Musiikki nuorten tunteiden säätelynä. 46 p. (119 p.) Yhteenveto 2 p. 2007.

68 Hujanen, Erkki, Lukijakunnan rajamailla. Sanomalehden muuttuvat merkitykset arjessa. - On the fringes of readership. The changing meanings of newspaper in everyday life. 296 p. Summary 4 p. 2007.

69 Tuokko, Eeva, Mille tasolle perusopetuksen englannin opiskelussa päästään? Perusopetuksen päättövaiheen kansallisen arvioinnin 1999 eurooppalaisen viitekehyksen taitotasoihin linkitetyt tulokset. - What level do pupils reach in English at the end of the comprehensive school? National assessment results linked to the common European framework. 338 p. Summary 7 p. Sammanfattning 1 p. Tiivistelmä 1 p. 2007.

70 Tuikka, Timo, "Kekkosen konstit". Urho Kekkosen historia- ja politiikkakäsitykset teoriasta käytäntöön 1933–1981. - "Kekkonen´s way". Urho Kekkonen's conceptions of history and politics from theory to practice, 1933–1981 413 p. Summary 3 p. 2007.

71 Humanistista kirjoa. 145 s. 2007.

72 Nieminen, Lea, A complex case: a morphosyntactic approach to complexity in early child language. 296 p. Tiivistelmä 7 p. 2007.

73 Torvelainen, Päivi, Kaksivuotiaiden lasten fonologisen kehityksen variaatio. Puheen ymmärrettävyyden sekä sananmuotojen tavoittelun ja tuottamisen tarkastelu. - Variation in phonological development of two-year-old Finnish children. A study of speech intelligibility and attempting and production of words. 220 p. Summary 10 p. 2007.

74 Siitonen, Marko, Social interaction in online multiplayer communities. - Vuorovaikutus verkkopeliyhteisöissä. 235 p. Yhteenveto 5 p. 2007.

75 Stjernvall-Järvi, Birgitta, Kartanoarkkitehtuuri osana Tandefelt-suvun elämäntapaa. - Manor house architecture as part of the Tandefelt family´s lifestyle. 231 p. 2007.

76 Sulkunen, Sari, Text authenticity in international reading literacy assessment. Focusing on PISA 2000. - Tekstien autenttisuus kansainvälisissä lukutaidon arviointitutkimuksissa: PISA 2000. 227 p. Tiivistelmä 6 p. 2007.

77 Kőszeghy, Péter, Magyar Alkibiadés. Balassi Bálint élete. - The Hungarian Alcibiades. The life of Bálint Balass. 270 p. Summary 6 p. 2007.

78 Mikkonen, Simo, State composers and the red courtiers - Music, ideology, and politics in the Soviet 1930s - Valtion säveltäjiä ja punaisia hoviherroja. Musiikki, ideologia ja politiikka 1930-luvun Neuvostoliitossa. 336 p. Yhteenveto 4 p. 2007.

79 Sivunen, Anu, Vuorovaikutus, viestintäteknologia ja identifioituminen hajautetuissa tiimeissä. - Social interaction, communication technology and identification in virtual teams. 251 p. Summary 6 p. 2007.

80 Lappi, Tiina-Riitta, Neuvottelu tilan tulkinnoista. Etnologinen tutkimus sosiaalisen ja materiaalisen ympäristön vuorovaikutuksesta jyväskyläläisissä kaupunkipuhunnoissa. - Negotiating urban spatiality. An ethnological study on the interplay of social and material environment in urban narrations on Jyväskylä. 231 p. Summary 4 p. 2007.

81 Huhtamäki, Ulla, "Heittäydy vapauteen". Avantgarde ja Kauko Lehtisen taiteen murros 1961–1965. - "Fling yourself into freedom!" The Avant-Garde and the artistic transition of Kauko Lehtinen over the period 1961–1965. 287 p. Summary 4 p. 2007.

82 Kela, Maria, *Jumalan kasvot* suomeksi. Metaforisaatio ja erään uskonnollisen ilmauksen synty. - God's face in Finnish. Metaphorisation and the emergence of a religious expression. 275 p. Summary 5 p. 2007.

83 Saarinen, Taina, Quality on the move. Discursive construction of higher education policy from the perspective of quality. - Laatu liikkeessä. Korkeakoulupolitiikan diskursiivinen rakentuminen laadun näkökulmasta. 90 p. (176 p.) Yhteenveto 4 p. 2007.

84 Mäkilä, Kimmo, Tuhoa, tehoa ja tuhlausta. Helsingin Sanomien ja New York Timesin ydinaseuutisoinnin tarkastelua diskurssianalyyttisesta näkökulmasta 1945–1998.

- ”Powerful, Useful and Wasteful”. Discourses of Nuclear Weapons in the New York Times and Helsingin Sanomat 1945–1998. 337 p. Summary 7 p. 2007.

85 Kantanen, Helena, Stakeholder dialogue and regional engagement in the context of higher education. - Yliopistojen sidosryhmävuoropuhelu ja alueellinen sitoutuminen. 209 p. Yhteenveto 8 p. 2007.

86 Almonkari, Merja, Jännittäminen opiskelun puheviestintätilanteissa. - Social anxiety in study-related communication situations. 204 p. Summary 4 p. 2007.

87 Valentini, Chiara, Promoting the European Union. Comparative analysis of EU communication strategies in Finland and in Italy. 159 p. (282 p.) 2008.

88 Pulkkinen, Hannu, Uutisten arkkitehtuuri - Sanomalehden ulkoasun rakenteiden järjestys ja jousto. - The Architecture of news. Order and flexibility of newspaper design structures. 280 p. Yhteenveto 5 p. 2008.

89 Meriläinen, Merja, Monenlaiset oppijat englanninkielisessä kielikylpyopetuksessa - rakennusaineita opetusjärjestelyjen tueksi. - Diverse Children in English Immersion: Tools for Supporting Teaching Arrangements. 197 p. 2008.

90 Vares, Mari, The question of Western Hungary/Burgenland, 1918-1923. A territorial question in the context of national and international policy. - Länsi-Unkarin/Burgenlandin kysymys 1918–1923. Aluekysymys kansallisen ja kansainvälisen politiikan kontekstissa. 328 p. Yhteenveto 8 p. 2008.

91 Ala-Ruona, Esa, Alkuarviointi kliinisenä käytäntönä psyykkisesti oireilevien asiakkaiden musiikkiterapiassa – strategioita, menetelmiä ja apukeinoja. – Initial assessment as a clinical procedure in music therapy of clients with mental health problems – strategies, methods and tools. 155 p. 2008.

92 Oravala, Juha, Kohti elokuvallista ajattelua. Virtuaalisen todellisen ontologia Gilles Deleuzen ja Jean-Luc Godardin elokuvakäsityksissä. - Towards cinematic thinking. The ontology of the virtually real in Gilles Deleuze's and Jean-Luc Godard's conceptions of cinema. 184 p. Summary 6 p. 2008.

93 Kecskeméti, István, Papyruksesta megabitteihin. Arkisto- ja valokuvakokoelmien konservoinnin prosessin hallinta. - From papyrus to megabytes: Conservation management of archival and photographic collections. 277 p. 2008.

94 Suni, Minna, Toista kieltä vuorovaikutuksessa. Kielellisten resurssien jakaminen toisen kielen omaksumisen alkuvaiheessa. - Second language in interaction: sharing linguistic resources in the early stage of second language acquisition. 251 p. Summary 9 p. 2008.

95 N. Pál, József, Modernség, progresszió, Ady Endre és az Ady–Rákosi vita. Egy konfliktusos eszmetörténeti pozíció természete és következményei. 203 p. Summary 3 p. 2008.

96 Bartis, Imre, „Az igazság ismérve az, hogy igaz”. Etika és nemzeti identitás Sütő András Anyám könnyű álmot ígér című művében és annak recepciójában. 173 p. Summary 4 p. 2008.

97 Ranta-Meyer, Tuire, Nulla dies sine linea. Avauksia Erkki Melartinin vaikutteisiin, verkostoihin ja vastaanottoon henkilö- ja reseptiohistoriallisena tutkimuksena. - *Nulla dies sine linea*: A biographical and reception-historical approach to Finnish composer Erkki Melartin. 68 p. Summary 6 p. 2008.

98 Koivisto, Keijo, Itsenäisen Suomen kanta-aliupseeriston synty, koulutus, rekrytointitausta ja palvelusehdot. - The rise, education, the background of recruitment and conditions of service of the non-commissioned officers in independent Finland. 300 p. Summary 7 p. 2008.

99 Kiss, Miklós, Between narrative and cognitive approaches. Film theory of non-linearity applied to Hungarian movies. 198 p. 2008.

100 Ruusunen, Aimo, Todeksi uskottua. Kansandemokraattinen Neuvostoliitto-journalismi rajapinnan tulkkina vuosina1964–1973. - Believed to be true. Reporting on the USSR as interpretation of a boundary surface in pro-communist partisan journalism 1964–1973. 311 p. Summary 4 p. 2008.

101 Härmälä, Marita, Riittääkö *Ett ögonblick* näytöksi merkonomilta edellytetystä kielitaidosta? Kielitaidon arviointi aikuisten näyttötutkinnoissa. – Is *Ett ögonblick* a sufficient demonstration of the language skills required in the qualification of business and administration? Language assessment in competence-based qualifications for adults. 318 p. Summary 4 p. 2008.

102 Coelho, Jacques, The vision of the cyclops. From painting to video ways of seeing in the 20th century and through the eyes of Man Ray. 538 p. 2008.

103 Brewis, Kielo, Stress in the multi-ethnic customer contacts of the Finnish civil servants: Developing critical pragmatic intercultural professionals. – Stressin kokemus suomalaisten viranomaisten monietnisissä asiakaskontakteissa: kriittis-pragmaattisen kulttuurienvälisen ammattitaidon kehittäminen. 299 p. Yhteenveto 4 p. 2008.

104 Belik, Zhanna, The Peshekhonovs' Workshop: The Heritage in Icon Painting. 239 p. [Russian]. Summary 7 p. 2008.

105 Moilanen, Laura-Kristiina, Talonpoikaisuus, säädyllisyys ja suomalaisuus 1800- ja 1900-lukujen vaihteen suomenkielisen proosan kertomana. – Peasant values, estate society and the Finnish in late nineteenth- and early

and early twentieth-century narrative literature. 208 p. Summary 3 p. 2008.

106 Päärnilä, Ossi, Hengen hehkusta tietostrategioihin. Jyväskylän yliopiston humanistisen tiedekunnan viisi vuosikymmentä. 110 p. 2008.

107 Kangasniemi, Jukka, Yksinäisyyden kokemisen avainkomponentit Yleisradion tekstitelevision Nuorten palstan kirjoituksissa. - The key components of the experience of loneliness on the Finnish Broadcasting Company's (YLE) teletext forum for adolescents. 388 p. 2008.

108 Gajdó, Tamás, Színháztörténeti metszetek a 19. század végétől a 20. század közepéig. - Segments of theatre history from the end of the 19th century to the middle of the 20th century. 246 p. Summary 2 p. 2008.

109 Catani, Johanna, Yritystapahtuma kontekstina ja kulttuurisena kokemuksena. - Corporate event as context and cultural experience. 140 p. Summary 3 p. 2008.

110 Mahlamäki-Kaistinen, Riikka, Mätänevän velhon taidejulistus. Intertekstuaalisen ja -figuraalisen aineiston asema Apollinairen L'Enchanteur pourrissant teoksen tematiikassa ja symboliikassa. - Pamphlet of the rotten sorcerer. The themes and symbols that intertextuality and interfigurality raise in Apollinaire's prose work L'Enchanteur pourrissant. 235 p. Résumé 4 p. 2008.

111 Pietilä, Jyrki, Kirjoitus, juttu, tekstielementti. Suomalainen sanomalehtijournalismi juttutyyppien kehityksen valossa printtimedian vuosina 1771-2000. - Written Item, Story, Text Element. Finnish print journalism in the light of the development of journalistic genres during the period 1771-2000. 779 p. Summary 2 p. 2008.

112 Saukko, Päivi, Musiikkiterapian tavoitteet lapsen kuntoutusprosessissa. - The goals of music therapy in the child's rehabilitation process. 215 p. Summary 2 p. 2008.

113 Lassila-Merisalo, Maria, Faktan ja fiktion rajamailla. Kaunokirjallisen journalismin poetiikka suomalaisissa aikakauslehdissä. - On the borderline of fact and fiction. The poetics of literary journalism in Finnish magazines. 238 p. Summary 3 p. 2009.

114 Knuutinen, Ulla, Kulttuurihistoriallisten materiaalien menneisyys ja tulevaisuus. Konservoinnin materiaalitutkimuksen heritologiset funktiot. - The heritological functions of materials research of conservation. 157 p. (208 p.) 2009.

115 Niiranen, Susanna, «Miroir de mérite». Valeurs sociales, rôles et image de la femme dans les textes médiévaux des *trobairitz*. - "Arvokkuuden peili". Sosiaaliset arvot, roolit ja naiskuva keskiaikaisissa *trobairitz*-teksteissä. 267 p. Yhteenveto 4 p. 2009.

116 Aro, Mari, Speakers and doers. Polyphony and agency in children's beliefs about language learning. - Puhujat ja tekijät. Polyfonia ja agentiivisuus lasten kielenoppimiskäsityksissä. 184 p. Yhteenveto 5 p. 2009.

117 Jantunen, Tommi, Tavu ja lause. Tutkimuksia kahden sekventiaalisen perusyksikön olemuksesta suomalaisessa viittomakielessä. - Syllable and sentence. Studies on the nature of two sequential basic units in Finnish Sign Language. 64 p. 2009.

118 Särkkä, Timo, Hobson's Imperialism. A Study in Late-Victorian political thought. - J. A. Hobsonin imperialismi. 211 p. Yhteenveto 11 p. 2009.

119 Laihonen, Petteri, Language ideologies in the Romanian Banat. Analysis of interviews and academic writings among the Hungarians and Germans. 51 p. (180 p) Yhteenveto 3 p. 2009.

120 Mátyás, Emese, Sprachlernspiele im DaF-Unterricht. Einblick in die Spielpraxis des finnischen und ungarischen Deutsch-als-Fremdsprache-Unterrichts in der gymnasialen Oberstufe sowie in die subjektiven Theorien der Lehrenden über den Einsatz von Sprachlernspielen. 399 p. 2009.

121 Paraczky, Ágnes, Näkeekö taitava muusikko sen minkä kuulee? Melodiadiktaatin ongelmat suomalaisessa ja unkarilaisessa taidemusiikin ammattikoulutuksessa. - Do accomplished musicians see what they hear? 164 p. Magyar nyelvü összefoglaló 15 p. Summary 4 p. 2009.

122 Elomaa, Eeva, Oppikirja eläköön! Teoreettisia ja käytännön näkökohtia kielten oppimateriaalien uudistamiseen. - Cheers to the textbook! Theoretical and practical considerations on enchancing foreign language textbook design. 307 p. Zusammanfassung 1 p. 2009.

123 Helle, Anna, Jäljet sanoissa. Jälkistrukturalistisen kirjallisuuskäsityksen tulo 1980-luvun Suomeen. - Traces in the words. The advent of the poststructuralist conception of literature to Finland in the 1980s. 272 p. Summary 2 p. 2009.

124 Pimiä, Tenho Ilari, Tähtäin idässä. Suomalainen sukukansojen tutkimus toisessa maailmansodassa. - Setting sights on East Karelia: Finnish ethnology during the Second World War. 275 p. Summary 2 p. 2009.

125 Vuorio, Kaija, Sanoma, lähettäjä, kulttuuri. Lehdistöhistorian tutkimustraditiot Suomessa ja median rakennemuutos. - Message, sender, culture. Traditions of research into the history of the press in Finland and structural change in the media. 107 p. 2009.

126 Bene, Adrián Egyén és közösség. Jean-Paul Sartre *Critique de la raison dialectique* című műve a magyar recepció tükrében. - Individual and community. Jean-Paul Sartre's

Critique of dialectical reason in the mirror of the Hungarian reception. 230 p. Summary 5 p. 2009.

127 Drake, Merja, Terveysviestinnän kipupisteitä. Terveystiedon tuottajat ja hankkijat Internetissä. - At the interstices of health communication. Producers and seekers of health information on the Internet. 206 p. Summary 9 p. 2009.

128 Rouhiainen-Neunhäuserer, Maijastiina, Johtajan vuorovaikutusosaaminen ja sen kehittyminen. Johtamisen viestintähaasteet tietoperustaisessa organisaatiossa. - The interpersonal communication competence of leaders and its development. Leadership communication challenges in a knowledge-based organization. 215 p. Summary 9 p. 2009.

129 Vaarala, Heidi, Oudosta omaksi. Miten suomenoppijat keskustelevat nykynovellista? - From strange to familiar: how do learners of Finnish discuss the modern short story? 317 p. Summary 10 p. 2009.

130 Marjanen, Kaarina, The Belly-Button Chord. Connections of pre-and postnatal music education with early mother-child interaction. - Napasointu. Pre- ja postnataalin musiikkikasvatuksen ja varhaisen äiti-vauva-vuorovaikutuksen yhteydet. 189 p. Yhteenveto 4 p. 2009.

131 Bőhm, Gábor, Önéletírás, emlékezet, elbeszélés. Az emlékező próza hermeneutikai aspektusai az önéletírás-kutatás újabb eredményei tükrében. - Autobiography, remembrance, narrative. The hermeneutical aspects of the literature of remembrance in the mirror of recent research on autobiography. 171 p. Summary 5 p. 2009.

132 Leppänen, Sirpa, Pitkänen-Huhta, Anne, Nikula, Tarja, Kytölä, Samu, Törmäkangas, Timo, Nissinen, Kari, Kääntä, Leila, Virkkula, Tiina, Laitinen, Mikko, Pahta, Päivi, Koskela, Heidi, Lähdesmäki, Salla & Jousmäki, Henna, Kansallinen kyselytutkimus englannin kielestä Suomessa: Käyttö, merkitys ja asenteet. - National survey on the English language in Finland: Uses, meanings and attitudes. 365 p. 2009.

133 Heikkinen, Olli, Äänitemoodi. Äänite musiikillisessa kommunikaatiossa. - Recording Mode. Recordings in Musical Communication. 149 p. 2010.

134 Lähdesmäki, Tuuli (Ed.), Gender, Nation, Narration. Critical Readings of Cultural Phenomena. 105 p. 2010.

135 Mikkonen, Inka, "Olen sitä mieltä, että". Lukiolaisten yleisönosastotekstien rakenne ja argumentointi. - "In my opinion…" Structure and argumentation of letters to the editor written by upper secondary school students. 242 p. Summary 7 p. 2010.

136 Nieminen, Tommi, Lajien synty. Tekstilaji kielitieteen semioottisessa metateoriassa. - Origin of genres: Genre in the semiotic metatheory of linguistics. 303 p. Summary 6 p. 2010.

137 Kääntä, Leila, Teacher turn allocation and repair practices in classroom interaction. A multisemiotic perspective. - Opettajan vuoronanto- ja korjauskäytänteet luokkahuonevuorovaikutuksessa: multisemioottinen näkökulma. 295 p. Yhteenveto 4 p. 2010. HUOM: vain verkkoversiona.

138 Saarimäki, Pasi, Naimisen normit, käytännöt ja konfliktit. Esiaviollinen ja aviollinen seksuaalisuus 1800-luvun lopun keskisuomalaisella maaseudulla. - The norms, practices and conflicts of sex and marriage. Premarital and marital sexual activity in rural Central Finland in the late nineteenth century. 275 p. Summary 12 p. 2010.

139 Kuuva, Sari, Symbol, Munch and creativity: Metabolism of visual symbols. - Symboli, Munch ja luovuus – Visuaalisten symboleiden metabolismi. 296 p. Yhteenveto 4 p. 2010.

140 Skaniakos, Terhi, Discoursing Finnish rock. Articulations of identities in the Saimaa-Ilmiö rock documentary. - Suomi-rockin diskursseja. Identiteettien artikulaatioita Saimaa-ilmiö rockdokumenttielokuvassa. 229 p. 2010.

141 Kauppinen, Merja, Lukemisen linjaukset – lukutaito ja sen opetus perusopetuksen äidinkielen ja kirjallisuuden opetussuunnitelmissa. - Literacy delineated – reading literacy and its instruction in the curricula for the mother tongue in basic education. 338 p. Summary 8 p. 2010.

142 Pekkola, Mika, Prophet of radicalism. Erich Fromm and the figurative constitution of the crisis of modernity. - Radikalismin profeetta. Erich Fromm ja modernisaation kriisin figuratiivinen rakentuminen. 271 p. Yhteenveto 2 p. 2010.

143 Kokkonen, Lotta, Pakolaisten vuorovaikutussuhteet. Keski-Suomeen muuttaneiden pakolaisten kokemuksia vuorovaikutussuhteistaan ja kiinnittymisestään uuteen sosiaaliseen ympäristöön. - Interpersonal relationships of refugees in Central Finland: perceptions of relationship development and attachment to a new social environment. 260 p. Summary 8 p. 2010.

144 Kananen, Heli Kaarina, Kontrolloitu sopeutuminen. Ortodoksinen siirtoväki sotien jälkeisessä Ylä-Savossa (1946-1959). - Controlled integration: Displaced orthodox Finns in postwar upper Savo (1946–1959). 318 p. Summary 4 p. 2010.

145 Nissi, Riikka, Totuuden jäljillä. Tekstin tulkinta nuorten aikuisten raamattupiirikeskusteluissa. – In search of the truth. Text interpretation in young adults' Bible study conversations. 351 p. Summary 5 p. 2010.

146 Lilja, Niina, Ongelmista oppimiseen. Toisen aloittamat korjausjaksot kakkoskielisessä keskustelussa. – Other-initiated repair sequences in Finnish second language interactions. 336 p. Summary 8 p. 2010.

147 Váradi, Ildikó, A parasztpolgárosodás „finn útja". Kodolányi János finnországi tevékenysége és finn útirajzai. – The "Finnish Way" of Peasant-Bourgeoization. János Kodolányi's Activity in Finland and His Travelogues on Finland. 182 p. Summary 3 p. 2010.

148 Hankala, Mari, Sanomalehdellä aktiiviseksi kansalaiseksi? Näkökulmia nuorten sanomalehtien lukijuuteen ja koulun sanomalehtiopetukseen. – Active citizenship through newspapers? Perspectives on young people´s newspaper readership and on the use of newspapers in education. 222 p. Summary 5 p. 2011.

149 Salminen, Elina, Monta kuvaa menneisyydestä. Etnologinen tutkimus museokokoelmien yksityisyydestä ja julkisuudesta. – Images of the Past. An ethnological study of the privacy and publicity of museum collections. 226 p. Summary 5 p. 2011. HUOM: vain verkkoversiona.

150 Järvi, Ulla, Media terveyden lähteillä. Miten sairaus ja terveys rakentuvat 2000-luvun mediassa. – Media forces and health sources. Study of sickness and health in the media. 209 p. Summary 3 p. 2011.

151 Ullakonoja, Riikka, Da. Eto vopros! Prosodic development of Finnish students´ read-aloud Russian during study in Russia. – Suomalaisten opiskelijoiden lukupuhunnan prosodinen kehittyminen vaihto-opiskelujakson aikana Venäjällä. 159 p. (208 p.) Summary 5 p. 2011.

152 Marita Vos, Ragnhild Lund, Zvi Reich and Halliki Harro-Loit (Eds), Developing a Crisis Communication Scorecard. Outcomes of an International Research Project 2008-2011 (Ref.). 340 p. 2011.

153 Punkanen, Marko, Improvisational music therapy and perception of emotions in music by people with depression. 60 p. (94 p.) Yhteenveto 1 p. 2011.

154 Di Rosario, Giovanna, Electronic poetry. Understanding poetry in the digital environment. – Elektroninen runous. Miten runous ymmärretään digitaalisessa ympäristössä? 327 p. Tiivistelmä 1 p. 2011.

155 Tuuri, Kai, Hearing Gestures: Vocalisations as embodied projections of intentionality in designing non-speech sounds for communicative functions. – Puheakteissa kehollisesti välittyvä intentionaalisuus apuna ei-kielellisesti viestivien käyttöliittymä-äänien suunnittelussa. 50 p. (200 p.) Yhteenveto 2 p. 2011.

156 Martikainen, Jari, Käsitettävä taidehistoria. Kuvalähtöinen malli taidehistorian opetukseen kuvallisen ilmaisun ammatillisessa perustutkinnossa. – Grasping art history. A picture-based model for teaching art history in the vocational basic degree programme in visual arts. 359 p. Summary 10 p. 2011.

157 Hakanen, Marko, Vallan verkostoissa. Per Brahe ja hänen klienttinsä 1600-luvun Ruotsin valtakunnassa. – Networks of Power: Per Brahe and His Clients in the Sixteenth-Century Swedish Empire. 216 p. Summary 6 p. 2011.

158 Lindström, Tuija Elina, Pedagogisia merkityksiä koulun musiikintunneilla perusopetuksen yläluokkien oppilaiden näkökulmasta. – Pedagogical Meanings in Music Education from the Viewpoint of Students of Junior High Grades 7-9. 215 p. 2011.

159 Anckar, Joanna, Assessing foreign language listening comprehension by means of the multiple-choice format: processes and products. – Vieraan kielen kuullun ymmärtämistaidon mittaaminen monivalintatehtävien avulla: prosesseja ja tuloksia. 308 p. Tiivistelmä 2 p. 2011.

160 Ekström, Nora, Kirjoittamisen opettajan kertomus. Kirjoittamisen opettamisesta kognitiiviselta pohjalta. – The story of writing teacher. Teaching writing from cognitive base. 272 p. Tiivistelmä 4 p. 2011.

161 Huovinen, Mika, Did the east belong to the SS? The realism of the SS demographic reorganisation policy in the light of the germanisation operation of SS- und Polizeiführer Odilo Globocnik. – Kuuluiko Itä SS-järjestölle? SS-järjestön uudelleenjärjestelypolitiikan realistisuus SS- ja poliisijohtaja Odilo Globocnikin germaanistamisoperaation valossa. 268 p. Tiivistelmä 1 p. 2011.

162 Pakkanen, Irene, Käydään juttukauppaa. Freelancerin ja ostajan kohtaamisia journalismin kauppapaikalla. – Let's do story business. Encounters of freelancers and buyers at the marketplace of journalism. 207 p. 2011.

163 Koskela, Heidi, Constructing knowledge: Epistemic practices in three television interview genres. – Tietoa rakentamassa: episteemiset käytänteet kolmessa eri televisiohaastattelugenressä. 68 p. (155 p.) Yhteenveto 3 p. 2011.

164 Pöyhönen, Markku O., Muusikon tietämisen tavat. Moniälykkyys, hiljainen tieto ja musiikin esittämisen taito korkeakoulun instrumenttituntien näkökulmasta. – The ways of knowing of a musician: Multiple intelligences, tacit knowledge and the art of performing seen through instrumental lessons of bachelor and post-graduate students. 256 p. Summary 4 p. 2011.

JYVÄSKYLÄ STUDIES IN HUMANITIES

165 Rautavuoma, Veera, Liberation exhibitions as a commemorative membrane of socialist Hungary. 251 p. Yhteenveto 3 p. 2011.
166 Lehtonen, Kimmo E., Rhetoric of the visual – metaphor in a still image. – Visuaalisen retoriikka – metafora still-kuvan tarkastelussa. 174 p. Yhteenveto 1 p. 2011.
167 Sarkamo, Ville, Karoliinien soturiarvot. Kunnian hallitsema maailmankuva Ruotsin valtakunnassa 1700-luvun alussa. – Carolean warrior values: an honour-dominated worldview in early-eighteenth-century Sweden. 216 p. Summary 11 p. 2011.
168 Rynkänen, Tatjana, Русскоязычные молодые иммигранты в Финляндии – интеграция в контексте обучения и овладения языком. – Russian-speaking immigrant adolescents in Finnish society – integration from the perspective of language and education. 258 p. Tiivistelmä 9 p. Summary 9 p. 2011.
169 Tiainen, Veikko, Vähentäjää vähentämässä. Tehdaspuu Oy puunhankkijana Suomessa. – Tehdaspuu Oy in Finnish Wood Procurement. 236 p. Summary 5 p. 2011.
170 Stolp, Marleena, Taidetta, vastustusta, leikkiä ja työtä? Lasten toimijuus 6–vuotiaiden teatteriprojektissa. – Art, resistance, play and work? Children's agency in a six-year-olds' theatre project. 79 p. (142 p.) 2011.
171 Cools, Carine, Relational dialectics in intercultural couples' relationships. – Kulttuurienvälisten parisuhteiden relationaalinen dialektiikka. 282 p. 2011.
172 Saario, Johanna, Yhteiskuntaopin kieliympäristö ja käsitteet – toisella kielellä opiskelevan haasteet ja tuen tarpeet. – The language environment and concepts in social studies – challenges and need of support for a second language learner. 290 p. Summary 7 p. 2012.
173 Alluri, Vinoo, Acoustic, neural, and perceptual correlates of polyphonic timbre. – Polyfonisen sointivärin hahmottamisen akustiset ja hermostolliset vastineet. 76 p. (141 p.) Yhteenveto 1 p. 2012.
174 Vuoskoski, Jonna Katariina, Emotions represented and induced by music: The role of individual differences. – Yksilöllisten erojen merkitys musiikillisten emootioiden havaitsemisessa ja kokemisessa. 58 p. (132 p.) Tiivistelmä 1 p. 2012.
175 Leinonen, Jukka, The beginning of the cold war as a phenomenon of realpolitik – U.S. secretary of state James F. Byrnes in the field of power politics 1945–1947. – Kylmän sodan synty reaalipoliittisena ilmiönä – James F. Byrnes suurvaltapolitiikan pelikentällä Jaltasta Stuttgartiin 1945–1947. 393 p. Yhteenveto 8 p. 2012.
176 Thompson, Marc, The application of motion capture to embodied music cognition research. - Liikkeenkaappausteknologian soveltaminen kehollisen musiikkikognition tutkimuksessa. 86 p. (165 p.) Yhteenveto 1 p. 2012.