

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Measuring musical engagement

Permalink

<https://escholarship.org/uc/item/73x8c66v>

Author

Leslie, Grace

Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Measuring Musical Engagement

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Music and Cognitive Science

by

Grace Leslie

Committee in charge:

Miller Puckette, Chair
Charles Curtis
Diana Deutsch
Scott Makeig
Steven Schick

2013

Copyright
Grace Leslie, 2013
All rights reserved.

The dissertation of Grace Leslie is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Figures	vi
List of Tables	ix
Acknowledgements	x
Vita	xi
Abstract of the Dissertation	xiii
Chapter 1	Introduction and Theoretical Motivations	1
	1.1 A Definition of Musical Engagement	2
	1.2 Past approaches to measuring emotional response to music	7
	1.3 Proposed Study: Measuring Musical Engagement	9
Chapter 2	An Expressive “Conducting” Experiment	17
	2.1 Method	18
	2.1.1 Demographics	18
	2.1.2 Stimuli	18
	2.1.3 Data Recording	22
	2.1.4 Experiment Design and Presentation of Stimuli	24
	2.1.5 Data Streams Recorded	27
	2.1.6 Additional Participant Data Collected	28
	2.1.7 Motion Capture Processing Pipeline	28
	2.1.8 EEG Processing Pipeline	29
	2.2 Results	35
	2.2.1 Motion Capture Data	35
	2.2.2 Discussion	41
	2.2.3 EEG	43
	2.2.4 Discussion	46
	2.3 Conclusion	51
Chapter 3	Audience Experiment	52
	3.1 Method	53
	3.2 Results	56
	3.2.1 Correlation of Internet and Conductor Responses	56
	3.2.2 Distinguishing Engaged vs. Non-Engaged Conditions	57
	3.3 Conclusion	58

Chapter 4	Conclusion	59
	4.1 Discussion	59
	4.2 Implications	62
	4.3 Future Directions	63
Appendix A	Experiment Narrative Scripts	66
	A.1 Initial Instructions for the Active Music Listening Task	66
	A.2 Guided Relaxation Script	68
	A.3 Conducting Experiment Episode Instructions	69
	A.4 Music Feeling Questionnaire	71
	A.5 Music Experience Questionnaire	72
	A.6 Internet Experiment Video Feedback Form	73
Appendix B	Single Subject Motion Plots	75
Bibliography	91

LIST OF FIGURES

Figure 1.1:	The proposed model of musical engagement involves an interdependence between the mechanisms that support attention, emotion, and action.	2
Figure 1.2:	(a) Cortical brain structures involved in attention, auditory processing, and motor planning. (b) Subcortical brain structures involved in emotion response and motor planning.	4
Figure 1.3:	Experiment Design Overview.	10
Figure 2.1:	Music stimuli used in the conducting experiment plotted along the Valence and Energy dimensions. Two clusters of excerpts emerge, one containing the high energy pieces, and another containing the low energy pieces.	20
Figure 2.2:	A participant wears the full-body motion capture suit with an additional sensor worn on the conducting hand, and a 128-channel EEG system with cables connected to an amplifier worn in a backpack. The participant’s movements are animated as a white dot on the facing screen.	23
Figure 2.3:	Channel locations for one subject wearing the 128-channel MoBI cap.	23
Figure 2.4:	Still from a film shown to the participant illustrating how a music listener can express their musical feeling through gesture. Copyright Warner Brothers, 1968.	24
Figure 2.5:	Each experiment block was divided into training, engaged, and not-engaged conditions. Participants completed a music emotion rating questionnaire at the end of each of the first set of blocks, before a new excerpt was presented.	26
Figure 2.6:	During each excerpt presentation ‘bout,’ the excerpt was preceded by four repetitions of a bandpassed noise ‘whoosh’ sound designed to show the participant the beat of the excerpt.	26
Figure 2.7:	EEG processing pipeline, for individual (left) and group (right) level analysis.	30
Figure 2.8:	Plotted x-y trajectories of the average swings for subject 414 colored by acceleration trajectory, with red representing maximum acceleration, and blue representing minimum acceleration. The average trajectories over the engaged trials are shown in the left column, and not-engaged shown in the right column.	37

Figure 2.9:	(a) Plot of time- and space-normalized trajectories for the engaged condition only, averaged across all subjects, with the global mean removed. The x- and y- axes of the plot correspond to the horizontal and vertical movement of the participants. The x-positions of the second half of the trajectory have been reversed to be able distinguish it from the first half of the trajectory. Each trace is colored by the acceleration along the trajectory path. (b) Global mean subtracted from each within-excerpt mean in order to produce the traces above. (c) time points at which difference between songs is significant at $p < .05$	39
Figure 2.10:	MDS plot of swing data shows that swings cluster similarly to emotion measures of excerpts. Each ball represents the average across all subjects, and is colored by the energy rating for the excerpt as reported in [Eerola and Vuoskoski, 2011].	40
Figure 2.11:	A plot of average Right-Left-Right swing trajectory colored by acceleration difference, masked at a $p < .05$ significance level, shows that there are significant differences between the E and NE conditions in the motion capture data. These differences vary by song. . .	42
Figure 2.12:	The independent components found in each subject’s EEG data were clustered according to equivalent dipole map and spectral information. These topographic projections of each cluster of ICs onto a scalp surface show the areas of the frontal, temporal, parietal, and occipital lobes where the clusters were found.	44
Figure 2.13:	The averaged Event-Related Spectral Perturbation (ERSP) for cluster 4 differs between the engaged and not-engaged trials. The difference between the two ERSP plots, masked at a $p < .05$ significance level, is shown in the third column, revealing low alpha- and theta-synchronization specific to the engaged condition, and time-locked to the swing cycle.	47
Figure 2.14:	The estimated dipoles (blue) and their centroid (red) contributing to cluster 4, as projected onto axial, coronal, and sagittal plane sections from the MNI template.	48
Figure 3.1:	The motion capture data from eight of the Conducting Experiment participants were processed into animations that were uploaded to YouTube.	54
Figure B.1:	Movement Trajectories, Subject 357	76
Figure B.2:	Movement Trajectories, Subject 362	77
Figure B.3:	Movement Trajectories, Subject 370	78
Figure B.4:	Movement Trajectories, Subject 371	79
Figure B.5:	Movement Trajectories, Subject 372	80
Figure B.6:	Movement Trajectories, Subject 374	81

Figure B.7: Movement Trajectories, Subject 375	82
Figure B.8: Movement Trajectories, Subject 378	83
Figure B.9: Movement Trajectories, Subject 408	84
Figure B.10: Movement Trajectories, Subject 411	85
Figure B.11: Movement Trajectories, Subject 412	86
Figure B.12: Movement Trajectories, Subject 413	87
Figure B.13: Movement Trajectories, Subject 414	88
Figure B.14: Movement Trajectories, Subject 415	89
Figure B.15: Movement Trajectories, Subject 416	90

LIST OF TABLES

Table 2.1:	Label, Target Emotion, and Source information for excerpts used in the present study.	19
Table 3.1:	Example experiment block from the Internet audience study.	55
Table 3.2:	Correlation Coefficient (R) between song ratings from the Conducting Experiment and three Audience Experiments.	56

ACKNOWLEDGEMENTS

I would like to extend my thanks to my collaborators at the Swartz Center for Computational Neuroscience, including Alejandro Ojeda, whose talent and hard work is seen in these pages, and Makoto Miyakoshi, who was a patient teacher of all things EEG.

Veronique Larcher and Olivier Warusfel mentored and encouraged me during my first forays into music perception.

I would like to thank my committee chair Miller Puckette, and past and present committee members Steven Schick, Diana Deutsch, Dick Moore, Charles Curtis, and Adrienne Jenik, as well as my family, for their support.

Finally, I will be forever grateful to Scott Makeig for showing me how science can be used as a tool “to penetrate into the essence of all being and significance.”

Chapters 1, 2, 3, and 4 of this dissertation are currently being prepared for submission for publication of the material. Leslie, Grace; Ojeda, Alejandro; Makeig, Scott. “Measuring Musical Engagement Using Expressive Movement and EEG Brain Dynamics.”

VITA

2005	Bachelor of Arts with Honors, Stanford University, Stanford, CA.
2005	Master of Arts, Stanford University, Stanford, CA.
2005-2006	Calit2 Fellow, University of California, San Diego.
2006-2008	Teaching Assistant, Department of Music, University of California, San Diego.
2008-2009	Researcher, Room Acoustics Team, IRCAM, Paris, France.
2009-2013	Graduate Student Researcher, Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California, San Diego.
2013	Doctor of Philosophy, University of California, San Diego, La Jolla, CA.

PUBLICATIONS

Makeig, S., Leslie, G., Mullen, T., Sarma, D., Bigdely-Shamlo, N., Kothe, C. (2011). "First demonstration of a musical emotion BCI." *Affective Computing and Intelligent Interaction*, 487-496.

Makeig, S., Leslie, G., Mullen, T., Sarma, D., Bigdely-Shamlo, N., Kothe, C. (2011) "Analyzing Brain Dynamics of Affective Engagement." *Lecture Notes in Computer Science*, 6975.

Leslie, G., Mullen, T. (2011) MoodMixer: EEG-based Collaborative Sonification. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 296-299.

Leslie, G and Warusfel, O. (2010) "A Cognitive Test of Interactive Music Listening," In *Proceedings of the 2010 International Conference on Music Perception and Cognition*, Seattle.

Leslie, G., Zamborlin, B., Jodlowski, P., Schnell, N. (2010). Grainstick: A collaborative, interactive sound installation. In *Proceedings of the International Computer Music Conference*, New York, NY.

Leslie, G., Schwarz, D., Warusfel, O., Bevilacqua, F., Jodlowski, P. (2009)., "Wavefield Synthesis for Interactive Sound Installations," *Proc. 127th Audio Engineering Society Conference*, Oct. 2009, New York.

Leslie, G., Hassanpour, N. (2008) “A Game Theoretical Model for Musical Interaction,” in *Proceedings of the International Computer Music Conference*, Belfast, Ireland.

Chafe, C., Gurevich, M., Leslie, G., Tyan, S. (2004). “Effect of time delay on ensemble accuracy.” In *Proceedings of the International Symposium on Musical Acoustics* (Vol. 31).

Gurevich, M., Chafe, C., Leslie, G., Tyan, S. (2004). Simulation of networked ensemble performance with varying time delays: Characterization of ensemble accuracy. In *Proceedings of the 2004 International Computer Music Conference*, Miami, USA.

ABSTRACT OF THE DISSERTATION

Measuring Musical Engagement

by

Grace Leslie

Doctor of Philosophy in Music and Cognitive Science

University of California, San Diego, 2013

Miller Puckette, Chair

Currently little is known about the brain dynamics and expressive movements that support musical engagement. We hypothesize that repetitive expressive gestures play an important role in inviting musical engagement in listeners, and that these movements can reveal the feelings experienced by the listener. Furthermore, we hypothesize that brain dynamics supporting these expressive movements play a key role in musical engagement. We trained expert and non-expert participants to communicate the feeling of music they are hearing using simple rhythmic U-shaped hand/arm ‘conducting’ gestures that animate the 2-D movement of a spot of light on a video display while we use body motion capture and EEG to record their movements and brain activity. Periodically we introduced a “not engaged” condition during which a distractor task impedes the participant’s engagement in the engaged listening task. We then asked viewers to rate

the recorded 2-D spot animations of the recorded gestures on a musical emotion rating scale to test to what extent the musical affective experience of the ‘conductors’ can be conveyed by these animations to viewers who do not hear the music. The ratings from the conductor and viewer groups were well correlated, verifying that the affective intent of the conductors’ gestures are experienced by viewers. Statistically significant differences were found in the motion capture and EEG data between the fully engaged condition and the not-engaged condition. A comparison of the EEG data recorded during the two conditions revealed low alpha- and theta-synchronization in the parietal-temporal-occipital (PTO) junction which was specific to the engaged condition, and time-locked to the participants’ expressive movements. The results from the viewer experiment suggest that the feeling intention of the expressive gesture task is communicable through a single point-light display, and that viewers can distinguish engaged performances from not-engaged performances. Our EEG results suggest that brain dynamics supporting engaged music listening, located at the PTO junction, are co-modulated with the expressive, rhythmic movements made by the listener. The fact that we can non-invasively monitor musical engagement gives us a useful and general tool for music perception research, with possible wider applications to music classification, technology, and therapy.

Chapter 1

Introduction and Theoretical Motivations

This thesis examines the behavioral and brain dynamics supporting the experience of musical engagement, defined here as the experience of consciously “entering into” the experience of music as it is heard, imagined, or performed, a condition in which the listener is fully attentive and emotionally attuned to the musical environment and not attentive to extra-musical stimuli or concerns. This may occur (or not) when listeners are actively involved in creatively experiencing the musical stimulus they are hearing or imagining, and also when they are helping shape the musical experience by performing, improvising, or actively modulating the sound environment to enhance their own or others’ musical experience.

The intent of artistic musical performance is to allow or encourage listeners to enter into a musically engaged listening state readily and fully. The musical artist must not only avoid playing notes that are physically ‘wrong’ or out-of-place but must also create a sense in the listeners of an emotionally conducive musical ‘flow’ or ‘pulse,’ as pointed out by Clynes [Clynes, 1977]. The nature of this temporal flow is not understood, though it is intuitively understood by musicians—as evidenced, for example, by the success of conductors who use continuous gestures to control the ‘pulse’ of an orchestral performance, even though the notes the orchestral instruments are playing typically have abrupt onsets and a distinctly pulsatile texture.

Our goal is to better understand the state of musical engagement that can arise

either during “passive” listening or during physically engaged musical performance. Through the experiments outlined below, I attempt to clarify relationships between the experience of musical emotion or affect and physical movement during states of musical engagement. The resulting conclusions about musical engagement may well have broad impact outside the cognitive neuroscience community. New media researchers commonly assume that musical engagement parallels so-called ‘immersion’ in the experience of virtual-reality environments. Testing these assumptions requires a model and measure of musical engagement. A method for modeling brain and body dynamics accompanying musical engagement should have wide-ranging applications in fields of music and new media studies research and development, as well as broad applications to musical learning and therapy.

1.1 A Definition of Musical Engagement

A model of musical engagement is illustrated in Figure 1.1. As presented in this model, musical engagement involves attention, emotion, and action, each of which is dependent on the others, though they employ different areas of the brain and body.

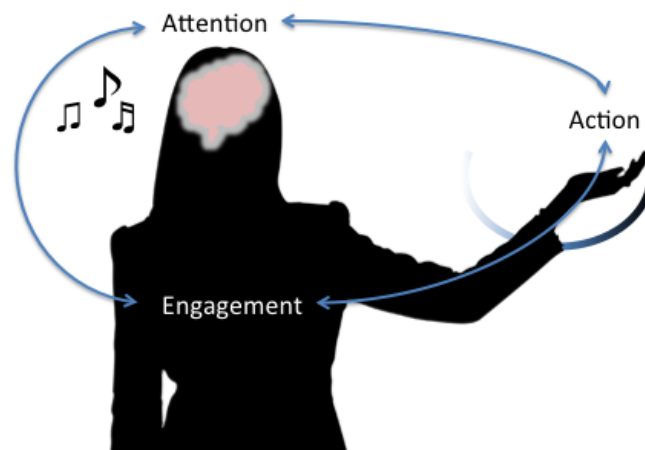


Figure 1.1: The proposed model of musical engagement involves an interdependence between the mechanisms that support attention, emotion, and action.

The musical attention process involves translating the physical properties of an

acoustic signal, found in the frequency and time information, into musical elements such as pitch and pulse. The primary auditory cortex is responsible for extracting these acoustic features, with tonal processing happening primarily in the right hemisphere, and time-related processing in the left hemisphere [Zatorre et al., 2002]. The secondary auditory cortex (located along the superior temporal gyrus, seen in Figure 1.2a) is responsible for higher-order processing of these auditory percepts. The right secondary auditory cortex is involved in the grouping of pitch relationships over time, whereas the left hemisphere has been implicated in the grouping of events based on their durations [Peretz and Zatorre, 2005]. Melodies and harmonies are created by grouping these musical elements based on melody, rhythm, timbre, and spatial location [Deutsch, 1999], much as syllables are combined to create syntax. In fact, the formation of the highest-level musical percepts involves associational areas of the temporal and parietal cortex that combine information from auditory, motor, and visual areas for the comprehension of language, and other complex cognitive functions [Patel, 2003]. These areas are shown in Figure 1.2a.

A musical stream, once attended to, is stored into short-term memory so that expectations can be formed about which musical elements may arise next. Frontal cortical areas (shown in Figure 1.2a) interact with auditory cortex in the holding of pitch sequences in short-term memory. It can also be stored as a long-term perceptual representation, helpful for pieces of music that span multiple movements, and for recalling melodies at later points in time. Inferior frontal areas, circled in Figure 1.2a, and the hippocampus in the limbic system, whose surrounding structures are seen in Figure 1.2b, are crucial for the recall of music sequences, much as they are with other memories. Memory and expectation can, in turn, modulate the emotional engagement of the listener, by one of many mechanisms, which are exhaustively reviewed in an article by Juslin and Vastfall [Juslin and Vastfjall, 2008]. It is likely that any particular music performance induces emotion by a combination of two or more of these mechanisms.

Particularly surprising events, are processed first along a deeper, subcortical emotional pathway starting with the limbic system (primarily the thalamus) and brain stem; these areas are shown in Figure 1.2b. The brain stem is designed to react to sudden events that may be of interest to survival by increasing feelings of arousal and unpleas-

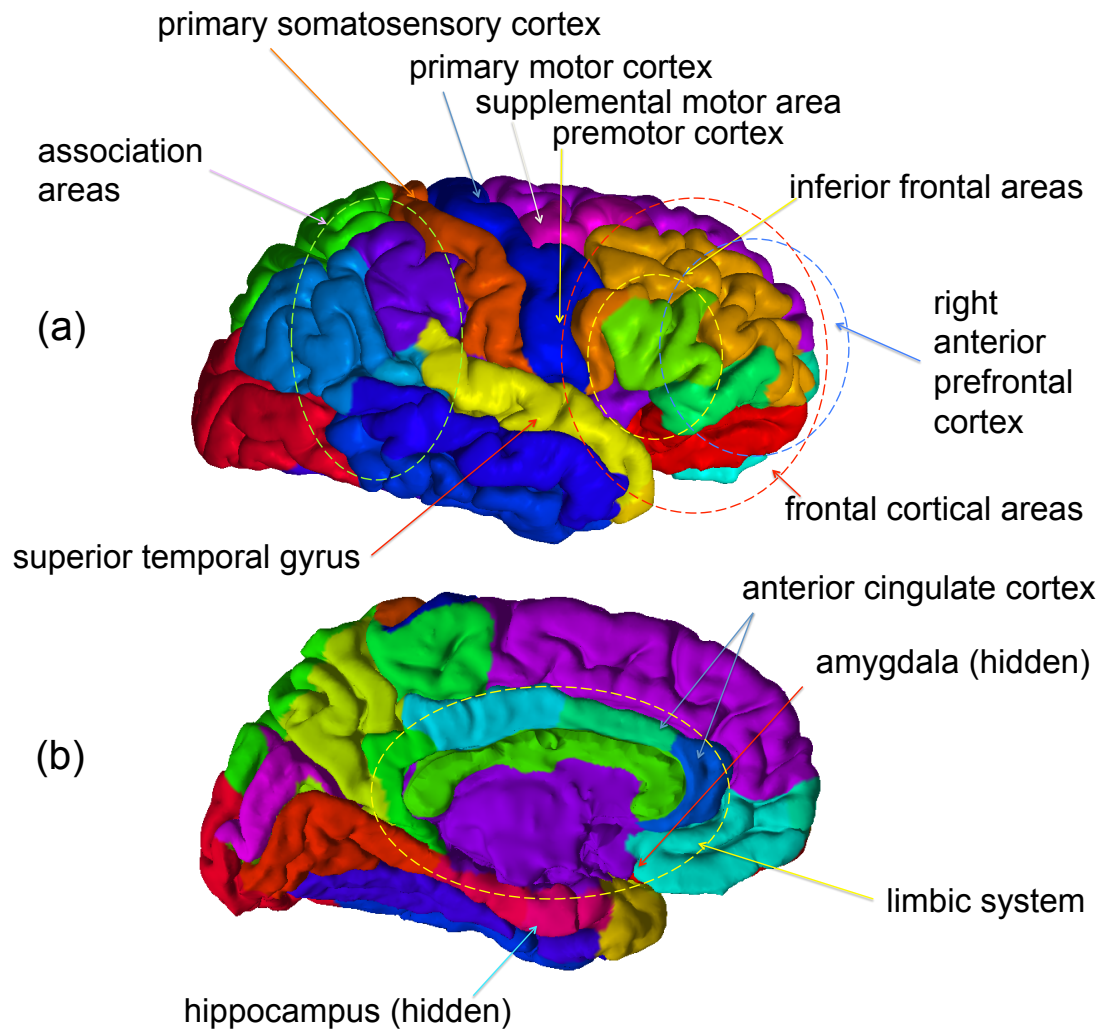


Figure 1.2: (a) Cortical brain structures involved in attention, auditory processing, and motor planning. (b) Subcortical brain structures involved in emotion response and motor planning.

antness. Higher level musical expectations can also be built based on short-memories, such as those accumulated over the course of a music piece, or long-term memory, where prior musical and cultural knowledge is recalled to make sense of a musical passage. Violations or fulfillment of these expectations can be formed, facilitated by higher-level processing in Broca's area, located in the left hemisphere and thus not visible in Figure 1.2, and the dorsal region of anterior cingulate cortex [Brown, 2000], shown in Figure 1.2b, supporting feelings of surprise and awe, or disappointment [Meyer, 1956][Huron, 2006a].

Emotions can also be induced by the memories we have that are linked to the music we are listening to. Episodic memory recall of associated feelings due to personal experience or cultural context will induce emotions. Episodic memory is in general supported by right anterior prefrontal cortex (circled in Figure 1.2a) and the hippocampus (surrounding structures shown in Figure 1.2b)[Brown, 2000]. Emotional conditioning, the repeated presentation of a piece of music along with a positive or negative stimulus, can cause a largely unconscious emotional response to the future presentation of that musical stimulus. Brain areas related to aversion, such as the amygdala (surrounding structures shown in Figure 1.2b) support the feelings of disgust we may feel when exposed to a piece of music that we don't want to listen to. The recall of these memories can also inspire new visual imagery that comes with its own associated feelings [Bonny and Savary, 1973].

Emotions are known to be transferred empathetically via facial expressions [Ekman and Oster, 1979] [Carr et al., 2003], gestures [Preston and De Waal, 2002], and non-verbal expressive gestures found in speech [Laukka et al., 2005] [Hatfield et al., 1993]. Darwin hypothesized that this invoking of emotion through empathy was a crucial component of social cohesion and mother-infant bonding [Darwin, 1872]. Recent studies of the hypothesized mirror-neuron system of the premotor cortex [Rizzolatti and Craighero, 2004] suggest that observation of an individual's actions will cause the same activation of motor areas of the brain that would be used to actually carry out that action. This has also been shown to be true for hearing sounds attributed to another's actions [Kohler et al., 2002]. It has been hypothesized that we internalize a performer's actions as we hear or see them, and it is the rehearsal of these actions in our own minds, via the

motor cortex and motor association areas, shown in Figure 1.2a, that invokes the feelings we experience as a consequence of our observations [Molnar-Szakacs and Overy, 2006]. This is the hypothesized mechanism by which body language inherent in live musical performance encodes feelings that are transferred between people [de Gelder, 2006]. An array of basic emotions, e.g., anger, fear, and happiness, are communicable through observation of gesture musical performance. Feelings of arousal and communion are also created when we internally entrain to a rhythm or pulse, with the help of internal oscillators [Clayton et al., 2005] that keep track of cyclic auditory events in the cerebellum, and the associational areas of the auditory and motor cortex (circled in Figure 1.2a).

Overt repetitive movements are an integral part of the listening practice in music traditions where highly engaged, trance-like mental states are desired [Becker, 2004]. However, the processing of musical events in motor cortex makes musical attention an inherently embodied process [Iyer, 2002], even when listeners do not move along with the music. The motor cortex and supplemental motor areas (Figure 1.2a) are active in musical novices [Zatorre et al., 2007a] and professional musicians as they listen to music [Haueisen and Knösche, 2001], suggesting motion-imitation underlies many aspects of music perception, beyond just emotion communication. Once an emotional response has been created, a listener is motivated to continue attending to the musical stream, closing the loop in our proposed model. This focused attention due to emotional response, modulated by expressive movement, characterizes the phenomenon of musical engagement that is studied in this dissertation.

The field of music perception and cognition, much like the field of linguistics, has chiefly concerned itself with the mapping of low-level psychoacoustical percepts into higher-level musical constructs. But comparatively little is known about what happens after we perceive pitch and rhythm: how brain and body interact to engage our attention and provoke emotion responses in uniquely musical ways.

1.2 Past approaches to measuring emotional response to music

Previous literature has attempted to describe the state of engaged music listening using concepts such as “tension” [Meyer, 1956] “activation” [Zentner et al., 2008] or “flow” [Csikszentmihalyi, 1990]. Numerous studies have been conducted in order to define methods by which these emotional responses to music can be measured, drawing upon psychological and neurological experiments [Eerola and Vuoskoski, 2013]. As our model of musical engagement focuses on the real-time engagement of the listener, we are concerned primarily with studies where participants provide continuous reports of their affective experience as they listen to music.

This self-report paradigm was first used by Nielsen in a study of musical tension [Nielsen, 1983], and was replicated by Masen and Fredrickson using a digital slider device [Madsen and Fredrickson, 1993]. It was extended to the use of a two-dimensional model of emotion by Schubert [Schubert, 1999] and Nagel [Nagel et al., 2007]. Krumhansl found convincing evidence that the self-reports obtained by this method are replicable and correlate with other measures of emotion. She used a computer interface to continually record listeners’ responses to emotional music stimuli as they listened [Krumhansl, 1996], and found that their measurements of tension as they listened to the first movement of a Mozart piano sonata correlated with the structure of the music as defined by Lerdahl’s Tree Model [Krumhansl, 2002]. In separate experiments, she found that their ratings of tension felt in various music excerpts correlated with their skin conductance level (SCL) and heart rate measurements [Krumhansl, 1997]. Timmers et al. [Timmers et al., 2006] presented video and/or audio recordings of three performances of a Scriabin piano etude to participants, and asked them to rate their emotional engagement with the performance in real time. Their emotional engagement was found to correlate with the measured dynamics of the performance, which in turn was connected to the movements of the pianists as measured by a video processing program.

The flaw in these measures is the fact that listeners must constantly monitor their affective state, an act that intuition tells us will impede musical engagement. Zentner

reported anecdotally that participants in his continuous report studies were not able to listen, monitor their emotions, and complete the physical task of reporting the emotion simultaneously [Zentner and Eerola, 2010]. In other words, the added monitoring and report tasks added a cognitive load that interfered with the participant's engagement in the music.

This could be explained empirically given evidence of the role of anterior cingulate cortex (ACC) in the regulation of affective and cognitive states [Mohanty et al., 2007]. In particular, it has been shown that motor-response selection tasks, such as the self-report paradigm presented above, activate the cognitive division and suppress the affective subdivision of the ACC in the presence of competing streams of information [Bush et al., 2000].

Evaluation can interfere with other types of emotional response. Taylor demonstrated that limbic system (amygdala and insula) response to aversive pictures is impacted when participants are asked to appraise them on a pleasant/unpleasant scale, when compared to passive viewing [Taylor et al., 2003]. Critchley found that limbic activation decreased when participants are asked to report on the expressions of faces, as opposed to their gender [Critchley et al., 2000]. Brattico found differing ERP responses to inappropriate chord sequences depending on whether a cognitive appraisal or subjective judgement task was given, and hypothesized that these different latencies represent differentiated systems for cognitive and affective music listening [Brattico et al., 2010].

Lastly, people are biased in their evaluations of behavior, tending to be much more willing to evaluate the traits of others over themselves. Konecni cites this well-known "attribution theory" [Jones and Nisbett, 1971] to explain that continuous self-report of felt emotion is likely to break down under cognitive load, since participants will likely report the intended feeling of the music, rather than report about their own internal feeling state [Konecni, 2008].

The evidence that cognitive load can interfere with emotional engagement warrants our exploration into non-cognitive means of measuring emotional engagement in musical listening tasks, which we present in the next section.

1.3 Proposed Study: Measuring Musical Engagement

Psychological experiments of music listening are typically performed in laboratory settings offering a listening experience that differs from the kind of experience that listeners may seek for themselves in every day life. This clinical presentation is one reason why experiments are criticized for lack of ecological validity. We developed an experiment protocol designed to invite a music listener to optimally experience and communicate musical feeling while in a laboratory setting. This experiment was designed to answer the following questions:

1. Can participants be trained to gesture expressively to music excerpts played for them, in a manner that is consistent with the feeling intention of music excerpts that are played for them? We hypothesize that we will be able to train participants to consistently perform expressive gestures that convey the feeling intention of the music they are listening to.

2. Can the feeling intention of participants' expressive movements be communicated through a vastly simplified, animated presentation of the performance? And furthermore, which motion dynamics are necessary to properly convey this feeling intention? We hypothesize that audience members viewing point-light animations of conducting participants' expressive gestures will be able to classify the feeling intention expressed by the conducting participants. We hypothesize that the responses given by audience members will not be affected by the manipulation of the tempo and overall trajectory length of the animations, since the feeling intention can be conveyed by the acceleration path within each movement cycle.

3. Can we characterize the brain dynamics that support engaged music listening? We hypothesize that the brain dynamics supporting engaged music listening will agree with findings in the fMRI literature.

4. Are the brain dynamics supporting music engagement modulated by expressive movement? We hypothesize that our co-analysis of the motion dynamics of expres-

sive movement and EEG of engaged music listening will uncover brain dynamics supporting engaged listening that co-vary with the expressive movements performed by the listeners.

The structure of the study is shown in Figure 1.3. It involves two main experiments; first, we recorded motion capture and EEG data from participants as they performed expressive gestures attempting to convey the feeling intention of the music they are listening to; second, we animated the recorded motion capture data and asked Internet viewers to rate these performances for engagingness and feeling intention.

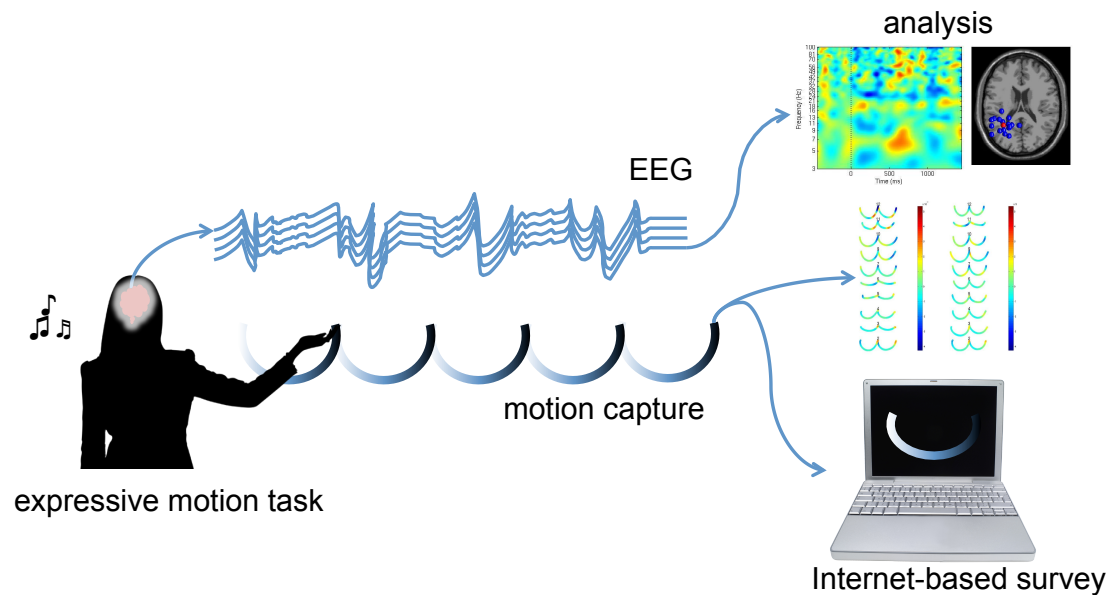


Figure 1.3: Experiment Design Overview.

First, we gave all of our participants a guided relaxation session which incorporated language from mindfulness meditation, Deep Listening, and the Guided Imagery in Music method. Mindfulness meditation has been shown to improve music engagement [Diaz, 2011] and increase efficacy of brain-computer interfaces used by clinical populations [Lakey et al., 2011]. Our relaxation exercise has similarities to The Deep Listening practice [Dempster et al., 1989] developed and disseminated by composer Pauline Oliveros. One goal of Deep Listening is the development of flexible listening attention so that the listener is able to switch between focused, foreground-based listening, and more global, environment-based listening. Our experiment protocol has

similarities to Deep Listening by incorporating repetitive motion exercises and breath training into preparation for a multidimensional listening experience. The script developed for our experiments (see Appendix A.1) borrows from a listening exercise used in Helen Bonny's Guided Imagery in Music method, developed for music therapy purposes [Bonny and Savary, 1973]. In this method, a listener seeking altered states of consciousness follows three steps: relaxation, attention, and induction, assisted by music excerpts specifically chosen to produce vivid mental images.

Second, we gave each participant an individualized training session in an expressive movement task, similar to the gesture a conductor uses to lead an orchestra, that was designed to allow participants to communicate their internal feeling state without adding any extra cognitive load. The expressive movement task had the added benefit of amplifying their experience of the music excerpts we played to them. This task was inspired by the work of Manfred Clynes, whose theory of sentics [Clynes, 1977] is built upon the principle that complex expressive gestures are composed of 'actons,' a small set of fundamental movement programs which are preprogrammed in the brain. Clynes developed a repetitive motion task in which participants were asked to express a set emotion, such as fear, anger, or love, using a single finger attached to a two-dimensional movement recording device. The results of his experiments demonstrated that a) remarkable consistencies between individuals and across cultures were found in how people expressed these emotions using this motion task, and b) the repetitive expression task was a reliable mechanism to invoke strong feeling states in the participants. We invited our participants to complete an expressive movement 'conducting' task as they listened to various music excerpts. Periodically, we introduced cognitive load by adding a non-rhythmic arithmetic task that the participants were expected to complete as they were moving expressively to the music. This distractor task was introduced so that we could later analyze the differences between the data collected in the fully engaged expressive movement trials and the distracted, not-engaged trials.

As we played music excerpts, we recorded the participants' expressive movements using a full-body motion capture suit, with an additional marker attached to the middle finger of the conducting hand. These movement traces were recorded with the intention of analyzing them for patterns that would correspond to the different music

pieces played to the participants, and the engaged or not-engaged conditions.

Our expressive intentions as artists are encoded in our gestures. The importance of these gestures is self-evident for dancers, but they are also vital to the visual arts and music. The Abstract Expressionists [Rosenberg, 1994] and Japanese brush and ink painters [Davey, 1999] painted with expressive intent that was encoded in the brush strokes we see on their canvases. Any music amateur who has attended a live performance by one of their favorite artists can attest to the power that the physical gestures of the musician performer have over the audience's musical experience. The added information encoded in these gestures has an impact on the reception of the musical performance that has been verified experimentally. Luck et al. [Luck et al., 2010] showed that ratings of valence, arousal, power, and expression given to point-light displays correlate with various computed features of the performances. Experiments by Wanderley [Wanderley et al., 2005] have shown that audiences use ancillary gestures made by instrumentalists as cues for interpretation of their performance. Vines et al. demonstrated by comparing audio, visual, and audiovisual presentations of instrumental performance, that the visual channel effects the interpretation of tension felt from the musical performance [Vines et al., 2006]. This can be extended to classical musical conducting, the practice that inspired the expressive movement task given to our participants. The Czech conductor Vilem Tausky wrote that a conductor's "greatest responsibility is to interpret the work by translating its meaning to the audience in such a way as to convey all its subtleties, even though the composer might not always give clear directions [Grindea and Menuhin, 1995]." In other words, conductor's movements are a vital tool to help the audience interpret the music they are listening to.

To some extent, the success of a painter, dancer, or musician hinges on their ability to communicate through these expressive gestures. Thus, an important part of the proposed study was a rating exercise given to Internet audience members to see if they were able to interpret the gestures made by the conducting participants. If the audience members were able to distinguish between the interpretations of the different musical excerpts, we can claim that the recorded motion traces of our expressive movement task are effective representations of musical feeling. Furthermore, if the audience members are able to distinguish between the motion traces recorded from the engaged trials and

the not-engaged trials, we can claim that the engagement of the music listeners was impacted by the cognitive distractor task we gave them.

This claim will be important when we analyze EEG data recorded from the conductor's scalps during the expressive engagement task. We will compare the EEG data from the engaged and not-engaged trials. When the EEG data from not-engaged trials is subtracted from the EEG data from the engaged trials, we will see in the data that remains, a representation of the brain dynamics that are specific to the musically engaged state in our conducting participants. An analysis of this data that is time-locked to the motion capture data can reveal information about the time course of these brain dynamics and their relationship to planning and execution of expressive movements.

Many EEG studies of music listening exist in the literature. However, for decades these studies have focused on Event-related potentials (ERPs), small voltage changes that are isolated from a recorded EEG signal. They represent a combination of the electrical fields generated by the group of neurons that respond synchronously to a particular event, whether it be sensory, motor, or cognitive. The resulting information has temporal resolution on the order of milliseconds, and is thus useful for studies of timing and sequence of brain activity. ERPs are calculated from an EEG using signal processing techniques, most commonly by averaging over windows of a signal that are time locked to a repeated stimulus. This procedure is necessary in order to obtain a proper signal-to-noise ratio since ERPs tend to be smaller than the amplitude of the EEG signal from which they are detected; however, it is limited in its ability to render information from individual trials.

These studies have mostly served to clarify the time course that various auditory processing modules follow in response to a musical stimulus [Koelsch and Siebel, 2005]. For instance, in [Fishman et al., 2004] it was demonstrated that within the first 100 ms (i.e., the N1 or P1 response), acoustic information about the stimulus (pitch, loudness, tone quality) is processed in the auditory cortex. A further stage of auditory grouping, or Gestalt processes likely occurs after these features have been extracted. The theory of auditory scene analysis developed by Bregman [Bregman, 1990], was developed using behavioral measures. An example of one of these grouping principles is grouping by pitch register. In [Bregman and Campbell, 1971], subjects were presented with a se-

quence of high and low tones, and asked the order in which they appeared. Given a fast presentation of tones, subjects were only able to reply first with the order of one set of tones (high or low), and then with the order of the other set. That is, they were unable to hear the sequence as a single stream of high and low notes, but rather they heard two separate streams separated in pitch. Given a slower presentation of tones, subjects were able to recall the order tones irregardless of pitch height. ERP studies have since been able to verify Bregman's assertion that these grouping mechanisms are performed at the preattentive stage of auditory processing. A mismatch negativity (MMN) is seen in the 100-200 ms range when regularities in an auditory stimulus, such as a repeating musical sequence, are violated [Tervaniemi and Huotilainen, 2003]. It is thought to reflect the processes by which the auditory cortex stores a repeated sound in memory. This MMN is thus related to grouping mechanisms in music listening in that it is evoked when a new auditory stream is created and attended to. However, this effect is preattentive: it has been shown to be equally strong in groups instructed to attend to a stimulus versus those instructed to read a book while listening [Sussman et al., 1999].

Later-stage ERPs have revealed information about higher-level processes involved in the processing of musical stimuli. Halpern [Halpern et al., 2008] showed that trained musicians have a P300 "oddball" response when hearing the lowered third scale degree indicative of minor modes; non-musicians are able to correctly identify the negative affect of the minor mode but do not exhibit the P3 response, probably because they do not process the needed information. ERPs have also been used to show that semantic processes similar to language are involved in the processing of musical stimuli. In particular, [Koelsch et al., 2004] demonstrated that the semantic processing N400 ERP is triggered by word stimuli when they do not correspond to a musical primer stimulus. The literature on ERP studies of music have illuminated some details about the sequential nature by which music is processed by the human brain; however, the underlying neural mechanisms associated with emotional responses to music remains insufficiently studied in the EEG literature.

The performed study used signal processing techniques for EEG developed at the Swartz Center for Computational Neuroscience at UCSD, namely the independent component analysis (ICA) of EEG data [Makeig et al., 1996] [Makeig et al., 2002].

The ICA algorithm extracts a set of processes (or components) which are maximally independent in the time domain; it does this by learning spatial filters that will maximize independence of independent components over the time domain. In the best case, the resulting map of independent components matches what we would expect to find if each were a single dipole, overlapping in space. However, some components are split into two distinct regions, possibly representing two patches of cortex that have the same sensory input or are connected directly through the corpus collosum [Makeig et al., 1996]. The ability to extract independent components of EEG activity, thereby being able to attribute frequency information to spatially separate sources, results in more detailed functional information than scalp recording alone.

The recent study of Onton and Makeig [Onton and Makeig, 2009] used ICA in classifying 15 different emotional states in participants primed using guided imagery techniques inspired by the work of Helen Bonny [Bonny and Savary, 1973]. A key finding in this study was that information about the imagined emotion was contained in the activities of many ICs, in particular by identifiable changes in source EEG spectral power in the beta, gamma, and high gamma bands. The latter two lie in the 30-90 Hz range, which overlaps strongly with electromyographic (EMG) activity from head and neck muscles. Without ICA, one would be forced to throw out information above the 30 Hz mark for fear it was due to artifacts from head movement. Using ICA one is able to make the distinction between muscle and cortical sources of gamma-band activity. In the [Onton and Makeig, 2009] study, this functionality was very important as gamma band activity had been linked previously to the processing of emotional stimuli; for instance, Buddhist monks have been shown to increase gamma-band power in the 30-60 Hz range when asked to meditate on love and compassion [Lutz et al., 2004]. We might therefore expect to find important activity in the gamma range when inducing strong, particularly positive, emotions in subjects.

A few studies using independent component analysis of EEG for music emotion have been performed. Lin et al. have demonstrated that EEG bandpower changes measured from the frontal midline area, suggestive of anterior cingulate cortex activity, is correlated with changes in tempo and mode of music stimuli presented to subjects [Lin et al., 2010a], and that valence of the musical stimuli is revealed by changes in

delta- and theta-band power measured at the frontal-central area [Lin et al., 2010b]. It is our hope that the experiments described in the following chapters, with the help of ICA analysis, will uncover new information about the brain dynamics that support engaged music listening and expressive movement.

Chapter 1, in part, is currently being prepared for submission for publication of the material. Leslie, Grace; Ojeda, Alejandro; Makeig, Scott. Measuring Musical Engagement Using Expressive Movement and EEG Brain Dynamics.

Chapter 2

An Expressive “Conducting” Experiment

In the experiment discussed in the present chapter, from now on referred to as the “Conducting Experiment,” novice and expert music listeners are invited to move expressively to music excerpts as we record their movements and EEG. We introduce an additional distractor task in some trials in an attempt to modulate their level of engagement in the emotional expression task. We analyze their recorded movements and EEG for patterns related to the feeling intent of the music, and their level of music engagement.

Our analysis attempts to answer the following questions:

- 1) How can music listening techniques be adopted in a laboratory setting to encourage engaged music listening?
- 2) What are the motion traces that are associated with the expression of feeling intent from various music excerpts?
- 3) How do these motion traces differ depending on the level of musical engagement of the performer?
- 4) What are the brain dynamics that are associated with the expression of feeling intent from various music excerpts?
- 5) How do these brain dynamics differ depending on the level of musical engagement of the performer?

2.1 Method

2.1.1 Demographics

Twenty-one right-handed participants were recruited from two different subject pools. The mean age over all 21 participants was 24.4 years with a standard deviation of 5.6 years. Eleven ‘novice’ participants, 5 male and 6 female, were recruited from a general subject pool consisting mainly of university undergraduate students. Their combined years of experience in music and 10 other expressive movement disciplines (see A.5: Music Experience Questionnaire) was a mean of 6.6 years of training summed across all 11 disciplines, with a standard deviation of 6.8 years. Ten ‘expert’ participants, eight male and two female, were recruited from a pool of graduate students in the Music Department at UCSD. These expert participants had a mean of 28.8 years of training summed across all 11 disciplines, with a standard deviation of 12.1 years. Written informed consent was obtained from all participants. All experimental procedures were carried out following the University of California Institutional Review Board.

2.1.2 Stimuli

Ten short film music excerpts of approximately 15 seconds in length were selected from a corpus of soundtracks developed by Tuomas Eerola and Jonna K. Vuoskoski at the University of Jyväskylä, Finland. The samples in this database were chosen as either moderately or highly representative of five discrete emotions (Anger, Fear, Happiness, Sadness, Tenderness) and three bipolar emotion dimensions (valence, energy, and tension). All samples were evaluated by 116 novice music listeners along these eight music emotion dimensions [Eerola and Vuoskoski, 2011]. The subset of ten excerpts chosen for the conducting experiment were chosen to span this music emotion space. The information for each of the excerpts is listed in Table 2.1. Figure 2.1 shows the organization of this subset of excerpts by plotting their ratings from [Eerola and Vuoskoski, 2011] along the valence and energy axes. Figure 2.1 demonstrates that the excerpts cluster by valence and energy, with one cluster containing the high-energy, low valence pieces, and the other cluster containing the low energy, high valence pieces.

We ran several experiment pilots during the development of the conducting ex-

Table 2.1: Label, Target Emotion, and Source information for excerpts used in the present study.

	Present Study Code	Eerola Study code	Emotion Target	Album Name	Track	Min:Sec
2	26		Happy	The English Patient	7	00:33 - 00:58
3	38		Sad	Psycho	3	00:58 - 01:24
4	41		Sad	Big Fish	22	00:00 - 00:20
5	65		Tender	Road to Perdition	1	00:35 - 00:53
6	74		Tender	The Untouchables	9	00:04 - 00:23
7	92		Fear	Psycho	1	00:00 - 00:22
9	125		Anger	The Alien Trilogy	9	00:03 - 00:16
10	134		Anger	Shakespeare in Love	15	00:40 - 00:56
11	151		Surprise	The Rainmaker	5	00:00 - 00:16
12	241		High Energy	The Godfather	11	00:49 - 01:06

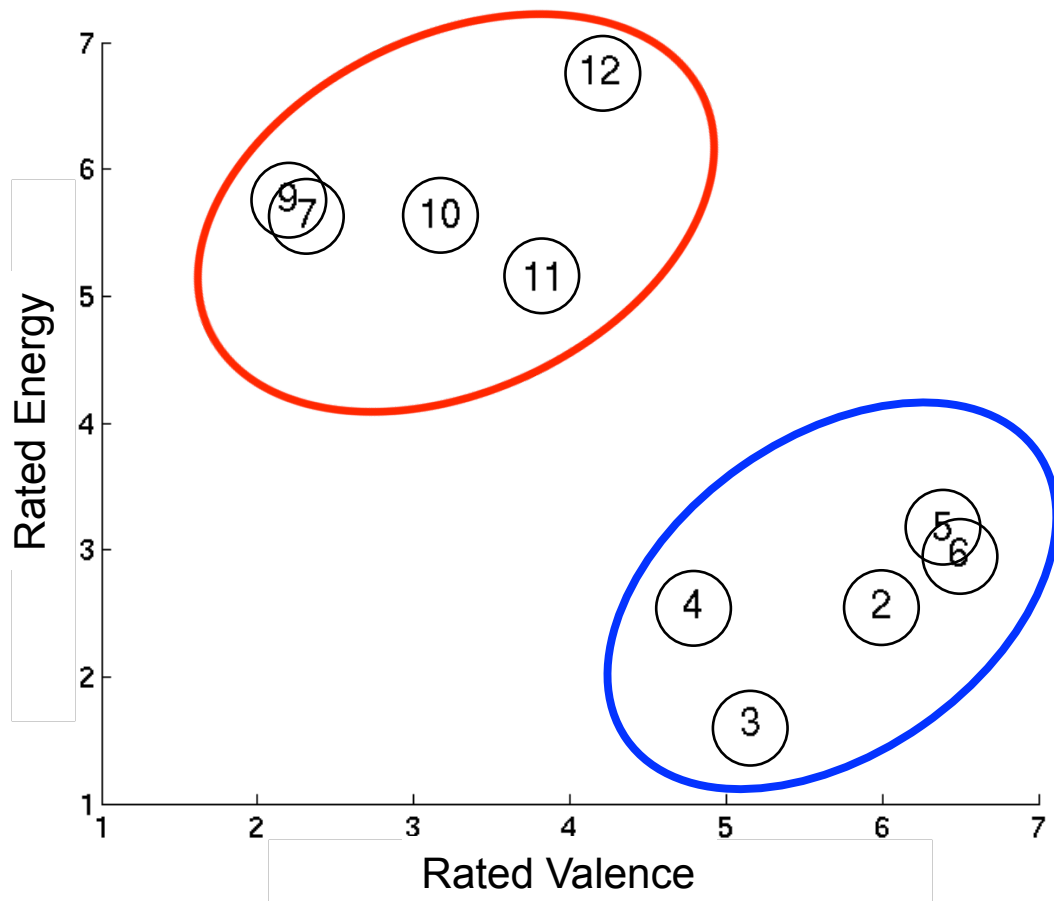


Figure 2.1: Music stimuli used in the conducting experiment plotted along the Valence and Energy dimensions. Two clusters of excerpts emerge, one containing the high energy pieces, and another containing the low energy pieces.

periment protocol. Originally, classical symphonic music excerpts were chosen as exemplars along the two bipolar emotion scales of valence and arousal, based on the author's music taste and experience. In subsequent pilots, excerpts from instrumental classical pieces included in Helen Bonny's suggested recordings for "altered states of conscious experience" were chosen, one from each of her defined "mood clusters:" solemn, sad, tender, leisurely, playful, gay, exciting, and vigorous [Bonny and Savary, 1973]. Longer (one minute or more) excerpts were chosen so that the music emotion intention of the composer was made clear during that excerpt. Even so, the feeling intent of these excerpts was not clear enough to the pilot participants for them to come to a consensus when rating. Also, these music samples were too long to be repeated multiple times as was required by the experiment protocol. Finally, the tempo of the music changed over the course of the excerpts, making it hard to train novices to beat along to the music.

The film music samples from [Eerola and Vuoskoski, 2011] were chosen because they solved these problems. First, film music composers are adept at eliciting emotional responses in their listeners using very few measures, and 16-second excerpts from the Eerola study sufficed. Second, film music tends to have sequences that are simple and highly regular in rhythm.

There are downsides to using film music for a study of musical emotion. Leonard Meyer famously distinguished between "absolutist" music, which derives its meaning exclusively from within the relationships set forth within the work, and "referentialist" music, whose meaning is created by reference to extramusical contexts such as love, suspense, or violence [Meyer, 1956]. In choosing film music excerpts as the stimuli for the final iteration of the conducting experiment, we've clearly chosen the referentialist path. These music sequences were composed to accompany visual images and narratives of love, war, and longing. When using these stimuli, it is impossible to attribute the feeling communicated by each musical excerpt to one particular factor. As mentioned in the introduction to this dissertation, musical feeling can be evoked by personal experience, cultural context, or preference. For our conducting experiment participants, these triggers could have been produced by the music excerpts themselves, but also in reference to the generic film narrative they suggest (i.e., a love story or war epic). This complicates the interpretation of the results, since any effects found by experimentation

cannot be attributed solely to the musical properties of the compositions used.

However, it is hard to imagine a music performance that doesn't contain some reference to extramusical meaning. Even a performance of the starkest, 12-tone music comes laden with cultural context, from the interpretation and gestures made by instrumentalists to the instruments on which they perform. It seems that one would have to generate music sequences electronically, bereft of any interpretable nuance, to avoid extramusical reference. Many studies of music emotion have been conducted using MIDI-generated stimuli. These studies have been criticized as lacking in ecological validity, since without a nuanced human interpretation, the music cannot adequately communicate the feeling intent of the composer. Our results from the present experiment and the audience experiment, described in Chapter 3, show that, even though the stimuli used refer to extramusical ideas, the conducting participants were able to communicate the feeling intention of these excerpts using only a moving white dot—a dramatically simplified communication channel.

2.1.3 Data Recording

Two research assistants helped each experiment participant fill out the necessary paperwork, and put on a PhaseSpace¹ full-body motion capture suit (see Figure 2.2 below). Once sitting in the SCCN Mobile Brain Imaging (MoBI) lab, the participant has a 128-channel EEG cap placed on his or her head. This cap was designed specifically for use in the MoBI lab and the electrode configuration, shown in Figure 2.3, did not conform to a standard system (e.g. the 10-20 system). One PhaseSpace LED was placed on the second joint of the middle finger of the participant's conducting hand, and the data stream containing the x, y, and z positions of the LED was streamed to a Producer program which animated the LED position as a point light display on a flat-screen television screen on the wall facing the participant.

The active EEG electrodes and amplifier used were part of a Biosemi² system. Data was streamed via optical connection from the amplifier to a computer running the

¹<http://www.phasespace.de>

²<http://www.biosemi.com>



Figure 2.2: A participant wears the full-body motion capture suit with an additional sensor worn on the conducting hand, and a 128-channel EEG system with cables connected to an amplifier worn in a backpack. The participant's movements are animated as a white dot on the facing screen.

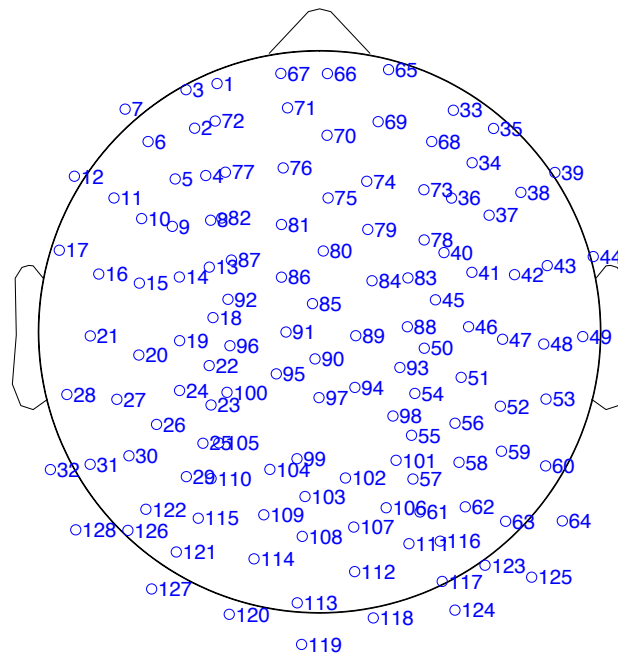


Figure 2.3: Channel locations for one subject wearing the 128-channel MoBI cap.

DataSuite³ environment. The locations of all electrodes were measured using a Zebris⁴ system. Conductive gel, used to establish a proper connection between the scalp and the active electrode, was inserted into each electrode in the cap using a syringe with a blunt-tip needle. The experimenter read the experiment description (see Section A.1) and played for the participant a short excerpt from the film *The Heart is a Lonely Hunter* [Miller, 1968], in which one character uses expressive arm and hand gestures to illustrate the feeling quality of an orchestral music recording for a deaf character so that he can hear the music through her gestures (see Figure 2.4).



Figure 2.4: Still from a film shown to the participant illustrating how a music listener can express their musical feeling through gesture. Copyright Warner Brothers, 1968.

2.1.4 Experiment Design and Presentation of Stimuli

The participant was invited to imagine a scenario similar to the one illustrated in the film excerpt played to them: they have a deaf friend in the adjacent room who longs to share in the experience of the music, but they cannot hear it. They can, however, see their movements via point light display, as the television screen acts as a window into the experimentation room.

³The DataSuite collection of experiment design and data recording software was developed by Andrey Vankov at SCCN.

⁴<http://www.zebris.de>

The experimenter demonstrated, with an example music excerpt looping in the background, the proper conducting hand gesture, a U-shaped pattern, similar to how a conductor would conduct in 2/4 time. The experimenter asked the participant to practice this gesture along with a music excerpt, and the experimenter intervened with further training if the participant was not reproducing the correct gesture. All further instructions were communicated to the participant via live audio feed from the adjacent control room.

Next, a guided body scan relaxation adapted for use in EEG experiments (see Appendix A.2) was read aloud to the participant, and the participant was asked to prepare to begin moving expressively to the music excerpts that followed. The participant's baseline brain (EEG) activity was measured while the experiment instructions and guided relaxation were presented.

The excerpts were presented to the participant in three blocks (see Figure 2.5 below). They were presented simultaneously with a specially designed "expressive" metronome beat consisting of enveloped bandpassed noise mimicking the whoosh of a conductor's baton, with the goal of minimizing the cognitive load in finding the musical beat. Each excerpt presentation was preceded and followed by four expressive metronome beats, as illustrated in Figure 2.6 below. The participant repeated the conducting hand gesture to express the feeling contained in the music sample.

They were allowed to repeat the music sample and their conducting as many times as they wished, until they were satisfied with their performance. Next, after the participant had completed a satisfactory run-through of the performance, they were invited to repeat this performance twice. Finally, they were invited to repeat the performance of the excerpt two more times, but while simultaneously performing a distractor task: they were read out loud a simple word (such as 'brain,' 'music,' or 'mind') and they were asked to convert each letter of the word to a number based on its position in the alphabet, and then add up those numbers. During all trials, the participants could not finish this task during the approximately one minute long stimulus presentation, but they were instructed to simply complete as much of the task as possible. The training, single task performance, and dual task performance sub-blocks were repeated for each of the 10 music excerpts.

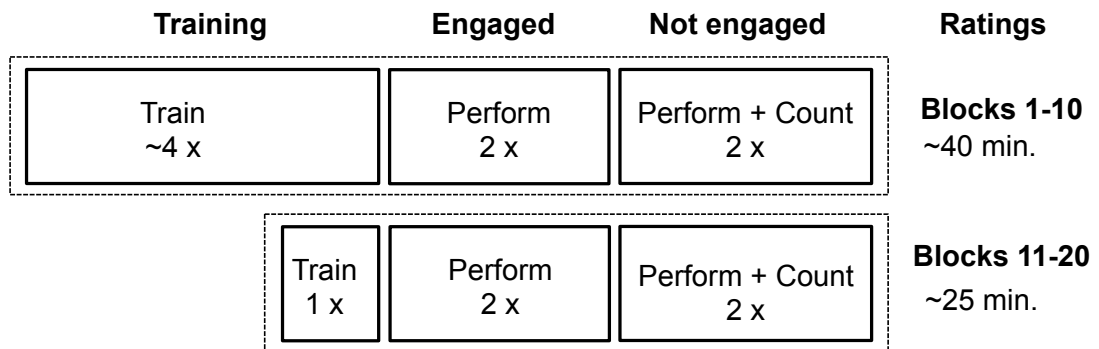


Figure 2.5: Each experiment block was divided into training, engaged, and not-engaged conditions. Participants completed a music emotion rating questionnaire at the end of each of the first set of blocks, before a new excerpt was presented.

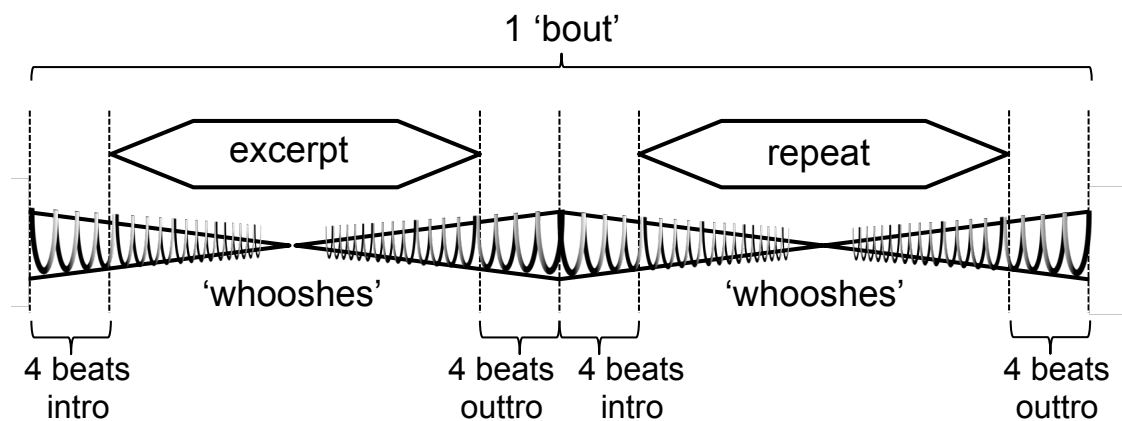


Figure 2.6: During each excerpt presentation 'bout,' the excerpt was preceded by four repetitions of a bandpassed noise 'whoosh' sound designed to show the participant the beat of the excerpt.

After the performance of the film music sample had been repeated 10 times, the participant filled out a short questionnaire that asked him to evaluate his emotional response to the music piece using the Geneva Emotional Music Scale (GEMS-9) [Zentner et al., 2008], in addition to rating the sample for valence and arousal using the Self-Assessment Manikin (SAM) [Bradley and Lang, 1994] (See Appendix A.4).

After a five to 10 minute break, a second experiment block was completed, this time with only one repetition of the excerpt for training. This experiment design yielded a total of 32 minutes of single-task, ‘engaged’ listening data, and 32 minutes of dual-task, ‘not engaged’ listening data.

2.1.5 Data Streams Recorded

EEG data were collected synchronously from 128 scalp, two infra-ocular, and two mastoid electrodes with an active reference at a sampling rate of 512 Hz with 24-bit A/D resolution. The stimulus onset and offset of each music sample and metronome track were recorded in simultaneously acquired event channels. In addition, the participant’s behavior was recorded both with a video camcorder and a 31 motion capture channels. Separate computers were used for audio stimulus presentation, animation display and video recording, and recording of the PhaseSpace and BioSemi data streams. Time codes were sent from the Macintosh running an automated Max/MSP⁵ experiment script, indicating the time at which music excerpts were played. The time codes were sent from Max/MSP to the DataSource program running on the “control” computer. This same method was used to send time codes from the VideoStream program recording digital video from a camcorder facing the participant in the experiment room from the “presenting” computer to the “control” computer. The DataSource program running on the “control” computer used the DataRiver protocol to store these time codes along with the EEG, motion capture, and video data as sample-synchronous streams.

⁵<http://www.cycling74.com>

2.1.6 Additional Participant Data Collected

After the experiment, the participant filled out a questionnaire packet consisting of the following surveys:

The Big Five Inventory [John et al., 1991] [John et al., 2008], a personality test.

The Tellegen Absorption Scale [Tellegen and Atkinson, 1974], a measure of the participant's ability to become absorbed in tasks.

Positive and Negative Affect Schedule (PANAS-Brief) [Watson et al., 1988], a measure of the participant's current mood.

Immersive Tendencies Questionnaire [Witmer and Singer, 1998], a measure of the participant's propensity to become immersed in their activities.

Short Test Of Music Preferences (STOMP) [Rentfrow and Gosling, 2003], a measure of the participant's preference for various genres of music.

A brief music experience questionnaire (See Appendix A.5).

2.1.7 Motion Capture Processing Pipeline

A common problem with optical motion capture data of body movements is that time markers often go missing from multiple cameras, making it very hard for the hardware to produce a reliable estimation of those values in real-time. Missing markers were estimated offline using piecewise cubic polynomial interpolation [Fritsch and Carlson, 1980] [Kahaner et al., 1989]. After interpolation, a 6 Hz lowpass, zero-lag filter was applied to smooth out non-movement related artifacts. Further analysis was restricted to the marker attached to the right hand used for conducting.

Because the participant was conducting at a fixed location always facing the same direction, this study was narrowed to the components of the motion along the vertical and horizontal plane orthogonal to the conductor's arm; we called this plane the action plane. Projecting the 3D trajectory to the action plane was done by PCA rotation [Jolliffe, 2005] and then discarding the dimension with the lowest variance. This component happened to be the one that accounted for the variability in depth. The rotation matrix was estimated using only those segments of data where the participant was conducting. To characterize the motion under different experimental conditions, the first,

second, and third time derivatives of the 2D trajectory were calculated, corresponding to the velocity, acceleration, and jerk profiles.

After the preprocessing described above, the swing cycles were identified using the velocity profile. The edges of the swings were identified as the time points when the magnitude of the velocity vector reached a local minima and changed direction. Then, non-overlapping trials of right-left-right cycles were identified. A trial was defined as a multivariate vector of x , y , velocity, acceleration, and jerk values at each time point in the cycle. Swing trials varied in number of cycles and scale, and in order to compare them statistically, they were scaled linearly to match right-left-right edges of the mean swing. After scaling, they were resampled along the time dimension to make them the same length. This way the time axis was converted from seconds to percent of the swing cycle. To identify and reject outlier trials, a pairwise Mahalanobis distance was computed using within subject cycles. Trials whose distance was larger than the 99% of the overall distribution were discarded.

2.1.8 EEG Processing Pipeline

The EEG data was passed through two stages of processing in order to be analyzed. In the first, seen in the left column of Figure 2.7, each individual participant's EEG data is passed through a rigorous set of de-noising steps, implemented in the EEGLAB toolbox for Matlab [Delorme and Makeig, 2004], to remove muscle and other artifacts in preparation for the ICA algorithm. In the second, the processed datasets are grouped together by condition, and clusters of independent components are computed across subjects. Statistics are applied to this across-subject time-frequency information to determine if any significant effects exist between conditions.

Individual level processing and analysis

First, the EEG data was first high-pass filtered at 1 Hz to remove gross drift patterns. A concatenated version of the EEG data was created using the timecodes from the audio presentation program. We discarded data from resting and training periods, retaining a concatenation of EEG recorded only while the music stimulus was being presented during the single-task 'engaged' and dual-task 'not-engaged' conditions. A

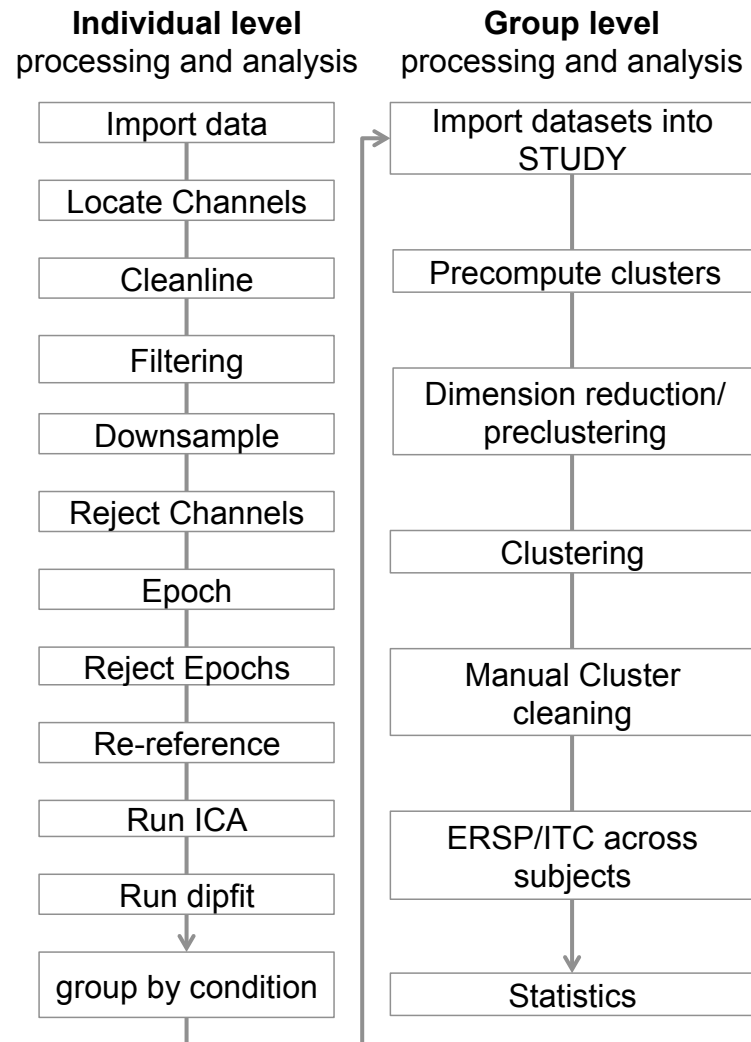


Figure 2.7: EEG processing pipeline, for individual (left) and group (right) level analysis.

second, continuous version of the EEG data was retained for the group-level processing stage, in which ERSF plots are calculated by averaging time-frequency information from epochs of varying lengths. The time points corresponding to the beginning of each Right-Left-Right swing cycle for the single-task and dual-task bouts were extracted from the motion capture data and embedded as events in both continuous and concatenated datasets. The time points corresponding to the training bouts were discarded.

The following steps were applied to both the concatenated and the continuous EEG datasets for each subject. First, we imported the EEG data into the EEGLAB Matlab toolbox. The first 136 channels were retained, discarding the other 8 channels containing motion capture data. The Zebris channel location file was read into the standard BEM 10-5 head model. The channel distribution was optimized around the center electrode Cz. Next, the non-EEG channels were discarded, leaving 128 channels to analyze. The cleanline program by Tim Mullen⁶ was used to remove electrical line noise at 60Hz and its harmonics. The data was downsampled to 256 Hz after application of a zero-phase FIR antialiasing filter. Noisy data channels were removed by rejecting channels with kurtosis greater than 5 ('bad' channels have been shown to have distributions of potential values that differ significantly from a Gaussian distribution [Delorme et al., 2001].)

Epochs from -.25 seconds to 1 second, with the 0 sec point time-locked to the beginning of each Right-Left-Right swing cycle, were extracted from the concatenated dataset only. All epochs whose power surpassed the -200 to 200 mV range were discarded. Epochs containing improbable data, determined to be channel data that was greater than 5 standard deviations from the mean within the channel, or all-channel data that was greater than 3 standard deviations from the mean, were rejected as well. This threshold resulted in discarding approximately 10% of the epochs from each dataset, as suggested in [Delorme et al., 2001].

After basic pre-processing, broken channels as well as short-time high-amplitude artifacts were identified and removed from the continuous EEG data. A cross-correlation matrix was computed within a two second sliding window across the entire session, and channels whose correlation coefficient with the rest was less than 0.45 for more than half

⁶<http://www.nitrc.org/projects/cleanline/>

of the session were labeled as broken and removed. Then, a two second sliding window PCA was used to statistically interpolate any high variance components exceeding a threshold relative to three times the covariance of clean segments of the data. Each affected time point of EEG was linearly reconstructed from the retained signal subspace based in the covariance structure observed in the clean segments.

One final preprocessing step, re-referencing, was applied to both the concatenated and the continuous datasets before running ICA. Typically a single channel EEG is calculated as the voltage difference measured between a point on the scalp and one or more reference electrodes. In the case of active electrode systems such as the Biosemi Active 2 system used in this experiment, the EEG is recorded without the reference and the reference calculation must be made during posthoc processing. During the conducting experiment, reference electrodes were placed over the mastoids. This resulted in a very noisy reference signal due to muscle artifact. The reference channels were discarded, and a reference signal calculated from the average of all EEG channels was used instead for re-referencing.

We ran both the concatenated dataset and the continuous data set through the CUDAICA algorithm, which was implemented on a GPU, and calculates the ICA weighting matrix and the data activation matrix using INFOMAX criteria [Raimondo et al., 2012].

Once the scalp map has been computed, it is possible to find the equivalent dipoles whose summed projections approximate the computed scalp map. Each dipole represents a patch of cortex whose neurons fire synchronously according to the time-frequency information contained in the row of the data activation matrix corresponding to that independent component (IC) [Makeig et al., 2011].

We calculated the equivalent dipole locations for all ICs contained in the concatenated and continuous datasets using the DIPFIT 2.2 plugin for EEGLAB by Robert Oostenveld [Delorme et al., 2011]. We used the head model from the Montreal Neurological Institute (MNI) standard coordinate system.

Finally, the concatenated and epoched dataset was divided into two datasets, one with all of the epochs centered around the beginning of the ‘engaged’ Right-Left-Right swings, and another with the ‘not-engaged’ epochs. Dividing the datasets by

condition allowed for easier automatic tagging of epochs by condition when loaded into the STUDY environment for group level processing, as discussed in the next section.

Group level processing and analysis

The epoched, concatenated datasets were loaded into the STUDY environment within EEGLAB for group level processing. The STUDY environment offers a group of functions that allow for the calculation of statistics evaluating the significance of differences between conditions across several subjects' data. The number of trials for each cell of the overall experiment design (Subject X Condition X Excerpt) was calculated. Any cells with zero trials were eliminated, either by excluding that subject's or that song's data from analysis, in order to correctly calculate statistics within EEGLAB. Three subjects and three songs were chosen for exclusion in this manner, leaving a total of seven songs and 13 datasets to analyze.

A group level processing pipeline was then followed, as illustrated in the right column of Figure 2.7. First, the dipoles within each dataset with residual variance greater than 15% were marked as excluded from analysis. The median number of remaining dipoles for each dataset was 21, the smallest number of remaining dipoles in an individual dataset was 12, and the largest was 42. Next, all measures used to cluster the ICs were computed, including the log-mean power spectrum and Event-Related Spectral Perturbation (ERSP). The ERSP is a valuable measure for estimating the event-related brain dynamics at multiple temporal scales [Makeig, 1993]. The ERSP measure allows us to test if the frequency components of the EEG signal coming from specific locations in the brain are correlated with the different parts of the swing cycle, and if these frequency patterns differ between the engaged and not-engaged conditions. The ERSP was calculated as the deviance from the mean time/frequency decomposition of the time course of each EEG independent components. The time/frequency decomposition was carried on by the continuous time wavelet transform [Mallat, 1999] [Kiebel et al., 2005] of the signal.

The Continuous Wavelet Transform (CWT) compares the signal to shifted and compressed versions of a mother wavelet. Stretching or compressing a function is collectively referred to as dilation or scaling, and corresponds to the physical notion of

scale. By comparing the signal to the wavelet at various scales and positions, a two-dimensional representation of a one-dimensional signal is obtained. For a scale parameter a and position b , the CWT is defined as follows:

$$C(a, b, f(t), \Psi(t)) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \Psi^* \left(\frac{t-b}{a} \right) dt \quad (2.1)$$

where $*$ denotes the complex conjugate. The scale parameter a is related inversely to the frequency, although no precise mapping exists; in wavelet analysis it is common to map scales to pseudo-frequencies [Mallat, 1999]. The more stretched the wavelet, the longer the portion of the signal with which it is being compared, and therefore the coarser the signal features measured by the wavelet coefficients. Mathematically the scale is defined as follows:

$$a = \frac{F_c}{F_a \cdot \Delta} \quad (2.2)$$

where Δ is the sampling period, F_c is the center frequency of the wavelet in Hz, and F_a is the pseudo-frequency correspondent to the scale a , in Hz. The shifting parameter b delays or advances the wavelet in time.

In our analysis we used a complex Morlet with center frequency 1.5 and bandwidth parameter 1 as a mother wavelet [Teolis, 1998]. The scale parameter used in our analysis varied logarithmically between 7.68 and 768 to cover a frequency grid from 1 Hz to 100 Hz. The logarithmic scale was used to maximize the frequency resolution at lower frequencies. This schema suits well the spectral characteristics of EEG signals where the power decreases toward the high frequencies. After computing ERSP matrices for each condition, they were co-registered in time following the same procedure used for co-registering mocap trials, as discussed in section 2.1.7.

After the ERSP computation step, we now had several independent components from each participants' dataset, with corresponding features calculated for each of these ICs. In order to find patterns in the behavior of these ICs across many subjects, we searched for clusters of independent components that existed across several subjects' data. We then grouped the information calculated from the ICs for each cluster, and tested them for significant differences between the engaged and not-engaged listening conditions in the experiment.

As a first step in the clustering method, we used principle components analysis to measure the abstract distance between each IC based on scalp maps, dipole models, spectra, and ERSPs. A multidimensional cluster position vector was computed for each IC based on these measures, and thus a global distance matrix was created which characterized how far apart each IC was from one another in this PCA-reduced abstract space. This dimension reduction was necessary because the k-means PCA clustering step which follows can take a limited number of independent variables as input. For our calculations, we gave relative weightings to each of the measures as follows: Spectra, 3, Dipoles, 10, ERSP, 2, ITC, 1. Thus, the computed dipole models were given the highest priority among the four measures included.

Next, the kmeans PCA-clustering algorithm [MacQueen et al., 1967] was used to find 21 clusters of ICs in the already PCA-reduced abstract space; this number was chosen so that on average, one IC per subject would be included in each cluster. Any ICs greater than 3 standard deviations from any of the cluster centroids were moved to a designated ‘outlier’ cluster. Next, each cluster was cleaned manually, by plotting the scalp map and spectrum for each IC included in the cluster. ICs with scalp maps indicative of muscle artifact or bad channels were moved to a designated outlier cluster. ICs with peak activity in the 20-30 Hz range were excluded, since these were likely muscle artifacts. We plotted the scalp maps for each of these clusters, and results comparing the spectral perturbation of a single cluster as seen in the engaged and not-engaged conditions, in section 2.2.3 below.

2.2 Results

2.2.1 Motion Capture Data

Variation between subjects

Within each subject, the movement trajectories varied significantly between songs, as shown in Figure 2.8, an example plot of one subject’s data. To create this plot, we averaged the Right-Left-Right movement trajectories within each subject’s data, and within each song. The resulting averages were divided by condition, single-task ‘en-

gaged’ and dual task ‘not-engaged.’ The x-y position of each trace corresponds to the position trajectory in the two dimensions of the dimension-reduced PCA analysis of the movement, whose calculation was described in section 2.1.7. Each trace is colored by the acceleration along the position trajectory path, so that moments of highest acceleration are colored in red, and moments of lowest acceleration are colored in blue. The movement trajectories varied greatly by subject, as demonstrated by the plots of the movement trajectories for all subjects included in Appendix B.

Difference between excerpts

We were interested in comparing the motion trajectories of the participants’ expressive movements to the different excerpts. In order to examine the differences between songs, we first normalized each Right-Left-Right trajectory in time, so that the beginning and end of each Right-Left-Right cycle corresponded. We then averaged across subjects, within excerpt. For this comparison, we only analyzed the swings from the engaged condition trials. These means are plotted and labeled by excerpt label in Figure 2.9. The axes of the plot correspond to the x and y position of the PCA-reduced movement trajectories. The second half of each trajectory has been reversed along the horizontal axis, creating a mirror image, so that it can be distinguished from the first half of the trajectory. Each trace is colored by the acceleration along the trajectory, so that red areas have maximum acceleration, and blue areas have minimum acceleration. We subtracted the grand mean across all subjects and excerpts, so that only the difference of each excerpt mean from the grand mean, the values of interest, would be plotted. The grand mean is plotted as the second-to-lowest curve in Figure 2.9.

To test for statistically significant differences, we computed a functional-ANOVA [Ramsay and Silverman, 1997] comparing the engaged subject means between each excerpt. To do this, we passed the swings through a normalization step in which they were approximated by the linear combination of b-spline basis functions that best reconstructs the original time series. The b-spline functions used to create the functional data objects from each dimension of the motion capture data consisted of fifth-degree polynomials that are joined together at spaced intervals along the time series. The defined set of basis functions, along with their coefficients for creating the linear combination, comprise

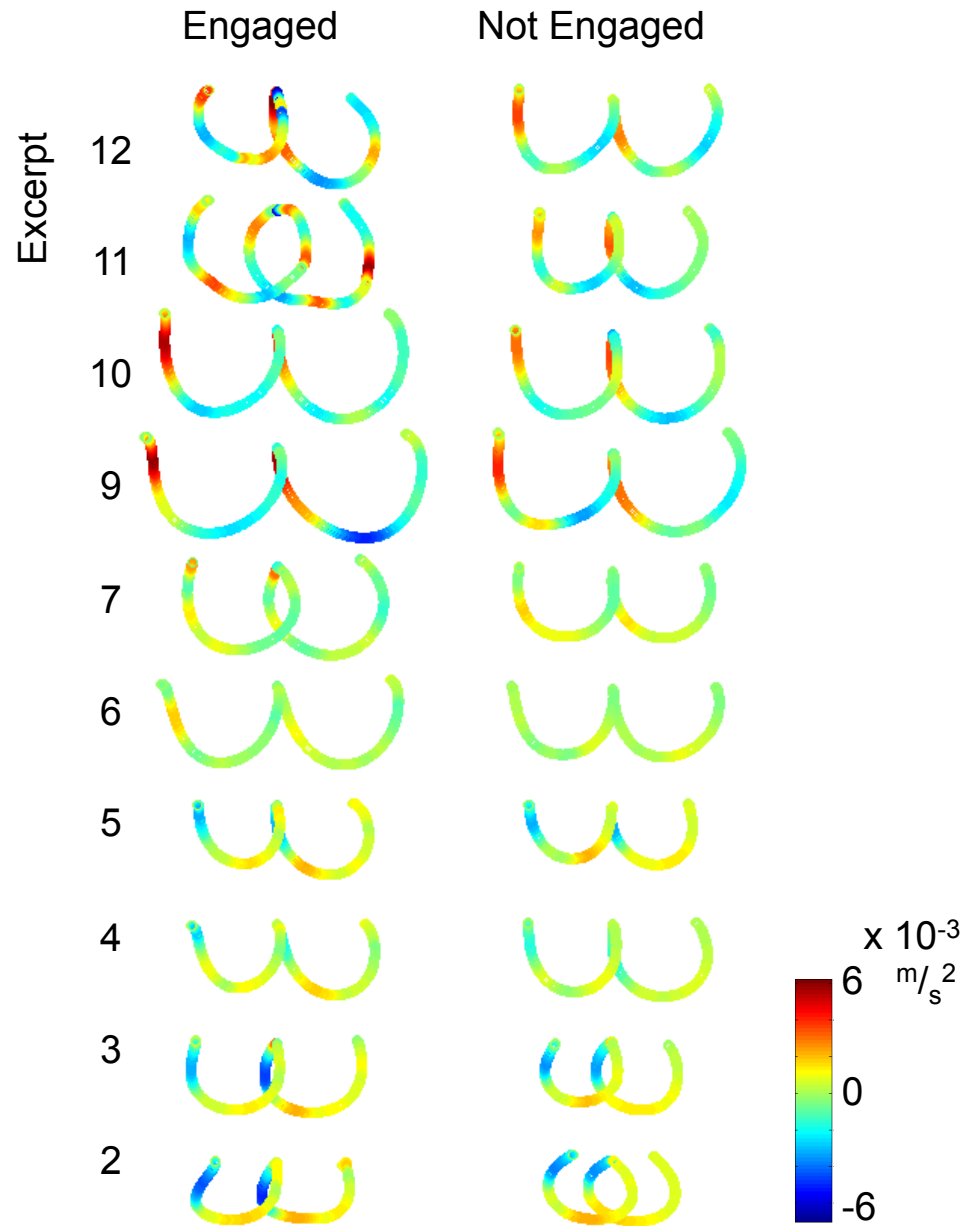


Figure 2.8: Plotted x-y trajectories of the average swings for subject 414 colored by acceleration trajectory, with red representing maximum acceleration, and blue representing minimum acceleration. The average trajectories over the engaged trials are shown in the left column, and not-engaged shown in the right column.

a functional data object that approximates the time series. This functional data object approximation is necessary for our statistical test, since points along a motion trajectory are not linearly independent observations, and thus they do not fulfill the independence requirement of traditional parametric statistical tests [Levitin et al., 2007]. We used a functional-ANOVA function to test the difference between the approximated motion capture paths for the engaged condition only, averaged across each excerpt, for each time point along the path. The time points at which the difference between songs was statistically significant at the $p < .05$ level are plotted in red on the bottom-most trace in Figure 2.9.

This plot demonstrates that the expressive movements of the participants varied significantly between excerpts, and that the acceleration paths for each excerpt contain clues as to the feeling expressed by that excerpt. The aggressive nature of excerpt seven is reflected in the severe and prolonged downward acceleration starting the swing. In contrast, excerpt two has a lilting quality that is reflected by a much smaller acceleration in the downward and upward portions of the swing, when compared to the grand mean across all excerpts. The significance curve at the bottom of Figure 2.9 shows us that the acceleration at three points along each swing represent much of the expressive difference between the excerpt performances.

We also performed a multidimensional scaling of the computed Mahalanobis distance of the acceleration and position trajectories for engaged condition only; these results are shown in Figure 2.10. These MDS results, when compared with the results in Figure 2.1, show that the recorded expressive movements corresponding to individual music excerpts cluster similarly to the emotion ratings given to those same excerpts. The balls representing individual music excerpts in Figure 2.10 are colored by rated energy to emphasize the influence of perceived energy on the clustering results. This clustering is a demonstration that the movements of the conducting experiment participants varied according to the feeling intent of the music excerpt. This supports our claim that the participants completed the emotional expression task we set out for them.

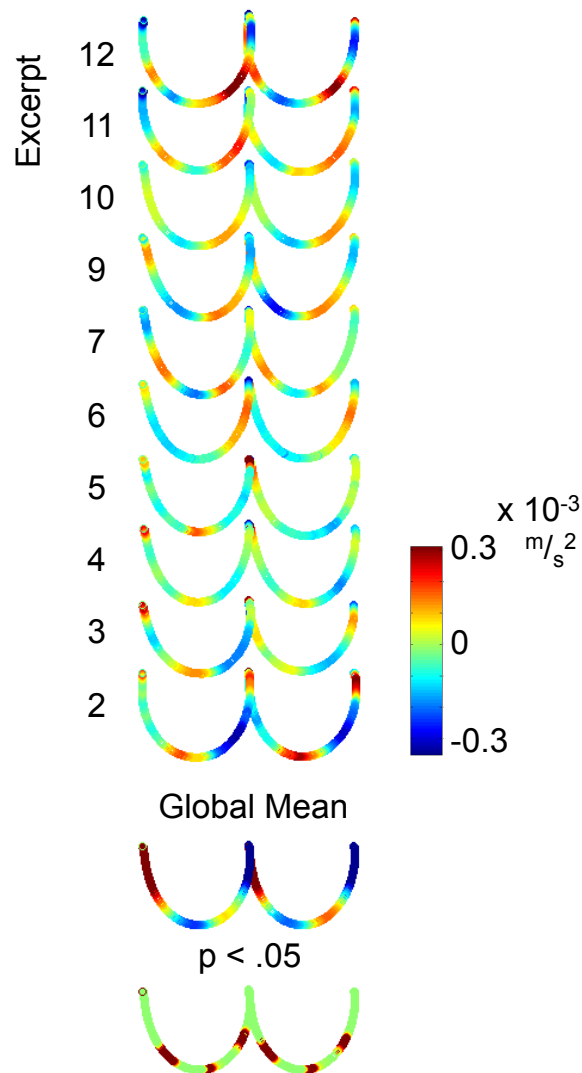


Figure 2.9: (a) Plot of time- and space-normalized trajectories for the engaged condition only, averaged across all subjects, with the global mean removed. The x- and y- axes of the plot correspond to the horizontal and vertical movement of the participants. The x-positions of the second half of the trajectory have been reversed to be able distinguish it from the first half of the trajectory. Each trace is colored by the acceleration along the trajectory path. (b) Global mean subtracted from each within-excerpt mean in order to produce the traces above. (c) time points at which difference between songs is significant at $p < .05$.

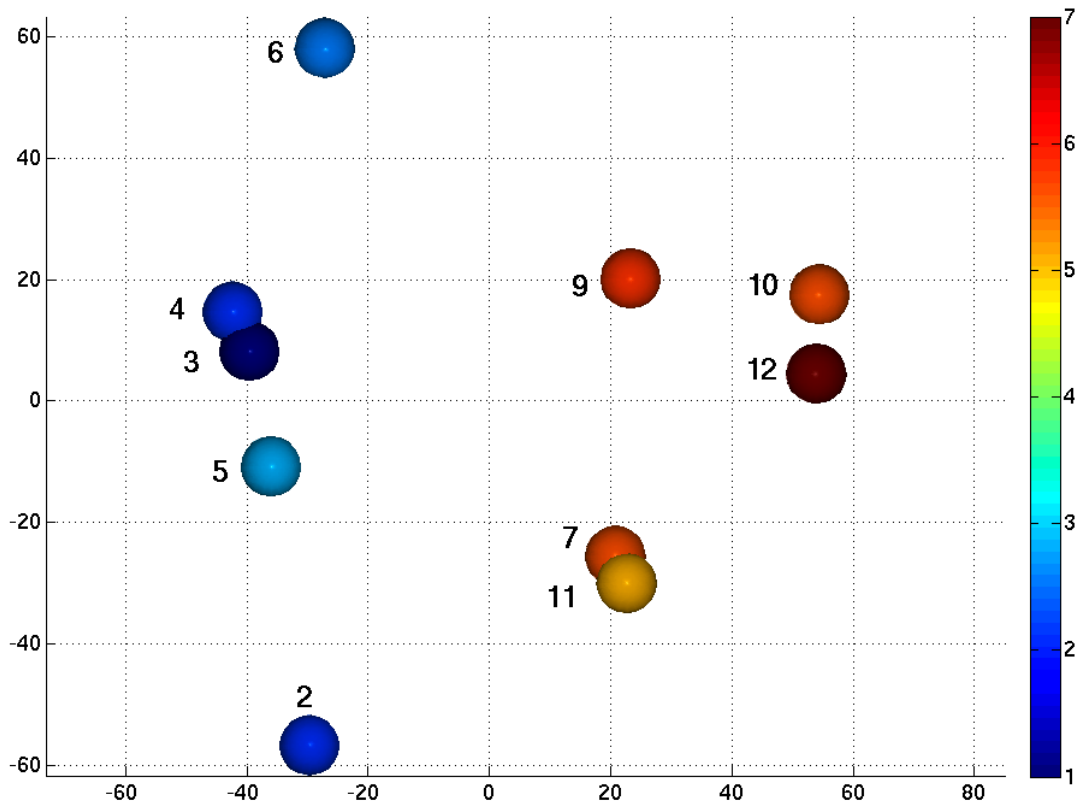


Figure 2.10: MDS plot of swing data shows that swings cluster similarly to emotion measures of excerpts. Each ball represents the average across all subjects, and is colored by the energy rating for the excerpt as reported in [Eerola and Vuoskoski, 2011].

Difference between engaged and not-engaged conditions

In order to examine the differences between the engaged and not-engaged conditions, we first normalized each Right-Left-Right trajectory in time. We then averaged across subjects, within excerpt. These means are plotted and labeled by excerpt label in Figure 2.11. Each trace is colored by acceleration. We subtracted the grand mean across all subjects and excerpts, so that only the difference of each excerpt mean from the grand mean, the values of interest, would be plotted. The grand mean is plotted as the second-to-lowest curve in Figure 2.11, and the mean of all engaged and all not engaged trials with the grand mean subtracted is plotted in the bottom-most trace in the left and right columns, respectively. The third column in Figure 2.11 shows the mean trajectory for the engaged and not-engaged conditions combined, colored by difference between the acceleration trajectory between the engaged and not-engaged conditions.

We computed a t-test between conditions within each excerpt using the functional data analysis method as described in Section 2.2.1. For this test, we used a functional-ANOVA function to test the difference between the engaged and not-engaged conditions by comparing the approximated motion capture paths within each condition, averaged within each excerpt, for each time point along the path. We masked the acceleration difference information in column three so that all time points at which the difference between songs are not statistically significant at the $p < .05$ level are plotted in green.

2.2.2 Discussion

The recorded motion capture data from the conducting experiment showed that individual performances of the expressive movement task varied widely. Significant differences between the acceleration trajectories were found between performances of the individual excerpts when comparing across subjects. These differences could be attributed to particular time points in the swing, located at the lowest part of the swing trajectory, and the transition to and from this lowest point. A MDS of this data shows that the performances of individual excerpts cluster based on the communicated feeling properties of the excerpts themselves. Taken together, the results examining the

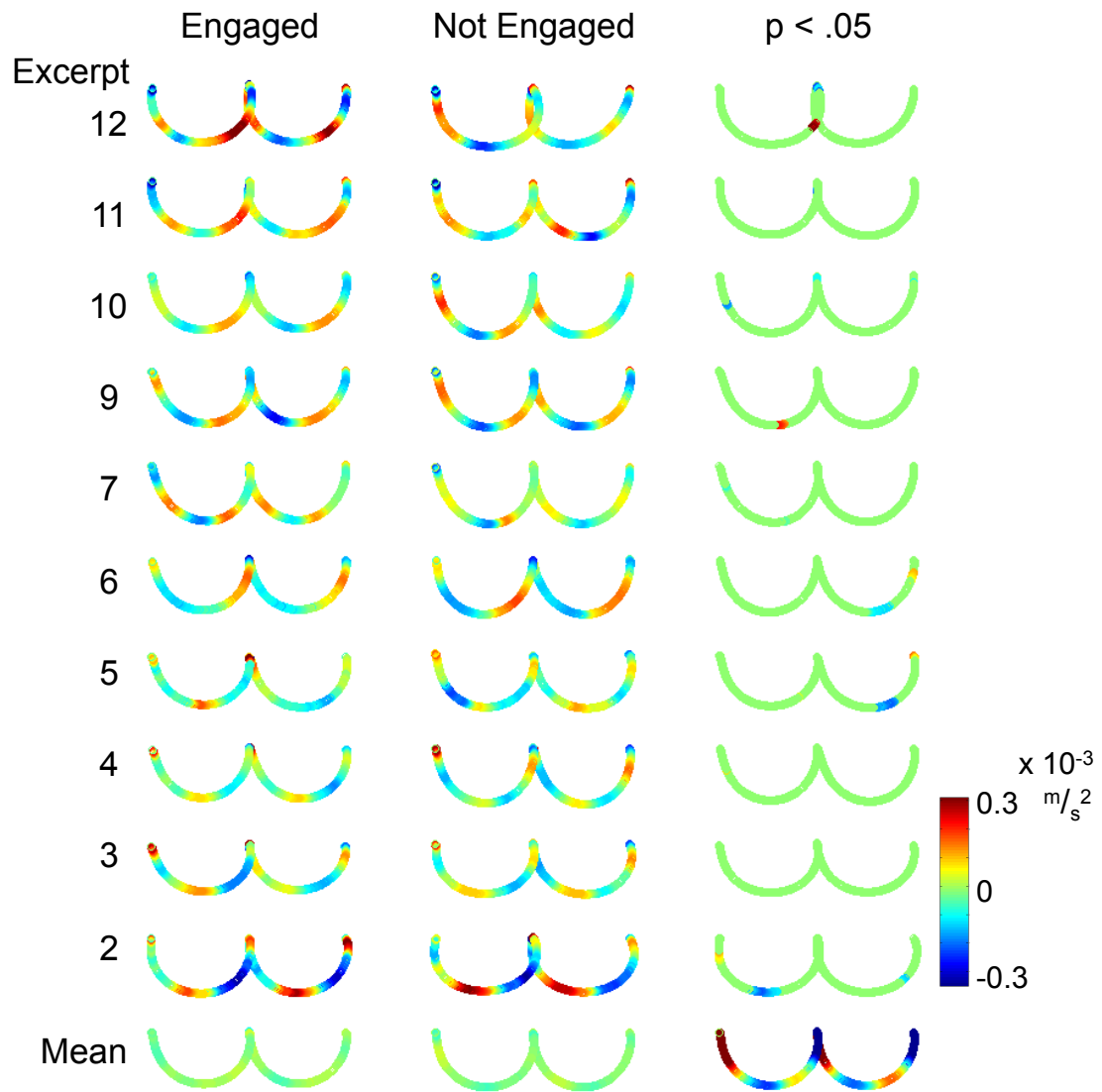


Figure 2.11: A plot of average Right-Left-Right swing trajectory colored by acceleration difference, masked at a $p < .05$ significance level, shows that there are significant differences between the E and NE conditions in the motion capture data. These differences vary by song.

engaged trials show that subjects performed each excerpt according to its feeling intention. However, individual differences still existed in how each subject translated that feeling intention into a repetitive movement sequence.

A comparison of the engaged single-task condition and the not-engaged dual-task condition show significant differences in the acceleration trajectory that vary by song. However, these trials did not cluster based on condition, suggesting individual differences in the quantity and nature of the distractor task's effect on the expressive movement made by the participants.

2.2.3 EEG

Figure 2.12 shows the clustered dipole results across all remaining subjects in the analysis plotted as topographic projections onto the scalp's surface. The clusters lie in the associational areas of the frontal, temporal, parietal, and occipital lobes, showing that a variety of functions relating to auditory and visual processing, motor planning, and emotional processing supported this engaged listening task.

Scalp Maps for Engaged Condition

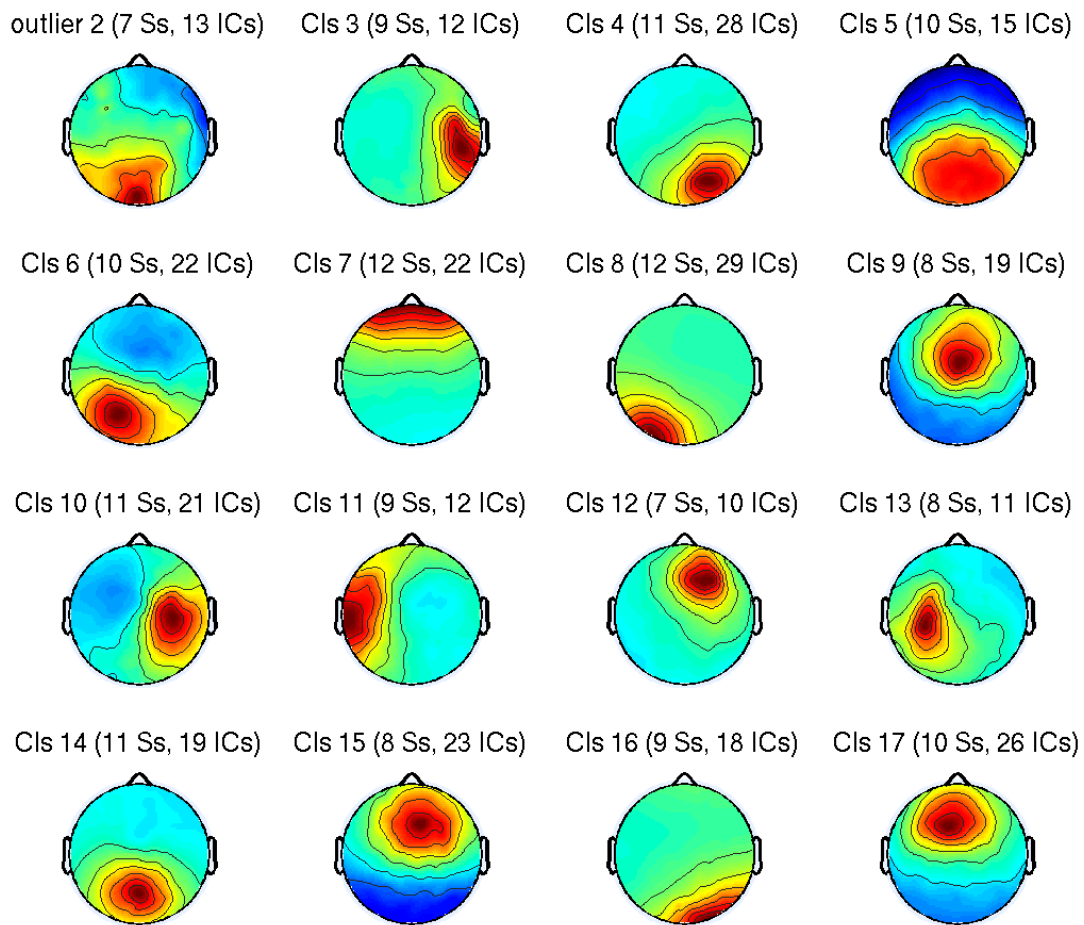


Figure 2.12: The independent components found in each subject's EEG data were clustered according to equivalent dipole map and spectral information. These topographic projections of each cluster of ICs onto a scalp surface show the areas of the frontal, temporal, parietal, and occipital lobes where the clusters were found.

Difference between engaged and not-engaged conditions

We sought statistically significant ways to distinguish the engaged trials from the not-engaged trials, and focused on the spectral perturbation patterns, time-locked to the repetitive expressive movement, that defined each condition, as calculated using the ERSP method described in section 2.1.8. It is possible to compare the spectral perturbation patterns between the engaged and not-engaged conditions by subtracting the not-

engaged condition's ERSP information from the engaged condition's ERSP information. This subtraction removes any patterns from analysis that exist in both conditions, and are thus irrelevant to the comparison between conditions. The resulting difference ERSP pattern characterizes the spectral perturbation that is specific to the engaged condition.

We calculated this difference in ERSP between the engaged and not-engaged conditions, for each cluster of independent components. We then ran the classical parametric statistical tests on these differences in ERSPs, always with the null hypothesis that there are no significant differences between the engaged and not-engaged conditions. However, since these tests were run pixel-by-pixel, yielding a large number of statistical inferences (the "multiple comparisons" problem), even with a $p < .05$ threshold applied, we would still expect to see 5% false positives in our results.

To correct for multiple comparisons, we applied a nonparametric permutation test originally developed for functional neuroimaging [Holmes et al., 1996] [Nichols and Holmes, 2001] and recently applied to ICA-decomposed EEG data in [Miyakoshi et al., 2010]. In this method, we first calculated uncorrected p-values for every pixel. The distribution of t-statistics, calculated from a t-test comparing the engaged and not-engaged conditions, was computed from 2,000 random permutations for each time point of each IC. The one-tailed cutoff value was set to 0.01, which for 2,000 permutations, corresponded to 20 values in from the end of the distribution. Any observed time-frequency responses at the pixels where the values exceed this limit suggested a statistically significant difference. Then, the correction for multiple comparisons was performed with a single-step nonparametric multiple comparisons permutation test, which has been proven to have fewer false positive outcomes [Holmes et al., 1996]. For the multiple comparisons test, the maximum value for every random permutation time-frequency map was extracted. We determined the one-tailed cutoff value ($p < .01$) for the distribution of these maxima by selecting the 20th highest value. Every pixel representing a time-frequency point with observed F- and t-statistics greater than this cutoff value was deemed statistically significant, and all other pixels were masked.

We plotted the difference in average ERSPs between the Engaged and Not Engaged Conditions, and masked at a $p < .01$ significance level. In cluster 4, a pattern of low alpha- and theta- synchronization time-locked to the swing onset emerged, as seen

in Figure 2.13. This dipole cluster is centered around Brodmann area 39, known as the angular gyrus of the parietal lobe, located near the top edge of the temporal lobe. Figure 2.14 shows the locations of the dipoles contributing to cluster 4, projected onto MRI images from the MNI template.

Cluster 4 lies in the parietal-temporal-occipital (PTO) association region of the right hemisphere, where the posterior, parietal, and occipital lobes meet, behind the primary sensory areas, and in front of the visual association areas. The result shown in Figure 2.13 suggests that emotionally engaged expressive movement planning causes a burst of alpha and theta synchronization in this area. In the following we describe this region of the brain in detail, and present functional neuroimaging studies that clarify the role this region of the brain plays in creating both the sense of agency and theory of mind.

2.2.4 Discussion

Each primary sensory area (auditory, touch, motor, and visual) of cortex is connected directly to its respective sensory organs. The association areas of cerebral cortex surround these primary sensory areas and are responsible for gathering the sensory information produced after a first level of processing (such as colors and edges) into a coherent percept (such as a human face). The associational cortical areas are in general the most high-level, developed part of the cerebral cortex. The most complex subset of these areas are the multimodal association areas, which receive input from multiple association areas and combine them into multimodal percepts. For instance, Wernicke's area, located in the posterior temporal area of the dominant hemisphere, uses information from auditory, visual, and motor pathways to facilitate the comprehension of speech.

Multimodal association areas exist in the frontal and limbic lobes, in addition to the parietal region where Cluster 4 was found. The anterior multimodal association area, located in the frontal cortex, links information from unimodal association areas for use in memory and planning, whereas the limbic association area links sensory input with emotional processing. The PTO association region is part of the posterior multimodal association area, which lies at the junction of visual, auditory and somatosensory

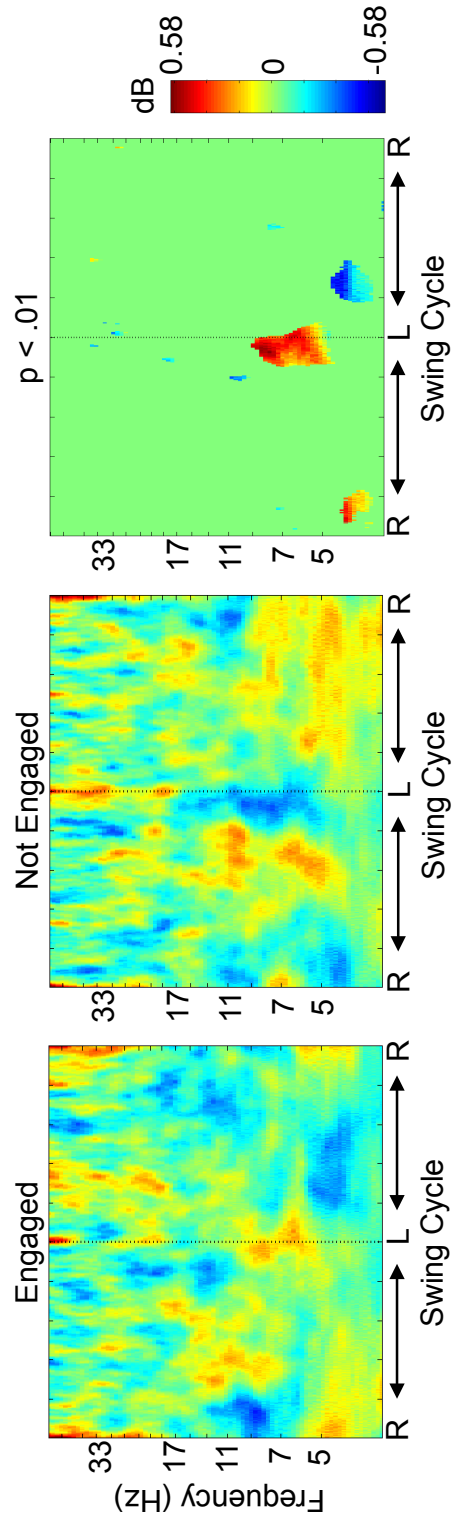


Figure 2.13: The averaged Event-Related Spectral Perturbation (ERSP) for cluster 4 differs between the engaged and not-engaged trials. The difference between the two ERSP plots, masked at a $p < .05$ significance level, is shown in the third column, revealing low alpha- and theta- synchronization specific to the engaged condition, and time-locked to the swing cycle.

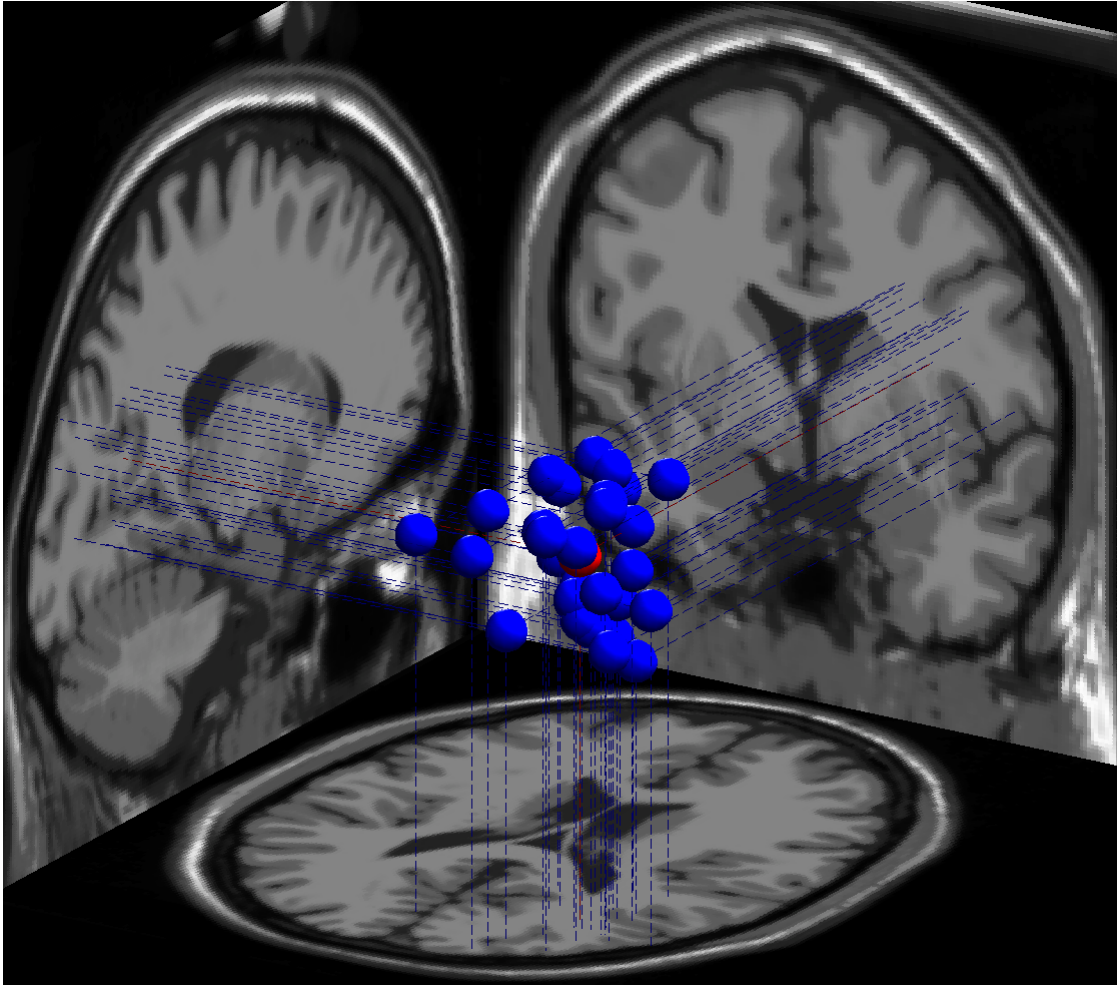


Figure 2.14: The estimated dipoles (blue) and their centroid (red) contributing to cluster 4, as projected onto axial, coronal, and sagittal plane sections from the MNI template.

association area. It is responsible for higher-order cognition, motor planning, spatial awareness of one's own body, and the maintenance of attention.

The PTO association region, as the name suggests, can be divided into its temporal, parietal, and occipital sub-regions, each of which is responsible for the integration of a different set of sensory inputs. The temporal subsection of the PTO area (middle and inferior temporal gyri) integrates visual and auditory information, making it essential for language comprehension. The posterior parietal subsection of the PTO area integrates visual and somatic information, making it an important area for the execution of complex motor plans. It receives input from the somatosensory system in order to track the locations of one's own body parts, and visual and auditory areas for tracking the locations of objects in one's immediate surroundings to be moved or avoided. Signals are sent directly from the posterior parietal cortex, where movements are planned and strategies are executed, to the basal ganglia, where the signal is given to initiate movement. The basal ganglia is, in turn, connected to the motor cortex and cerebellum, which are responsible for the more detailed, mid-level planning of motion trajectories. They communicate directly with the brain stem and spinal cord, where at this lowest level of motor control, movements are executed by sending messages to the skeletal muscles. Lesions to the posterior parietal area result in sensory neglect, an impaired ability to attend to stimuli arising from the side of the body contralateral to the lesion [Lezak, 2004], apraxia, the loss of ability to carry out complex planned movements such as reaching and grasping, and the more general loss of ability to perceive and remember spatial relationships [Culham and Kanwisher, 2001]. There are also hemispheric specializations within the PTO association area. In the left hemisphere, the PTO is primarily concerned with the integration of visual, auditory, and motor information into language comprehension. In the right hemisphere, where our result was found, the PTO is primarily concerned with spatial awareness and recognition of the spatial properties of visual percepts.

Spatially detailed functional analyses of this area show that it is important in creating the sense of agency, or the feeling that one is causing an action or controlling a movement with one's own body. In an fMRI study of subjects controlling an on-screen character whose movements periodically deviated from the instructions they gave it, ac-

tivation was observed in the right TPO junction, suggesting that this area serves as the neural bases of agency [Yomogida et al., 2010]. In fMRI and PET studies of subjects controlling a virtual hand, the less the subject felt in control of the movements of the virtual hand, the higher the level of activation in the right inferior parietal lobe [Farrer et al., 2003][Leube et al., 2003]. An fMRI study of schizophrenic patients shows a hyperactivation in the right parietal area associated with feelings that one's own movements originate from outside one's own body, out of one's control [Ganesan et al., 2005]. A PET study demonstrated that discriminating between one's own actions and those generated by others during a "perspective taking" exercise involves right inferior parietal cortex [Ruby and Decety, 2001]. When shown actions of dogs and humans in an FMRI, subjects showed increased right rostral parietal activation when shown human actions, suggesting that actions that are part of the motor repertory of the observer are processed here, rather than other actions, which are processed in more visual areas [Buccino et al., 2004]. In addition, creating "virtual lesions" in this area using transcranial magnetic stimulation (rTMS) interfered with a subjects ability to determine if one's own movements were made by their own hand, or a virtual hand [Preston and Newport, 2008], and with discriminating self-faces from other familiar faces [Uddin et al., 2006].

Other fMRI studies show that this area is critical for the formation of theory of mind, the capability to reason about another person's thoughts, which is important for social interaction [Decety and Lamm, 2007]. One study of participants reading stories involving physical detail of characters, mental states of characters, and non-social control stories, showed that the right temporo-parietal junction was specifically involved in the reasoning about the contents of another persons mind, as opposed to other socially relevant information about that character [Saxe and Kanwisher, 2005]. This role of the temporal-parietal junction in the theory of mind was also demonstrated in a similar fMRI task using non-verbal cartoons instead of stories [Vollm et al., 2006]. Another fMRI study compared adults with autism spectrum disorder, known to have deficits in theory of mind, with healthy controls; both were given the task of making physical judgements and judgments about the mental state of two target individuals. In the control subjects, the right temporal-parietal junction was selectively more responsive to the mentalizing task than the physical judgement task. This distinction between tasks was not as present

in the ASD subjects, with the degree of specialization of the RTPJ anticorrelated with the degree of social impairment [Lombardo et al., 2011].

In our experiment, the right-parietal temporal region was found to have increased alpha synchronization during the engaged listening condition, time locked to the initiation of the expressive movement task. This region is clearly important for constructing theory of mind, and maintaining one's sense of agency over one's body and actions. Our results suggest that engaged communication of musical feeling modulates attention towards the body and the inner mental life of others along the same time course as the actions taken to express those musical feelings.

2.3 Conclusion

We recorded the movements of our experiment participants during an expressive listening task, and analyzed the position and acceleration profiles of the resulting movement traces. We found that adding a distractor task interfered with the participants' communication of the feeling intention of the piece by altering their movement trajectories. Analysis of EEG data recorded during the experiment demonstrates that the distractor task interfered with activation in the temporal-parietal-occipital junction, an area known to be involved in formation of theory of mind, and the sense of agency over one's own movements and actions.

In the next chapter, we describe an experiment in which we show animations of our participants' performances to a second set of subjects, in order to see if our participants were able to effectively communicate the feeling intention of the music excerpt they listened to.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Leslie, Grace; Ojeda, Alejandro; Makeig, Scott. Measuring Musical Engagement Using Expressive Movement and EEG Brain Dynamics.

Chapter 3

Audience Experiment

In the previous chapter, we determined that the conducting study participants, when given the task of moving expressively to music excerpts, behaved in predictable ways based on the music excerpt, or their level of engagement.

In the experiment discussed in the present chapter, we showed silent animations of our conducting participant's movements to an Internet audience. The Internet study participants were then asked to describe the feeling tone of the silent animations, using a re-worded version of the Music Feelings Questionnaire given to the conducting study participants. They were also asked to compare animations from engaged and non-engaged trials, to see which better communicated the feeling intention of the performer.

We hypothesized that the results from the conducting study questionnaire and the results from the Internet study questionnaire would match. If confirmed, this correlation would show that the conductors were successful at the communication task we laid out for them. Furthermore, it would tell us that the point-light display was able to capture the feeling intention of the composer. We also hypothesized that the Internet viewers would choose the engaged trial animations as better communicating the feeling intention of the performer. Should this hypothesis be confirmed, it would tell us that the dual-task condition was able to correctly impede an engaged (and engaging) performance. It would also tell us that the movement dynamics characteristic of an engaged musical performance can be captured by a simple point-light display.

3.1 Method

90 participants were recruited via e-mail, Facebook, and Internet listings of on-line psychological experiments. The mean age was 35, with a standard deviation of 14.47 of the 90 participants were male, and 93% were right-handed. All participants completed the Music Experience Survey given to the Conducting Experiment participants. The mean combined number of years of musical training for the Internet Experiment participants was 17.7 years, with a standard deviation of 28.5 years.

The motion capture data from eight of the Conducting Experiment participants was processed into animations that were uploaded to YouTube. We projected the 3-dimensional motion capture data onto a two dimensional plane using principle components analysis (PCA). We then separated the motion capture data for each subject by music excerpt, and by condition (single vs. dual task), and then segmented each excerpt performance by movement cycle. One movement cycle was defined by a swing towards the left, and then a swing towards the right, bringing the hand back to the original position. We then averaged all movement cycles within the excerpt, so that for each subject, we had a single movement cycle that defined that subject's performance of the excerpt.

We then animated the movement cycles by plotting a white disc on a black screen, with a 1 second trail fading from white to black (see Figure 3.1). After excluding datasets with discarded trials due to missing data, we had 8 datasets with a complete set of animations (one for each music excerpt and condition combination, for a total of 160 animations). Six of the datasets were from novice conductors, and the remaining two were from experts.



Figure 3.1: The motion capture data from eight of the Conducting Experiment participants were processed into animations that were uploaded to YouTube.

We repeated each movement cycle 10 times to create an extended version and uploaded each video to YouTube. We embedded these videos into an Internet survey using the SurveyMonkey program. The survey began with an introduction page containing the following instructions:

Welcome to the SCCN Musical Communication and Engagement Experiment

Thank you for participating! You will see a sequence of five short videos, each showing a small white disc moving back and forth on a black background. Each video will play back a series of expressive conducting hand movements that were produced by a music lover as they listened to a short piece of film music.

Through their hand movements, they tried to communicate the feeling of the music they were listening to as they might do to give their deaf friend a feeling of what they were experiencing. Your task is, first, to watch the film clip carefully. As you watch, try to get into the feeling of the music the conductor was trying to express.

Watch the film clip as many times as youd like until you can clearly feel what the conductor was expressing. To replay the film clip, press the Previous button. When you can clearly feel the musical feeling the conductor was expressing, press the Next button to answer three questions about your experience of the feeling of the music.

Try to give answers that most closely match your experience of the musical feeling the hand movements conveyed to you. Thanks again!

The introduction page was followed by 20 pages, each containing one of the embedded YouTube videos on the top center of the page, followed by a short questionnaire consisting of a question inviting the participant to evaluate how well the “animation communicate[d] to you a distinct musical feeling?” followed by the music feeling questionnaire given to the conducting experiment participants, but reworded to fit the silent animation context (see Appendix A.6). Each of the questions was answered along a five point Likert scale.

One video was chosen for each survey page according to a block system. The eight conducting animation datasets were divided into eight blocks, so that the first four datasets were assigned to the first two blocks, and the second half of the datasets were assigned to the third and fourth blocks (see Table 3.1 below). Within each block, two of the datasets had half of their 20 music excerpt animations randomly sampled and assigned to their first block, with the other half of the excerpt animations assigned to the second block. Then, the remaining two datasets within each block had their half of the music excerpt chosen to complement those of the first two datasets, so that each music excerpt was played the same number of times within each block.

Table 3.1: Example experiment block from the Internet audience study.

Dataset	Excerpts					excerpts chosen by
403	6	3	2	11	10	Random Sampling
415	11	7	10	12	9	Random Sampling
408	4	9	12	7	5	Complement of 403
416	2	3	5	6	4	Complement of 415

The presentation of each page of the survey was not randomized, in order to preserve the grouping of the animations by conductor. This was made a priority so that the survey respondents would be able to adjust to the individual style of each conductor.

At the end of each block of five single-subject videos, a sixth page was added. On this page, two videos were presented, one the concatenated mean swings from the ‘engaged’ performances of a particular excerpt by the same subject from the previous five videos, and the other the mean swings from the ‘not engaged’ performances. The subject was asked “Which of the two animations below is better at communicating to you a distinct musical feeling?” They selected the video by clicking on a button placed

above the video.

The survey responses for the rating questions were normalized within subject and within question, by subtracting the mean of all the participant's responses for that question across all music excerpts, and dividing by the standard deviation. This ensured that any participant's individual sensitivity to, for example, sadness, would not skew the overall measure of sadness across subjects.

3.2 Results

3.2.1 Correlation of Internet and Conductor Responses

We ran correlation tests between the responses given by the conductors after hearing each music excerpt and the responses given by the audience experiment participants. The results are shown in Table 3.2. The Transcendence, Power, Nostalgia, Peacefulness, and Arousal excerpt ratings given by the two groups were well correlated ($p < .05$ or smaller).

Table 3.2: Correlation Coefficient (R) between song ratings from the Conducting Experiment and three Audience Experiments.

Rating	Concatenated	Full Performance	Normalized
Wonder	-0.02	-0.16	-0.52
Transcendence	0.72*	-0.24	-0.19
Power	0.73*	0.92*****	0.26
Nostalgia	0.88*****	0.76*	0.38
Peacefulness	0.78**	0.83****	0.39
Joyful Activation	0.59	0.67*	-0.06
Sadness	0.62	0.52	0.55
Tension	0.51	0.52	-0.11
Valence	0.21	0.22	0.41
Arousal	0.77**	0.89*****	0.10

* : $p < .05$, **: $p < .01$, ***: $p < .005$, ****: $p < .001$

In previous pilot experiments, full performances of the videos from the same subjects were used instead of concatenated average swings; the results from this experiment pilot are shown in the "Full Performance" column in Table 3.2. The fact that

only minor differences exist between these two sets of correlation results shows that the averaging and concatenating method that effectively “smooths out” the performance by repeating only the expressive variation recorded across all swings, does not have an effect on the Power, Nostalgia, Peacefulness, and Arousal Ratings. The swing averaging method did have an effect on the Joyful Activation method, suggesting that the expressive variation along this rating dimension is contained in the information between swings. Interestingly, the averaging method created a correlation for the Transcendence rating where there wasn’t one before—removing the inter-swing information somehow increased the regularity of the conductors’ and audiences’ ratings along this dimension.

Another experiment pilot was run using versions of the concatenated swings that were normalized in time (so that each swing took the same amount of time, corresponding to the median swing length across all excerpts, which was 900 msec for each cycle). The swing cycles were also normalized in space so that the moving dot traversed the same distance in that time for each excerpt. The correlation results are shown in the “Normalized” column in Table 3.2. There was no significant correlation between the ratings given by the conducting participants and the pilot participants. Removing the tempo and distance information from the movement trajectories appeared to remove most of the information the audience experiment participants used to judge the animations along the rating scales.

3.2.2 Distinguishing Engaged vs. Non-Engaged Conditions

The extra videos added to the end of each subject block comprised a 2-alternative forced choice (2AFC) test. The appropriate descriptive statistic for a 2AFC test is sensitivity, or the proportion of correct responses [Macmillan and Creelman, 2004]. The sensitivity value for a 2AFC test should be somewhere between 1/2 (random) and 1 (perfect detection). For our experiment, it was .67.

We can confirm that this statistic is significant by looking at where it lies on the cumulative binomial distribution centered around .5, corresponding to the null hypothesis (H_0): that the subjects answered the forced-choice task at random. In our case, the result of 60 out of 90 trials corresponds to a $p(H_0) = .0005$, which confirms our hypothesis that the “engagement signal” is detectable by human viewers.

3.3 Conclusion

We confirmed our hypothesis that the results from the conducting study questionnaire and the results from the Internet study questionnaire would match. This correlation demonstrates that the conductors were successful at the communication task we laid out for them. Furthermore, it tells us that the point-light display was able to capture the feeling intention of the composer. The forced-choice task goes one step further, in proving that that the movement dynamics characteristic of an engaged musical performance, regardless of feeling intention, can be captured by a simple point-light display. We also confirmed our hypothesis that the Internet viewers would choose the engaged trial animations as better communicating the feeling intention of the performer. This tells us that the dual-task condition was able to interrupt the communication involved in an engaged (and engaging) performance. This helps support our result from our EEG analysis by clarifying the behavioral differences between the engaged and not-engaged conditions.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Leslie, Grace; Ojeda, Alejandro; Makeig, Scott. *Measuring Musical Engagement Using Expressive Movement and EEG Brain Dynamics*.

Chapter 4

Conclusion

The experiments discussed in this thesis sought to examine the interplay between expressive movement and the creation of musical engagement. For the most part, our data confirmed our hypotheses concerning the communication of feeling intention through expressive gestures, and the brain dynamics that support musical engagement. In the following we will restate each of our hypotheses and explain how they are supported by the data collected in the conducting and audience experiments.

4.1 Discussion

Can participants be trained to gesture expressively to music excerpts are played for them, in a manner that is consistent with the feeling intention of music excerpts that are played for them? We hypothesized that our experiment participants would be able to convey the feeling intention of the music stimuli we played for them.

Our expressive gesture task was developed over a number of pilot experiments. The experiment discussed in Chapter 2 demonstrates the steps we took to show that participants can be trained to express the feeling intent of music they listen to using rhythmic gestures. The guided relaxation and individualized training sessions were critical in preparing participants to perform an engaged listening task regardless of their music training and music listening habits. Analysis of the motion capture data showed that participants' expressive gestures were significantly different between music excerpts, and that they clustered similarly to their own ratings of the feeling intention of the ex-

cerpts. A separate set of subjects recruited from the Internet viewed animations of this motion capture data, and their ratings of these performances also clustered according to the feeling intention of the excerpts. The similar relationships that exist between the ratings of the music excerpts, the recorded motion capture of the performances, and the ratings given to the performances, show that the feeling intention of the music excerpts was preserved in the expressive gestures of the conductors, and in the animations of these performances.

Can the feeling intention of participants' expressive movements be communicated through a vastly simplified, animated presentation of the performance? And furthermore, which motion dynamics are necessary to properly convey this feeling intention? We hypothesized that the feeling intention of the performers' expressive gestures could be transmitted through point-light animations, and viewers would be able to rate these animations correctly according to the feeling intention of the musical stimuli that inspired them. We also hypothesized that we could manipulate the tempo and overall length of the movement trajectories, and that these ratings would remain unchanged, since we believed the feeling intention could be transmitted through the change in acceleration within each repetition of the movement cycle.

We showed the animations of conductors' expressive movements to a separate group of subjects who then rated these performances on a scale similar to the one used by the conductors to rate the musical stimuli, only reworded to be appropriate for the characterization of visual stimuli. In three separate experiments, we displayed either full-length performances of each excerpt, repeated swings representing the average swing across each excerpt, or repeated average swings normalized for tempo and overall size of swing.

We reported significant correlations between the full excerpt performance animations and the conductors' excerpt ratings for the dimensions of Power, Nostalgia, Peacefulness, Joyful Activation, and Arousal. These correlations suggest that the feeling intention across these dimensions was properly conveyed even though the performers' movements were translated into an animated white dot on a black screen, confirming the first part of our hypothesis. Similar results were obtained for the concatenated av-

erage swing animations: significant correlations were found along the Transcendence, Power, Nostalgia, Peacefulness, and Arousal dimensions. This suggests that feeling intention along the dimensions of Power, Nostalgia, Peacefulness, and Arousal can be conveyed with dynamic variations within a movement cycle, and that global variations across many cycles are not needed to convey these feelings. When we showed versions of these animations that were normalized for tempo and swing length, we found no significant correlations between the conductors' and viewers' ratings. This suggests that either the tempo, the swing length, or both, are important factors for communication feeling intention through rhythmic movements, and acceleration patterns alone cannot transmit this information. This result disproves the second part of our hypothesis.

Can we characterize the brain dynamics that support engaged music listening?

We hypothesized that our engaged music listening task would reveal brain dynamics that agree with current fMRI research on music engagement. We discovered independent components in the frontal, temporal, parietal, and occipital areas that support this hypothesis. A cluster of measured dipoles in the temporal-parietal-occipital junction, an area where sensorimotor information is integrated across domains, showed low alpha- and theta- synchronization which was specific to the engaged trials. This result was supported by results from a second part of the Internet audience study, which showed us that animation viewers were able to distinguish an engaged performance from a not-engaged performance, providing evidence that the conducting participants behaved differently according to their level of engagement with the musical stimulus.

Are the brain dynamics supporting music engagement modulated by expressive movement?

We hypothesized that any brain dynamics revealed in our analysis would vary according to the expressive movements made by the participants. This is indeed what we found. When analyzing the brain dynamics along the time course of the averaged swing cycle, we found that the PTO alpha synchronization appeared as a burst immediately preceding the onset of each swing cycle. This suggests that the repetitive expressive movements performed by the participants played a role in the PTO brain dynamics can be attributed to their engaged listening state.

4.2 Implications

Our experience creating the engaged music listening experimental paradigm, and the motion capture and EEG results we have obtained, have a number of implications for future research of musical engagement.

First, we demonstrated that guided imagery and relaxation exercises are useful tools for preparing listeners for engaged music listening in a laboratory environment. Our reports of this success are mainly anecdotal, in that we did not perform controlled experiments comparing participant behavior between traditional listening experiments and more “mindful” ones. Future studies could quantitatively compare different means of encouraging engaged music listening.

The expressive movement task proved to be a useful tool for focusing listeners’ attention and allowing them to provide feedback about their experience without introducing a cognitive task. It provided valuable behavioral data that we used to evaluate listener’s engagement level and felt emotion. We showed that participants gestured expressively to music excerpts in consistent ways, regardless of their musical training. This should encourage more studies of expressive gesture and its connection to the experience—not just the production—of music. We also demonstrated that participants who are not necessarily trained in dance and music could reliably produce these gestures.

We demonstrated that point-light displays are effective means of removing extraneous variables from a performance, e.g., listener appearance, identity, and facial expressions, while transmitting the feeling intention of the participants’ gestures. This suggests they are a useful tool for the study of the underlying motion dynamics that communicate feeling intention across many domains.

We were able to show significant differences in our recorded motion capture data between music excerpts, and between engaged and not-engaged conditions. However, we have not attempted to describe those differences in quantitative ways. Further work needs to be done to develop methods to analyze the dynamics of expressive movement. Approaches from dynamical systems theory and deep learning could prove useful in revealing the underlying processes that distinguished an engaged performance from an not engaged one.

We demonstrated that brain activation in the right PTO junction is specific to the engaged condition, suggesting that this area is important for the support of musical engagement. This result is promising for the future development of an online EEG engagement-detection system that could focus on the PTO area. Our result was time-locked to the expressive movement, suggesting that the repetitive movement task was important part of the creation of the musical engagement. Future studies should examine if PTO activity is partially dependent on the expressive motion task, or if this area still shows activation selective to engaged listening in the absence of movement.

4.3 Future Directions

The fact that a non-invasive method to monitor musical engagement can be developed may give us a useful and general tool for music perception research, with possible wider applications to music classification, technology, and therapy.

Future experiments can monitor EEG signals recorded from the PTO junction as participants listen to full-length music performances that have been analyzed according to the MPEG-7 descriptors (e.g., spectral centroid, harmonicity, and power) to reveal the low-level features, and their patterns of change over time, that contribute to modulations in engagement. A similar approach could employ music pieces which have been analyzed according to Lerdahl and Jackendoff's Generative Theory of Tonal Music [Lerdahl and Jackendoff, 1983] and measure how the engagement level is modulated by the structure of the music, just as Krumhansl studied tension in her 1997 study [Krumhansl, 1997].

Our paradigm could also be applied to the field of Brain-Computer Interfaces (BCIs), which has expanded in recent years to include systems that track affective aspects of a user's daily experience. The BCI field has traditionally focused on providing control of computers to patients who have lost control of their peripheral nervous system. These systems have made use of signals derived from electrophysiological recordings such as slow cortical potentials, P300 potentials, and mu or beta rhythms recorded from the scalp, in addition to other signals that can be recorded by implanted electrodes [Wolpaw et al., 2002]. This information is then extracted and used for the real-time con-

trol of computers for spelling programs, or neuroprosthetic devices designed to replace paralyzed limbs.

The BCI field is now developing applications suited for healthy users who may have few physical limitations. Many of these proposed BCIs focus on the monitoring of attentional or emotional state, since for this population there can be little need to replace physical control devices. These proposed affective (aBCI) or passive BCIs can then be used to optimize the performance of software by detecting errors [Lehne et al., 2009], monitor the user's level of attention on his or her task [Fairclough et al., 2009] or modify interactive entertainment to modulate his level of attention in the task [Chanel et al., 2008]. A few models for affective brain-computer interfaces (aBCIs) based on musical stimuli have been proposed [Khosrowabadi et al., 2009].

A measure of engagement such as the one proposed in the following may be useful for this subset of the BCI or Cognitive Monitoring field. A measure of engagement computed online from an EEG signal could be particularly useful for the creation of aBCIs for the entertainment industry. Just as [Zander et al., 2010] have created a hybrid active/passive BCI by combining the active eye-tracking paradigm with a passive, EEG-based system, one could combine an active, affective gesture recognition system with a passive musical emotion-detector to create a hybrid music BCI. Such a system has been engineered by the author and colleague, operating both on covert measurements of emotional state to decide on the proper mood of the resulting musical mix, and overt gestural control for finer-scale decisions about musical form and style [Leslie and Mullen, 2010]. A commercially viable system is easily imagined by combining a commercially available, wearable EEG headset (e.g., NeuroSky¹ with a smartphone, many of which have 3-D accelerometer and gyroscope data available for calculation of gestural input.

Finally, many of the experiment participants in our study reported that the expressive, engaged listening task was a pleasurable and cathartic experience for them. A neurofeedback system encouraging a state of high musical engagement, much as previous systems encouraged users to manipulate their levels of alpha or theta rhythms, could be beneficial for those seeking stress relief. Further studies are needed to quantify the effect that engaged music listening sessions, perhaps using our methods, could have on

¹<http://www.neurosky.com>

overall wellness.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Leslie, Grace; Ojeda, Alejandro; Makeig, Scott. Measuring Musical Engagement Using Expressive Movement and EEG Brain Dynamics.

Appendix A

Experiment Narrative Scripts

A.1 Initial Instructions for the Active Music Listening Task

You are participating in a musical expression and guided imagery task. In this experiment, you will hear musical excerpts and move to them expressively.

Feelings are created in the depths of your mind and brain, but they act throughout your body. Because your mind creates them, feelings can be re-experienced at any time and may be powerfully triggered by listening to expressive music. When you dream, you will experience a feelingful world that is created by your mind and brain. In a similar way, while you listen to each piece of music, you will experience a strong, dream-like feeling state.

Imagine that the television screen you are facing acts as a window into the next room, and you have a close friend sitting facing the other side of this window. Your friend is deaf, and cannot hear the music you will be listening to in this room, but they long to share in the experience with you. The moving point of light on the screen is an animation of your arm and hand movements, and is the tool you will use to communicate to your friend the feeling of the music you are listening to or imagining. Here's a short film that should help illustrate the task. [play excerpt from *The Heart is a Lonely Hunter*.]

For each repetition of the music, your role is, first, to experience within yourself

the quality of human mood and feeling that is most naturally expressed by the music you are hearing, and then to express this mood and feeling yourself – through a completely expressive performance of the rhythmic arm and hand movement you have practiced.

Now, let's practice performing this gesture to a musical example.

First, we'll play a repetitive sound to let you know how fast you should perform your rhythmic arm and hand movement. Each repetition corresponds to one cycle of the U-shaped movement. Try practicing along with this sound.

[pause for a few seconds, then whooshes fade to almost silence]

Now, let's start the musical example.

[music example plays]

As you listen to each piece of music, let its feeling and mood permeate you. Deep inside of you is where both feeling and music begin. At first, let the feeling build up; allow enough time for it to fill you. As you enter into the feeling more and more, let it express itself within you, musically. Let it course through your body until every limb and organ are silently expressing its mood and feeling.

Now, feel how your own arm and hand are ready to express the same musical feeling that is now also pouring from within you. Slowly let your own arm and hand movements flow together with the feeling pulse of the music and the performer. Let them completely merge into and express the same musical mood and feeling – as they flow into and through the simple, U-shaped rhythmic pattern. Let the movements flow so as to completely express the feeling flowing from within you from the music you are hearing.

To make your rhythmic performance most deeply satisfying and communicative, invite all of you to take part in it. Let your whole body silently release the mood, pulse, and feeling of the music.

When the musical excerpt fades, imagine the same music continuing. Continue to express its mood, pulse, and feeling completely and expressively with your rhythmic arm and hand movements.

Sustain your feeling and expressive performance for as long as possible or until you hear my voice. Then, slowly return to your quiet, neutral, baseline feeling state. Because we will be recording your EEG, try to avoid making extreme facial movements

or expressions – but don't be rigid. Focus on your whole body experience, while your arm and hand express the feeling of the music as deeply as possible.

You'll perform each excerpt several times this way. I'll loop the samples repeatedly. The first time it plays, simply listen closely. Then, start practicing your expressive movements. You can practice as many times as you'd like, and when you are satisfied with your performance, wave at me through the camera. Then, I'll invite you to perform the excerpt two more times. This will be your most feelingful, expressive performance. Finally, I'll ask you to perform two more times, but while doing arithmetic in your head at the same time. Then, you can stop conducting and fill out the survey on the computer, which will ask you some questions about the music you just heard. We'll go through this process for each of the ten music excerpts. Then, we'll take a short break to allow you to relax and stretch. Finally, we'll do an express version of the first half, without the practicing or survey. In total, it should take about an hour and a half.

We will begin with a guided relaxation to help you achieve a quiet, neutral feeling state. Thank you for participating, and enjoy your journey through your musical feeling.

A.2 Guided Relaxation Script

As you stand comfortably, begin to focus on your breathing. Notice exactly how you are breathing at the present moment. Nothing is more fundamental to life than the breath, and your breathing pattern reflects your internal state. To become calm, slow down your breathing. Relax your upper torso and begin to breathe abdominally. Notice the immediate shift in your body feeling as you take charge of this basic function.

Now take a long, slow breath, inhaling for five long, long seconds

[pause for 4 sec]

then exhaling to the same count

[pause for 5 sec]

Be aware of your body, take a few moments; scan it's every component, shifting your awareness throughout your entire being as you renew your physiological state, moving into a state of quiet, calm relaxation.

[pause for 10 sec]

Allow your imagination to transport you away from your current location, leaving all physical sensation behind you as you do. As you breathe deeply and calmly, imagine yourself lying alone on white sheets in a quiet resort where you are on vacation. You are comfortable, warm and calm. Imagine a warm bath of sunlight coming through a window, gently warming your body. Let its warmth relax all your muscles.

Start at your feet and calves. Notice any tension there and let it relax. Next, move up your body to your thighs and torso. Allow the stillness and warmth to radiate through your muscles and give them a sense of numbness. As you move to your chest and arms, notice once again your breathing and the slow rhythm you are creating.

Remember that focusing on your breathing will always bring you back to this calm, centered state. Now, release any tension in your arms and fingers. Allow them to be still and relaxed, as if they were floating upwards. Now pay particularly close attention to your neck. While keeping your head still, let your shoulders drop and feel the muscles in your neck slowly relax. Take a moment to notice how this feels. [short pause]

Now, shift your attention to your face. Let your cheeks and jaw drop as you let all the muscles of your face relax. Notice how tension is so quickly relieved by this simple action. Remember that releasing the muscles of your face will also help to bring you back to this relaxed, neutral state. As you imagine lying in a bed in the sunlit room, completely relaxed, take a moment to notice how your body and mind feel. This is your calm, neutral state to which you will return following each musical episode.

A.3 Conducting Experiment Episode Instructions

Introduction to the first interval episode: We will now begin playing the musical excerpts. Try to feel the feeling the music expresses fully and completely, to appreciate its depth and sincerity. Pay attention to sensations that experiencing this feeling produces in your own body. Then begin to express this feeling with rhythmic musical arm and hand movements. Perform them in a very expressive way that fully and completely expresses the musical flow and pulse of feeling that you are experiencing.

Transition into each music excerpt: Now, prepare to hear and experience within yourself a new musical feeling and expression. In this episode, listen to the music, paying attention to feeling of the piece. As you perform your movements, concentrate on the animation in front of you. Imagine that you are communicating these feelings to your friend who is watching the animation through the window in front of you. They want to experience the feeling of the musical expression.

Now I'll play it twice through.

Now I'll play it two more times. Please add up the letters to the word [brain, mind, music, ...]

Return to neutral state from each musical excerpt: Now slowly return to your neutral, calm, soothing, and detached state while recalling in a more detached way what you imagined, heard, and experienced during this episode. Again become aware of your breathing as it returns to a slow neutral rhythm. Allow your face to relax. Dissolve all remnants of the emotion you were experiencing. Let go of its associated bodily and musical feelings. Let your entire being, mind, and body, experience the benefit of returning to your quiet, neutral and relaxed state.

Questionnaire: Please take a few moments to fill out the questionnaire on the computer next to you.

Post-session resting baseline: To allow us to record your post-session baseline brain activity, please sit quietly for the next two minutes until you hear my voice again.

Conclusion: Your emotional musical performance journey is now complete. Thank you for participating. Spend as many moments as you'd like now to rest and relax while we continue to record your EEG. When you are ready to move on, fill out the questionnaire, and to have the EEG cap measured and removed, please raise your hand.

A.4 Music Feeling Questionnaire













Please Respond to this Music Excerpt -- 1 of 11

When providing your ratings, please describe how the music you listen to makes you feel (e.g., this music makes me feel sad). Do not describe the music (e.g., this music is sad) or what the music may be expressive of (e.g. this music expresses sadness). Keep in mind that a piece of music can be sad or can sound sad without making you feel sad. Please rate the intensity with which you felt each of the following feelings on a scale ranging from 1 (not at all) to 5 (very much).

	Not at all	Somewhat	Moderately	Quite a lot	Very Much
Wonder <i>Filled with wonder, Dazzled, Moved</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Transcendence <i>Fascinated, Overwhelmed, Feelings of transcendence and spirituality</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Power <i>Strong, Triumphant, Energetic</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Nostalgia <i>Nostalgic, Dreamy, Melancholic</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Peacefulness <i>Serene, Calm, Soothed</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Joyful Activation <i>Joyful, Amused, Bouncy</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Sadness <i>Sad, Sorrowful</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5
Tension <i>Tense, Agitated, Nervous</i>	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5

For the following questions, pick the image that most closely matches how the music actually made you feel.

					
	1	2	3	4	5
How did this music make you feel? (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

					
	1	2	3	4	5
How did this music make you feel? (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Submit

A.5 Music Experience Questionnaire

Please indicate the approximate number of months (since the age of 6) you have participated in any of the activities listed below in the space provided. Music lessons in any style or genre of music on any instrument or voice.

Participation in an instrumental ensemble or choir (including vocal ensembles such as glee clubs, barbershop quartets, and musical theater troupes).

Dance lessons or classes (any style including, but not limited to, ballet, tap, modern, swing, jazz, polynesian, Hawaiian, or hip-hop).

Performance or practice with a dance troupe, gymnastics troupe, circus troupe, or competitive-style jump rope.

Synchronized swimming.

Synchronized diving.

Synchronized skating.

Marching band (including drum major or field major).

Baton twirling.

Group recitation of text.

A.6 Internet Experiment Video Feedback Form



Watch the next video carefully, until you get a sense of the musical feeling the conductor was experiencing.



1. How well did this animation communicate to you a distinct musical feeling?

- Not at All Somewhat Moderately Quite a Lot Very Much

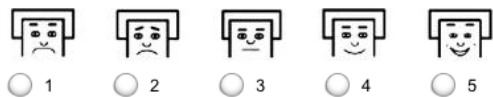
2.

How closely do each of these words describe the feeling tone of the performance you just watched?

	Not At All	Somewhat	Moderately	Quite a Lot	Very Much
Wonder <i>Filled with wonder, Dazzled, Moved</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Transcendence <i>Fascinated, Overwhelmed, Feelings of transcendence and spirituality</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Power <i>Strong, Triumphant, Energetic</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nostalgia <i>Nostalgic, Dreamy, Melancholic</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Peacefulness <i>Serene, Calm, Soothed</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Joyful Activation <i>Joyful, Amused, Bouncy</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sadness <i>Sad, Sorrowful</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tension <i>Tense, Agitated, Nervous</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

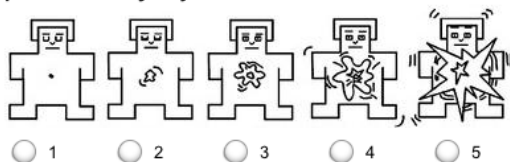
3.

Which image best describes the feeling tone of the performance you just watched?



4.

Which image best describes the feeling tone of the performance you just watched?



Appendix B

Single Subject Motion Plots

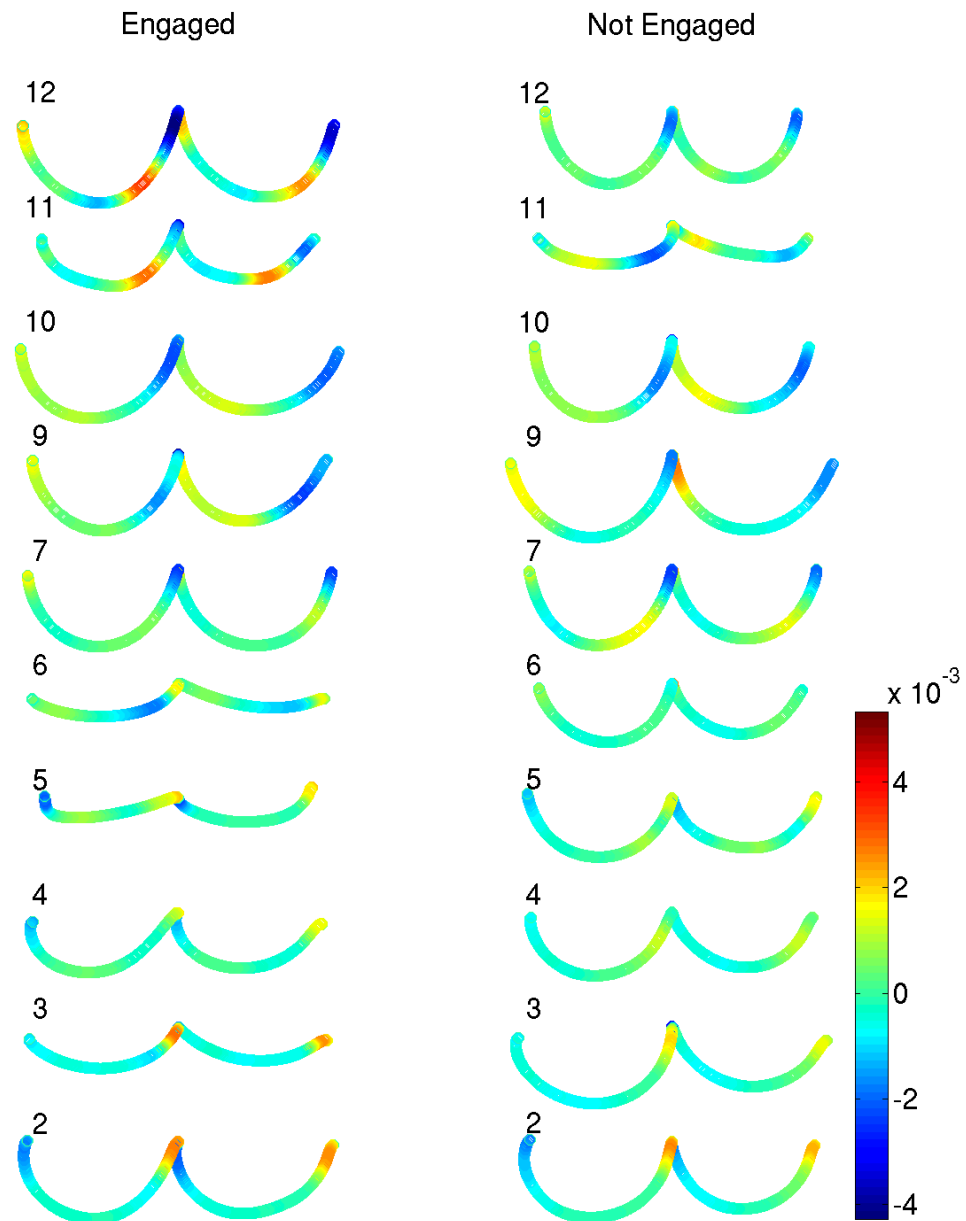


Figure B.1: Movement Trajectories, Subject 357

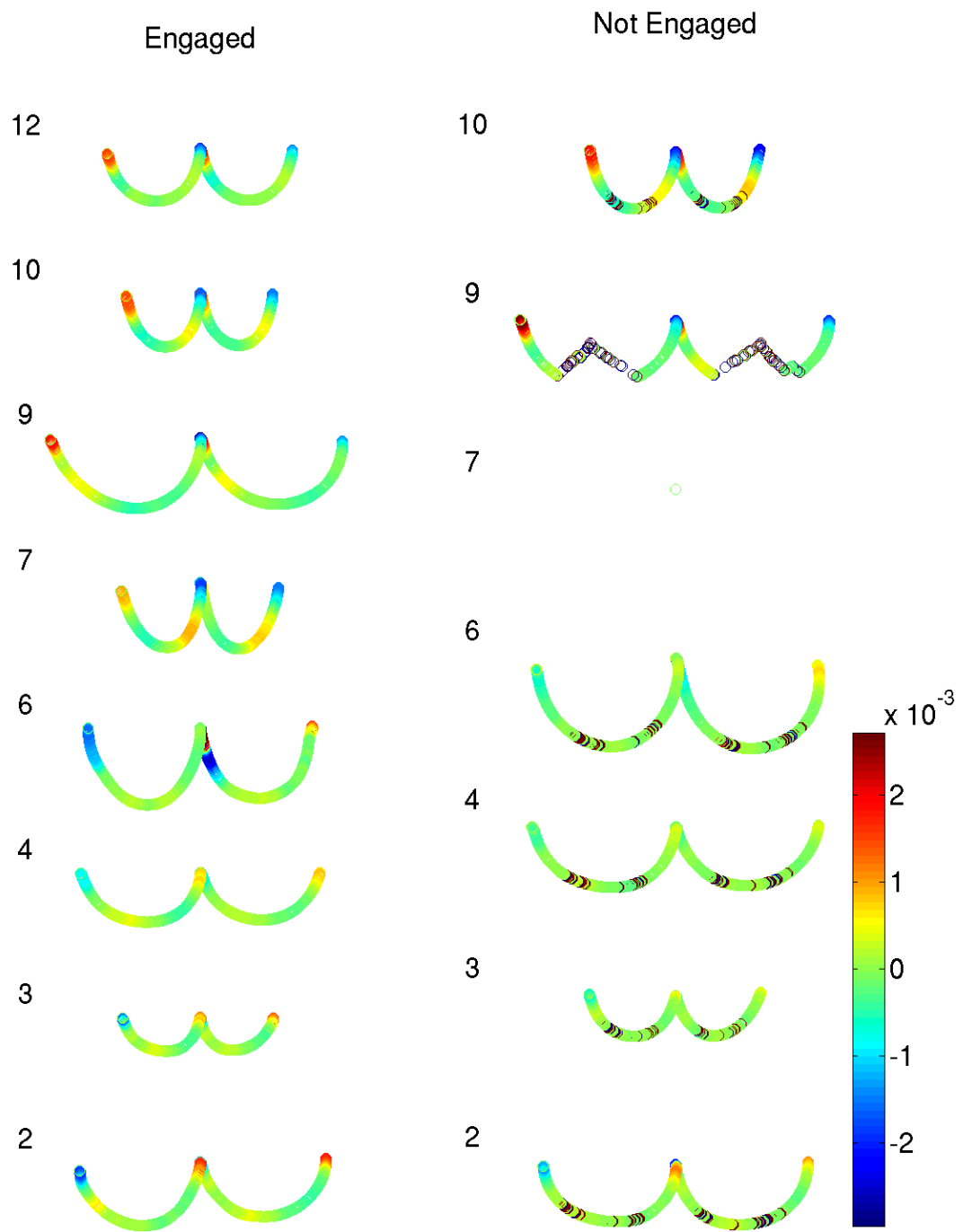


Figure B.2: Movement Trajectories, Subject 362

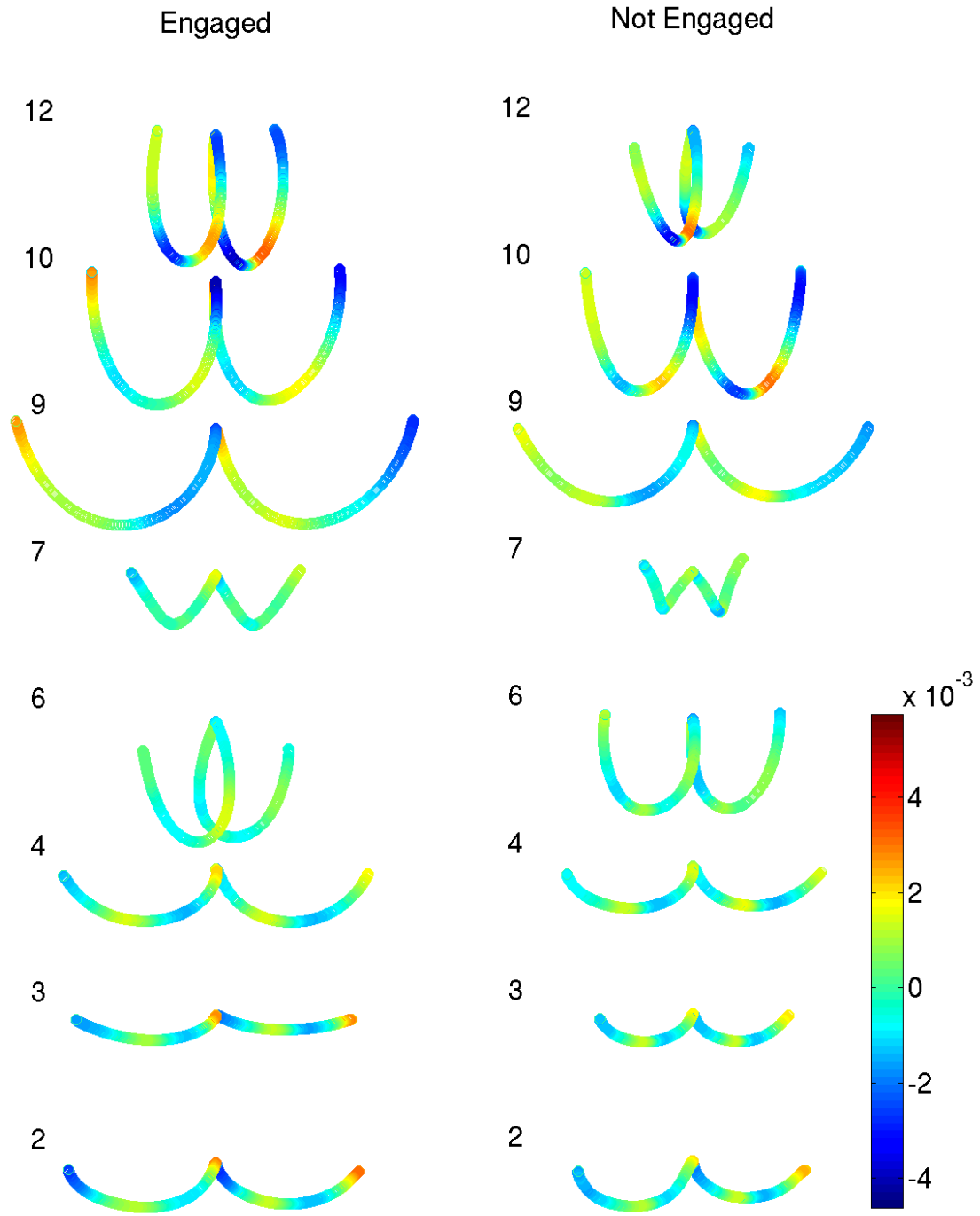


Figure B.3: Movement Trajectories, Subject 370

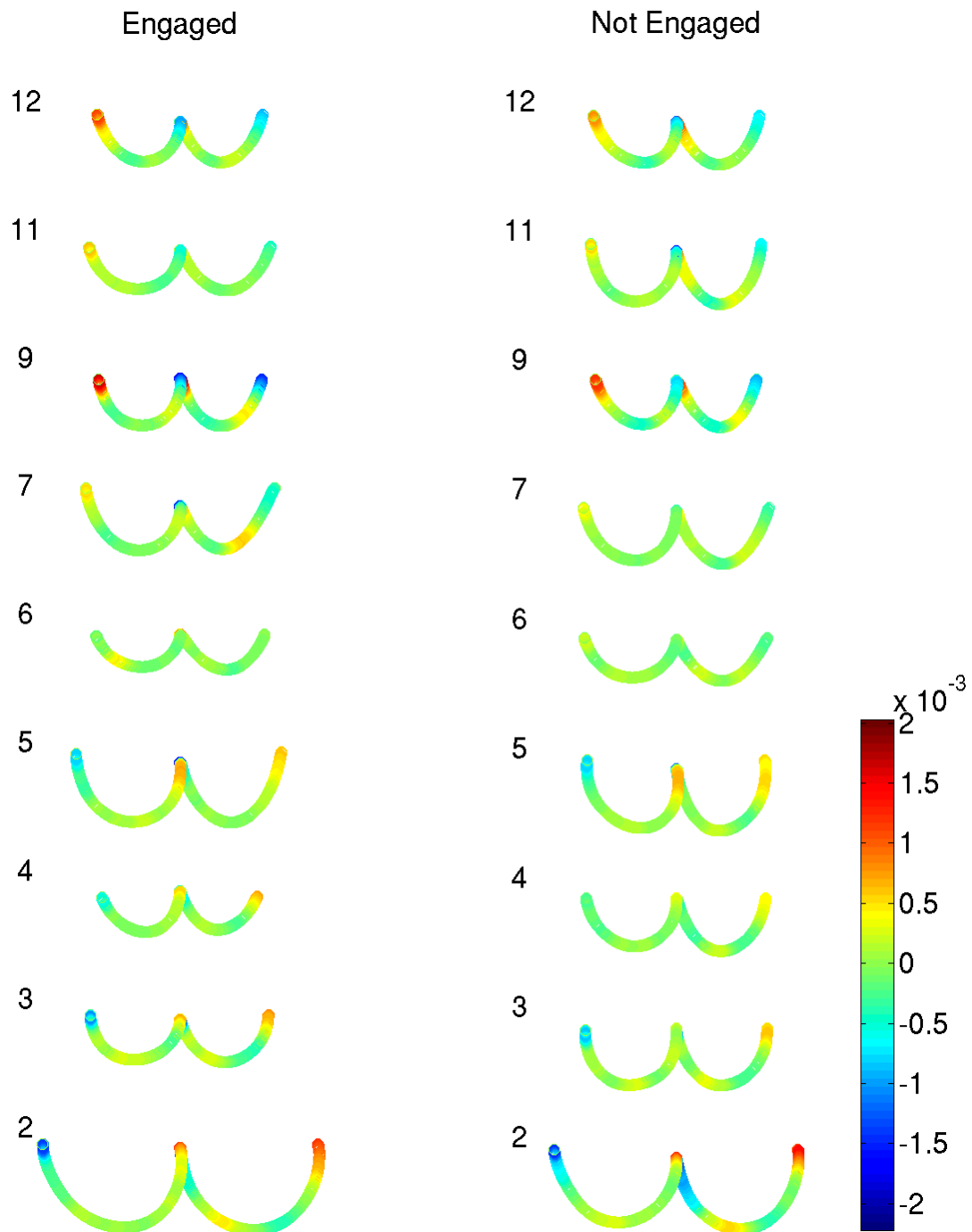


Figure B.4: Movement Trajectories, Subject 371

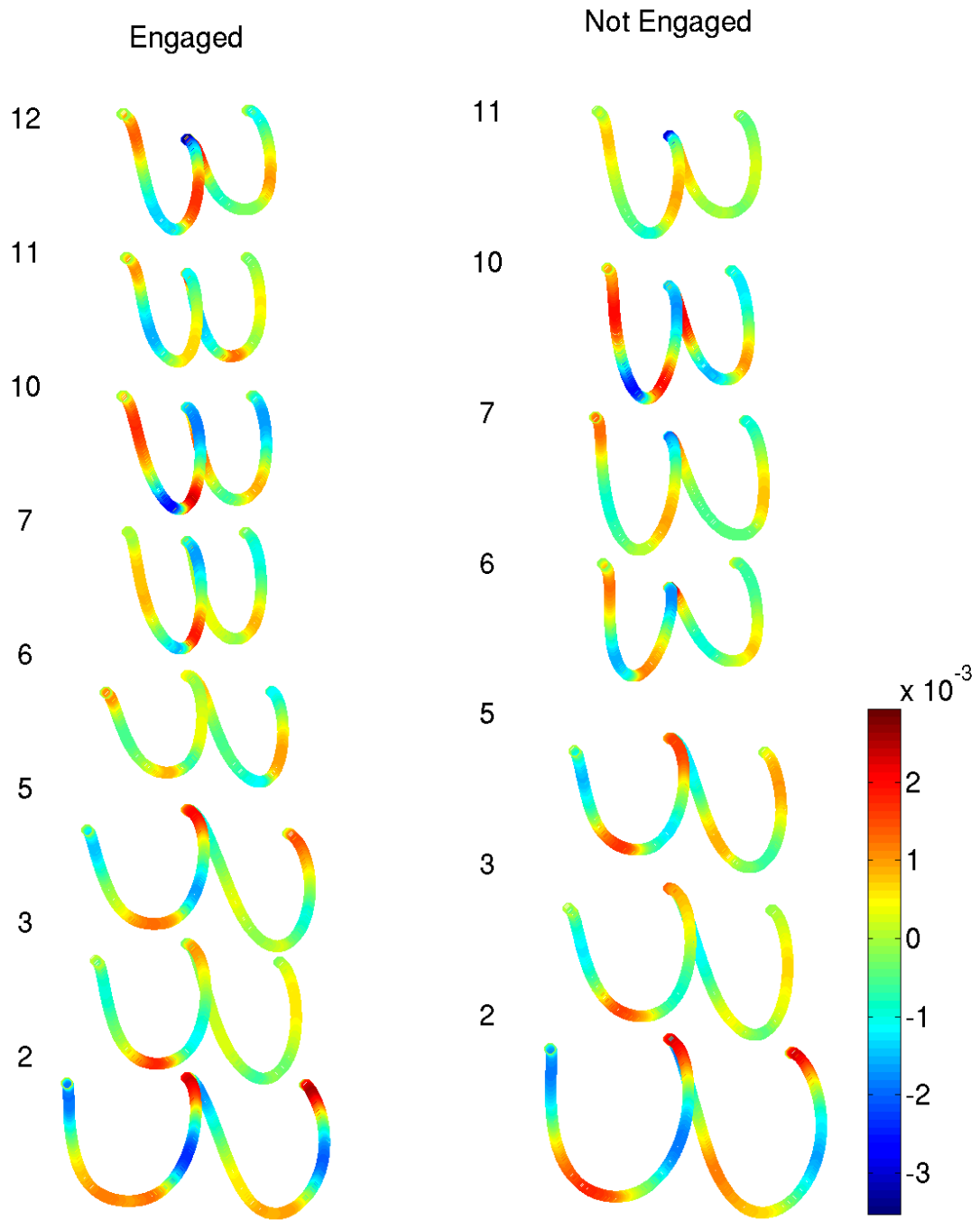


Figure B.5: Movement Trajectories, Subject 372

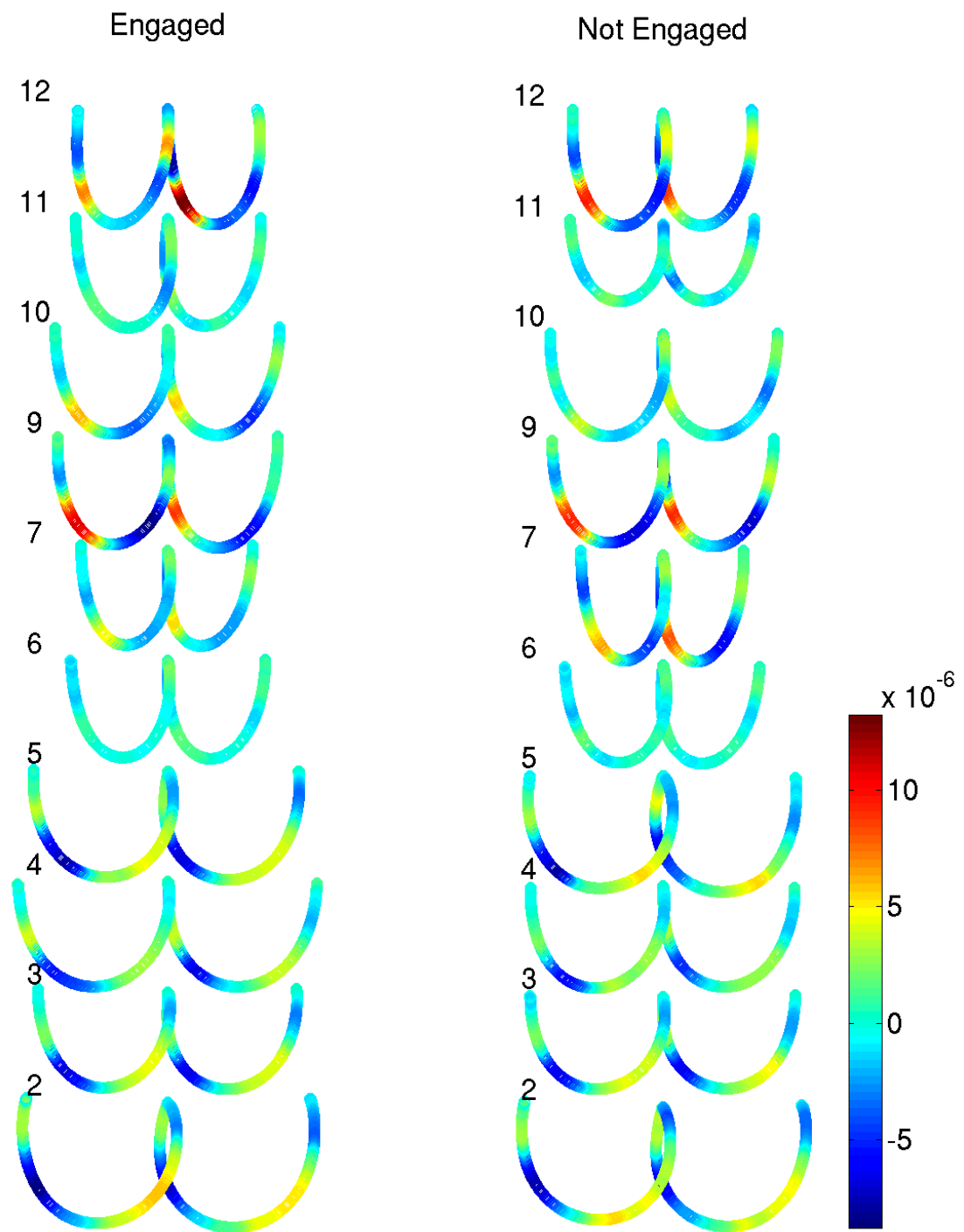


Figure B.6: Movement Trajectories, Subject 374

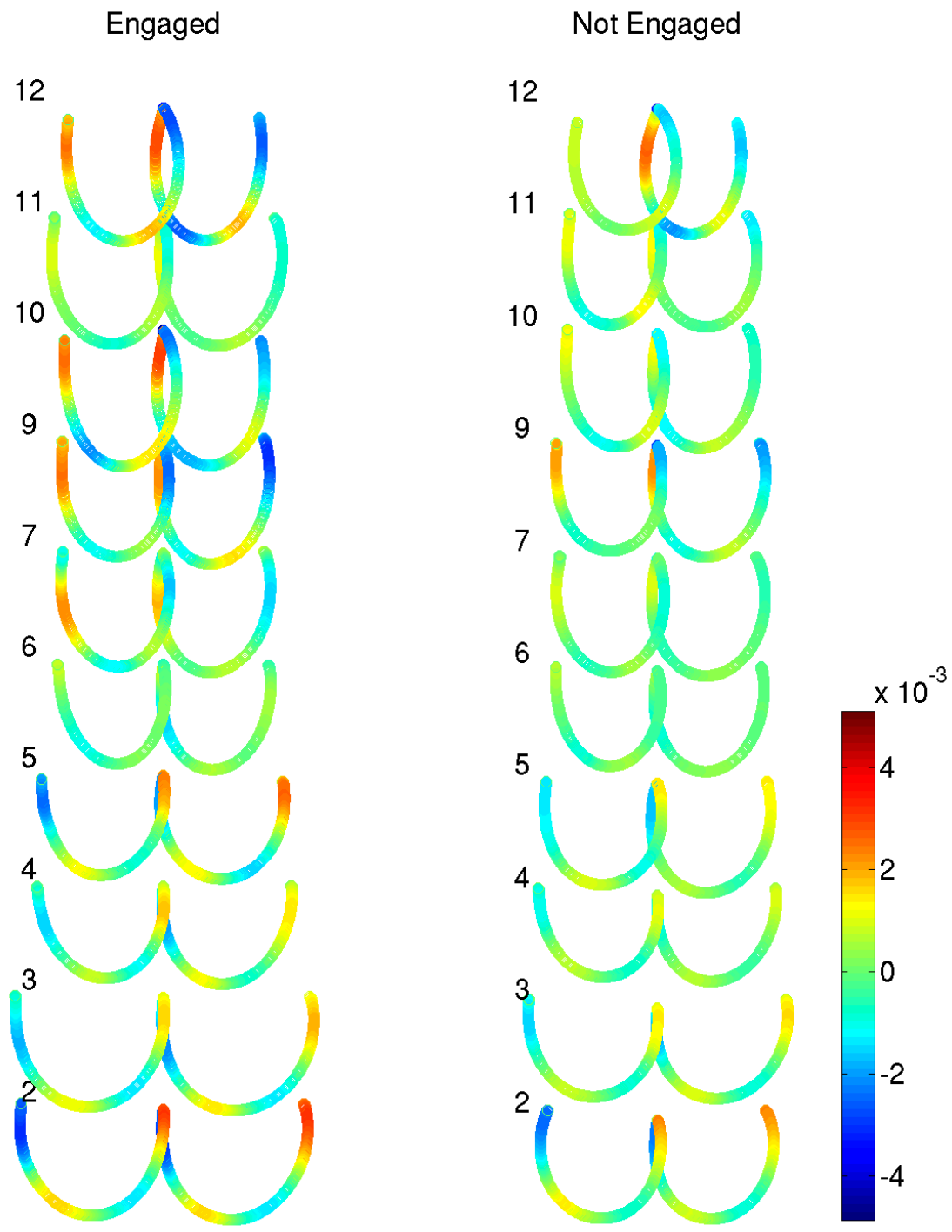


Figure B.7: Movement Trajectories, Subject 375

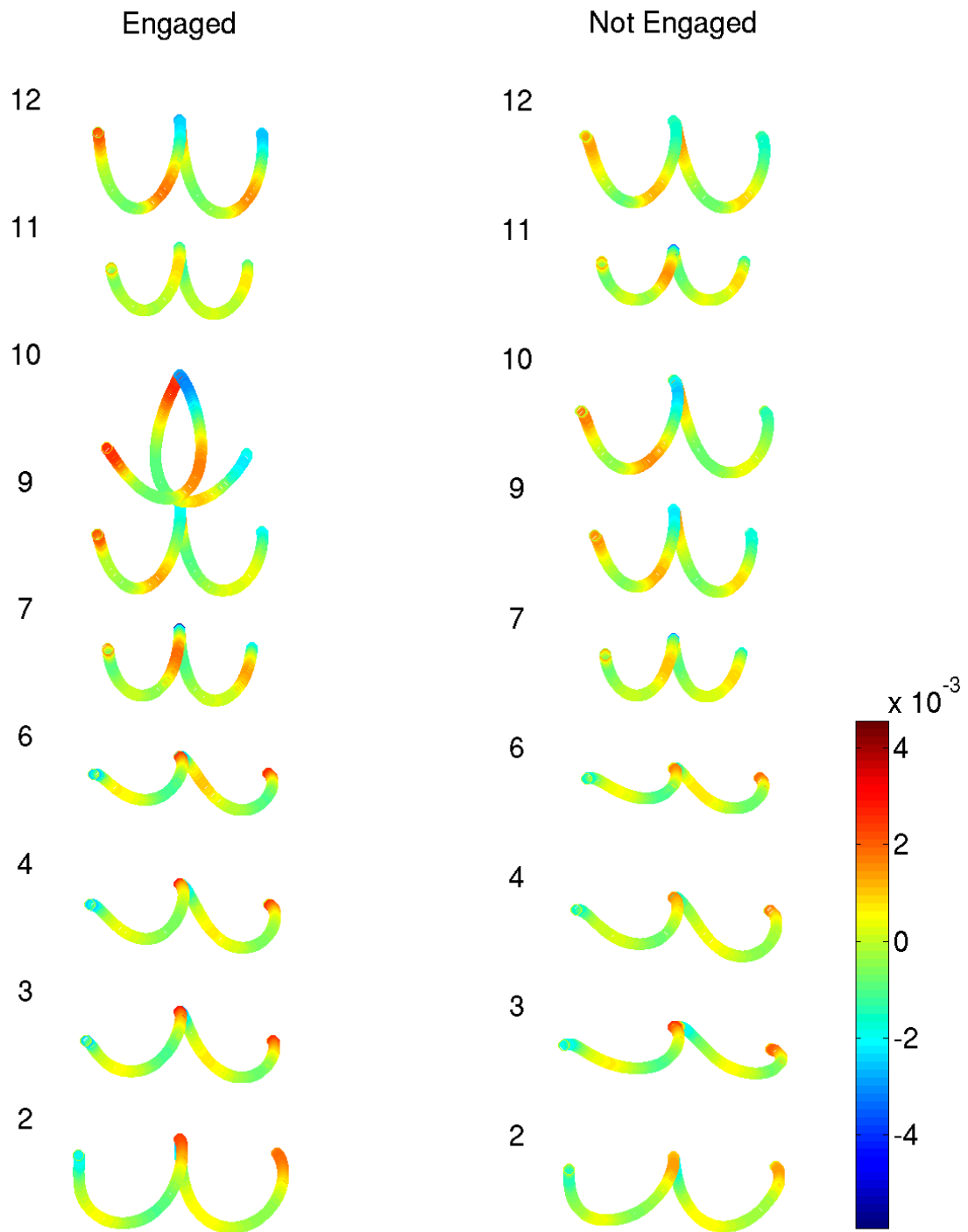


Figure B.8: Movement Trajectories, Subject 378

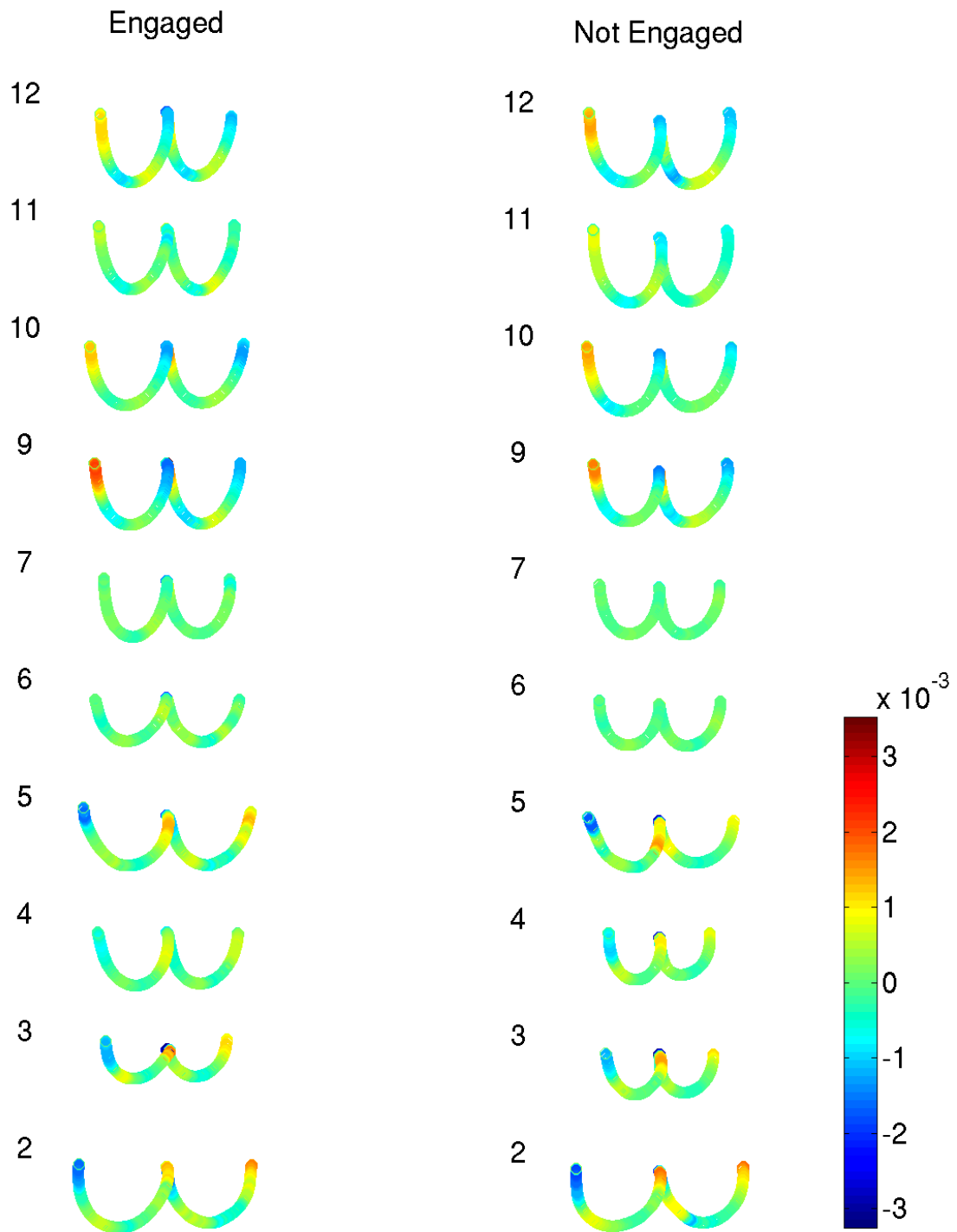


Figure B.9: Movement Trajectories, Subject 408

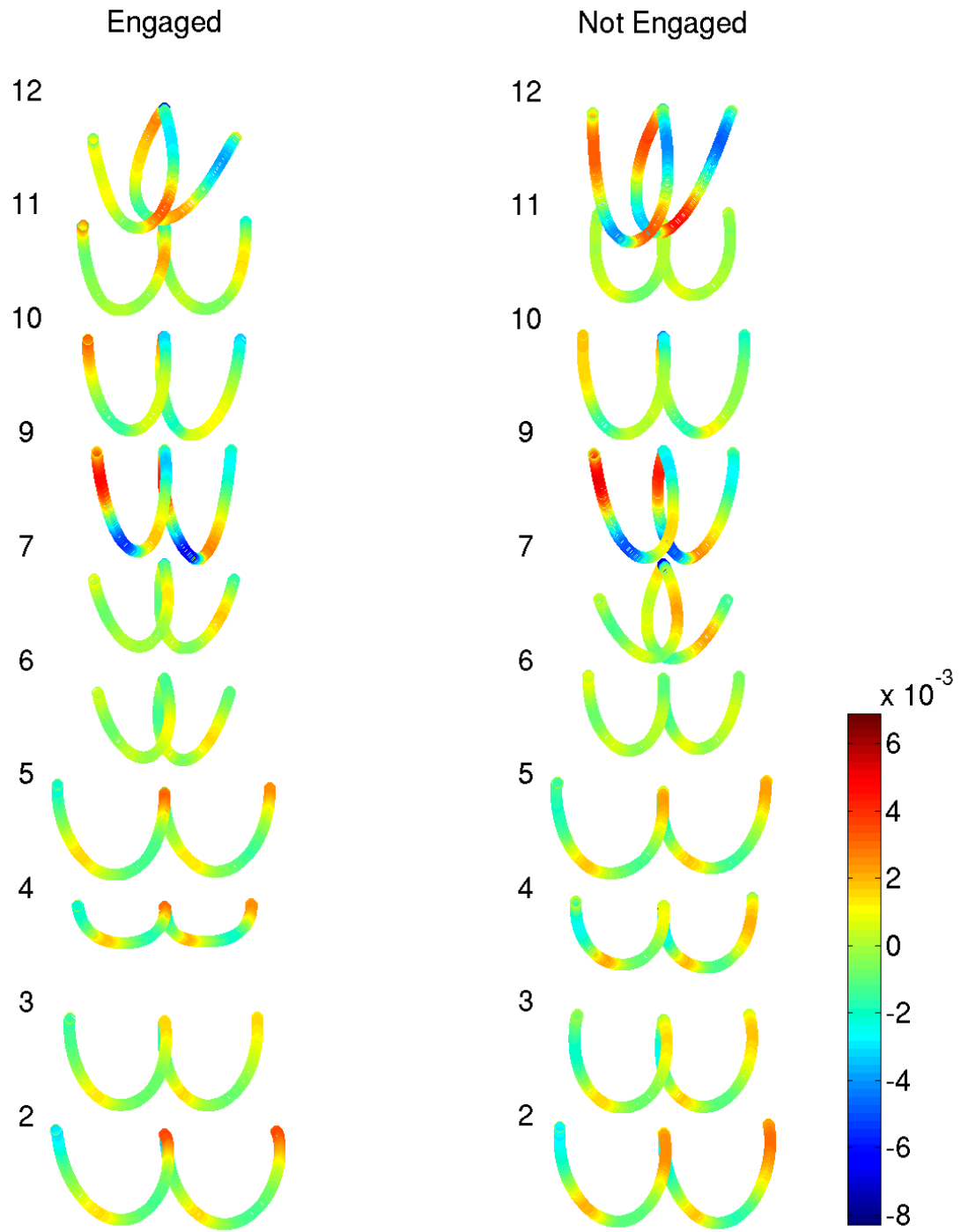


Figure B.10: Movement Trajectories, Subject 411

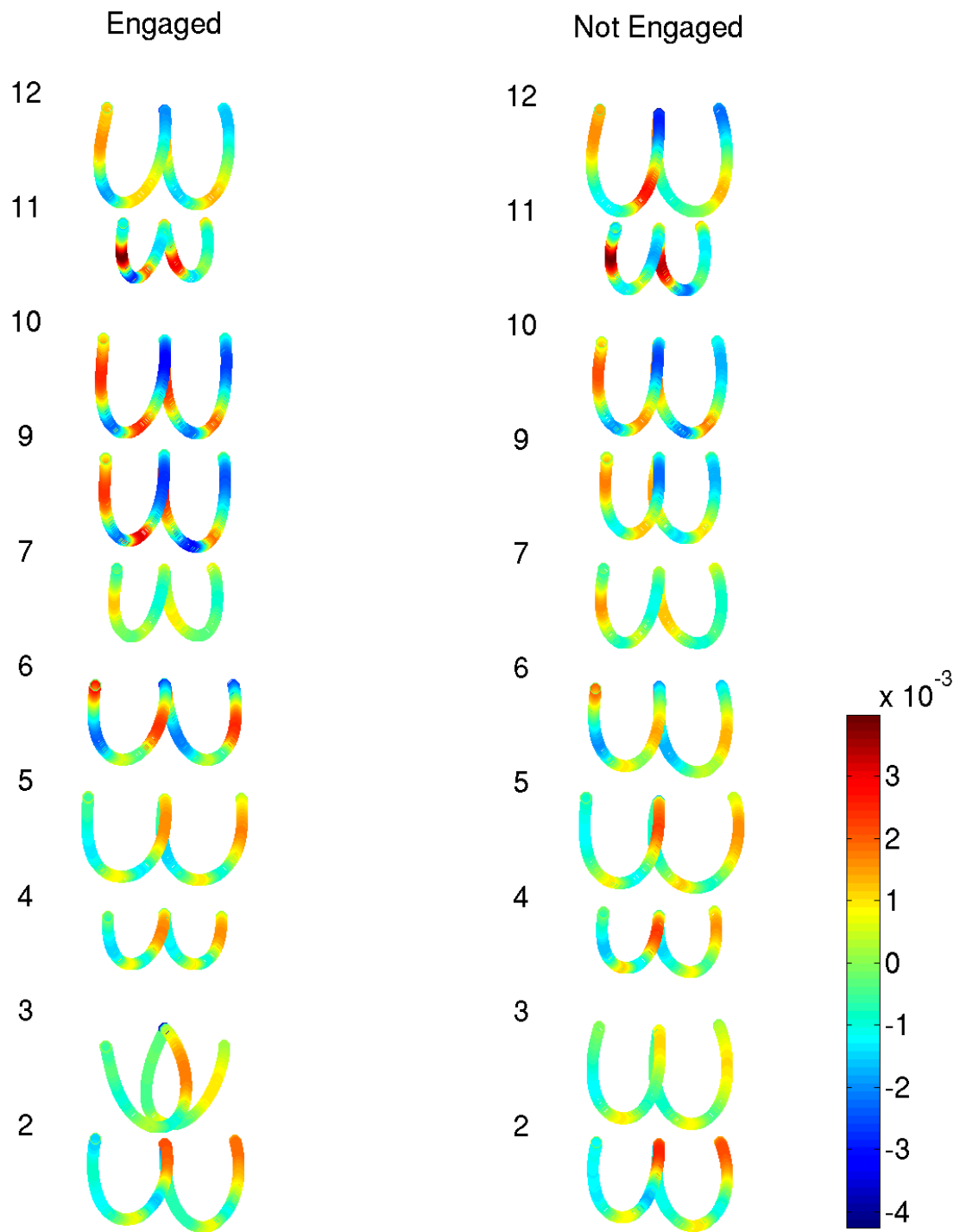


Figure B.11: Movement Trajectories, Subject 412

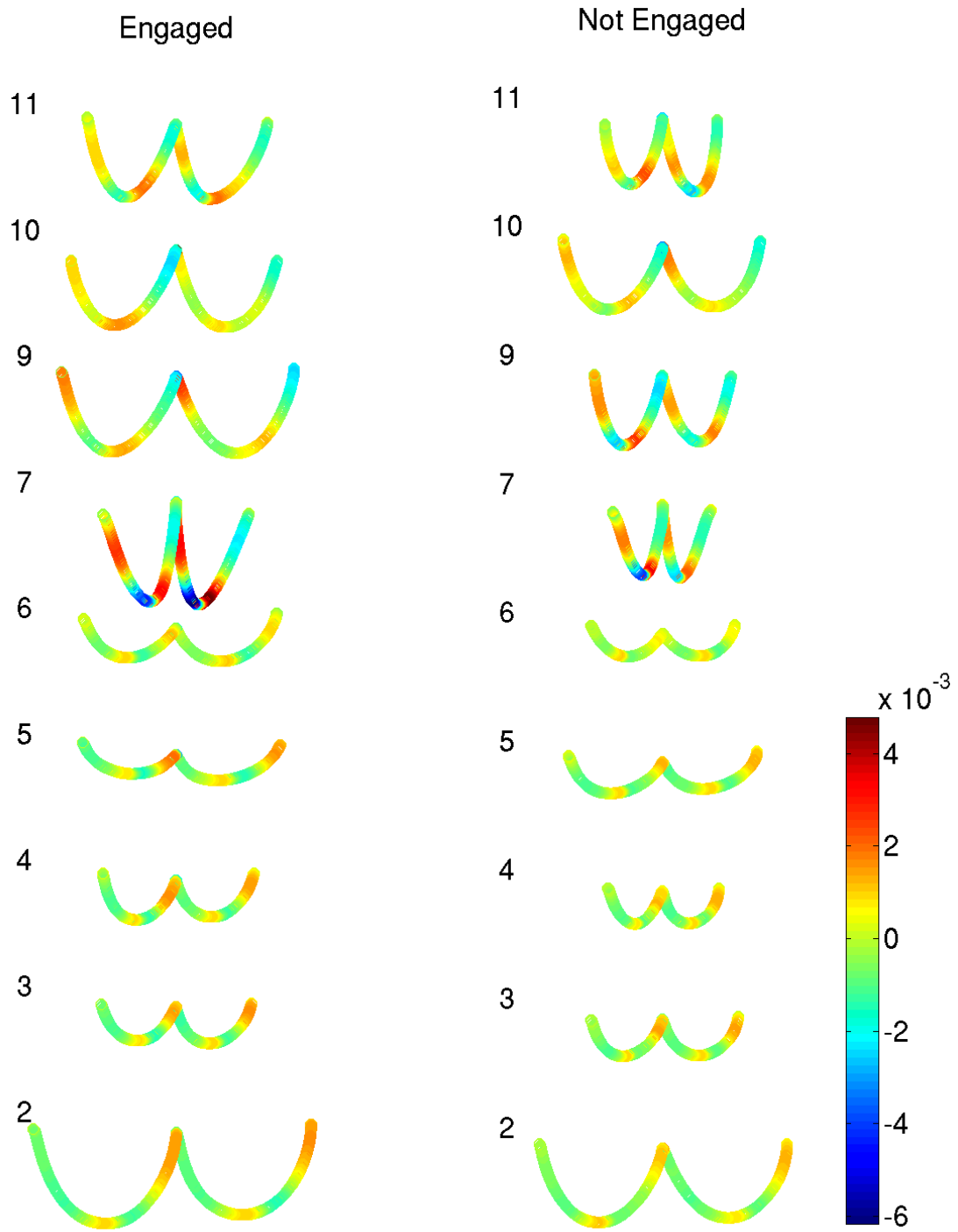


Figure B.12: Movement Trajectories, Subject 413

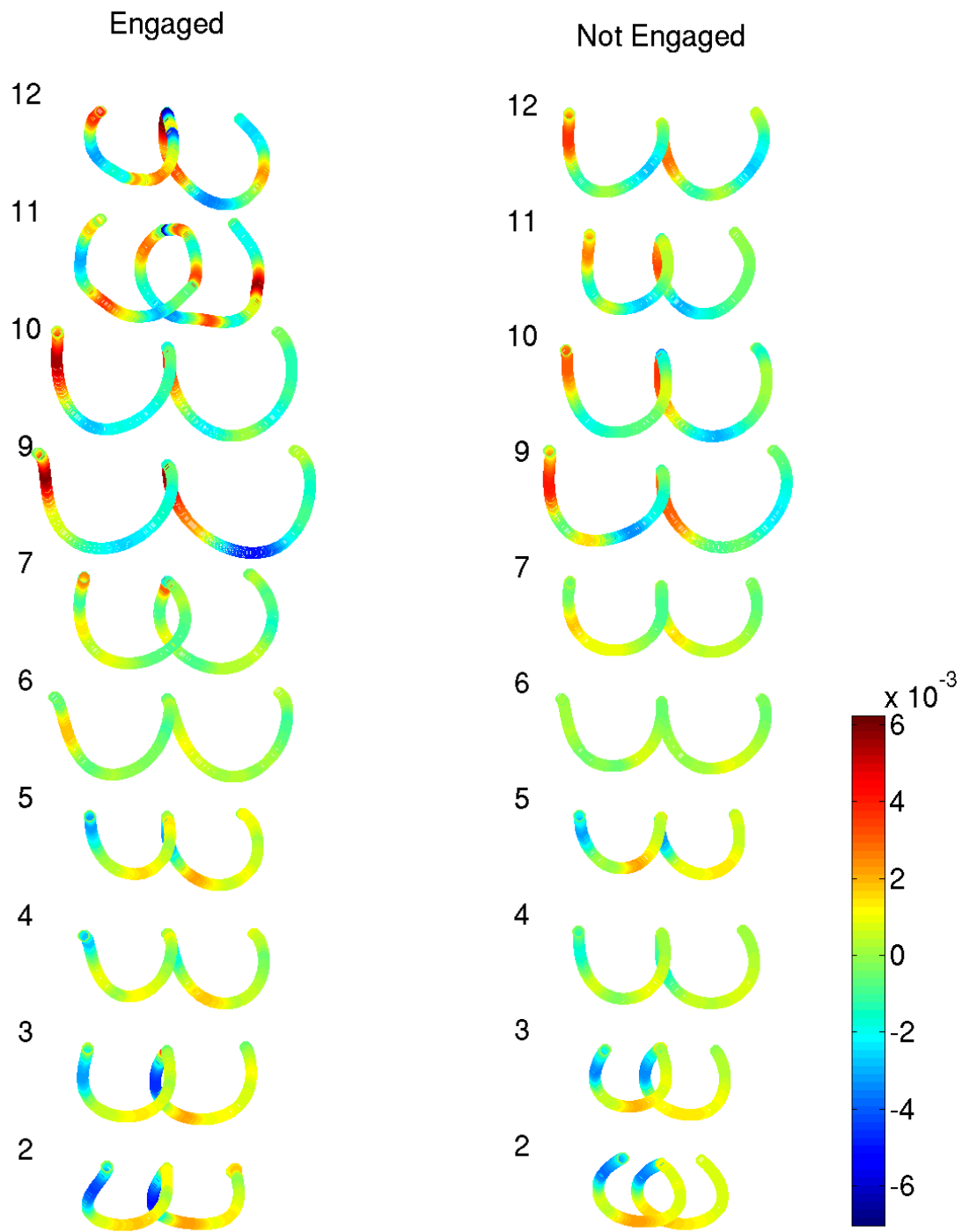


Figure B.13: Movement Trajectories, Subject 414

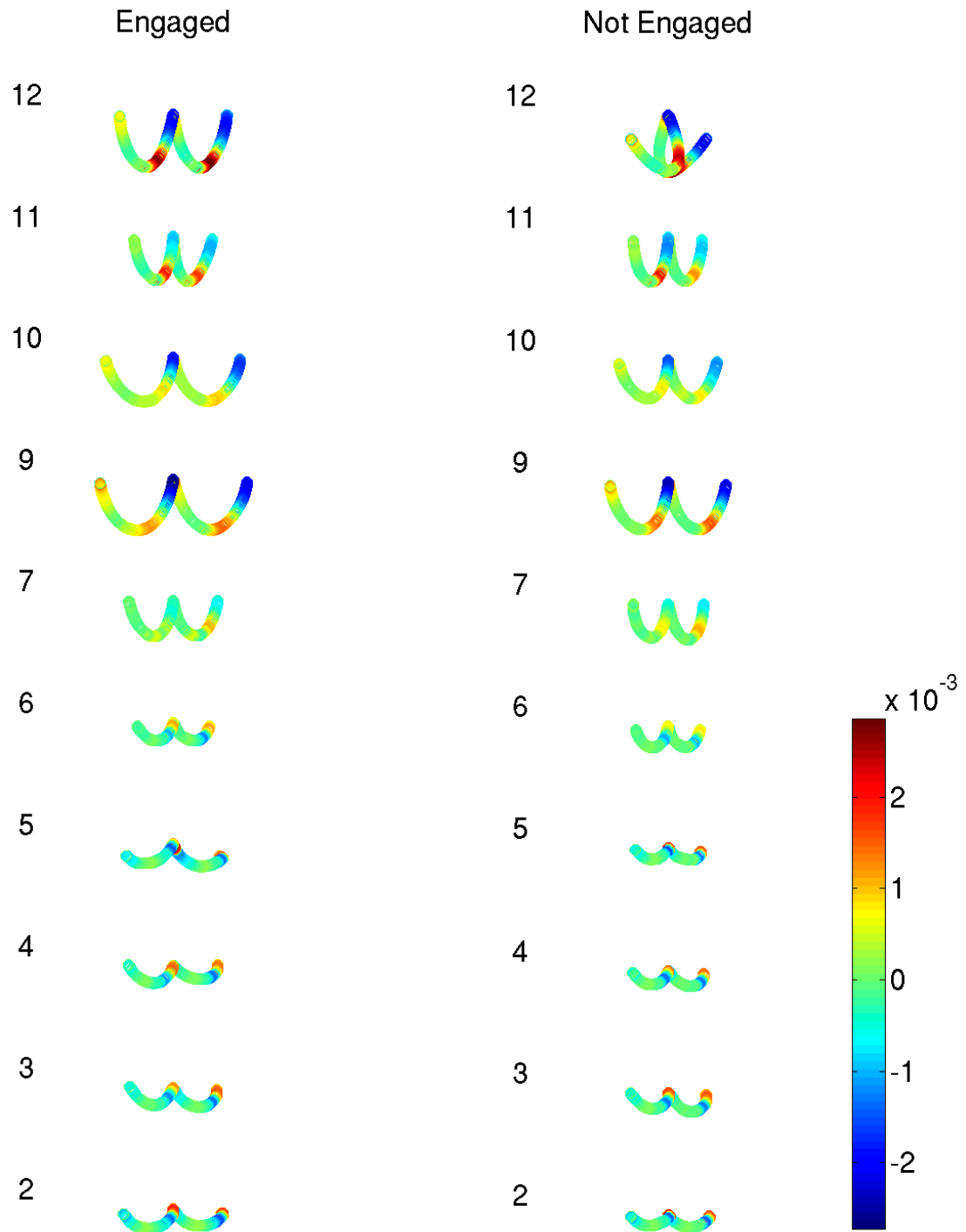


Figure B.14: Movement Trajectories, Subject 415

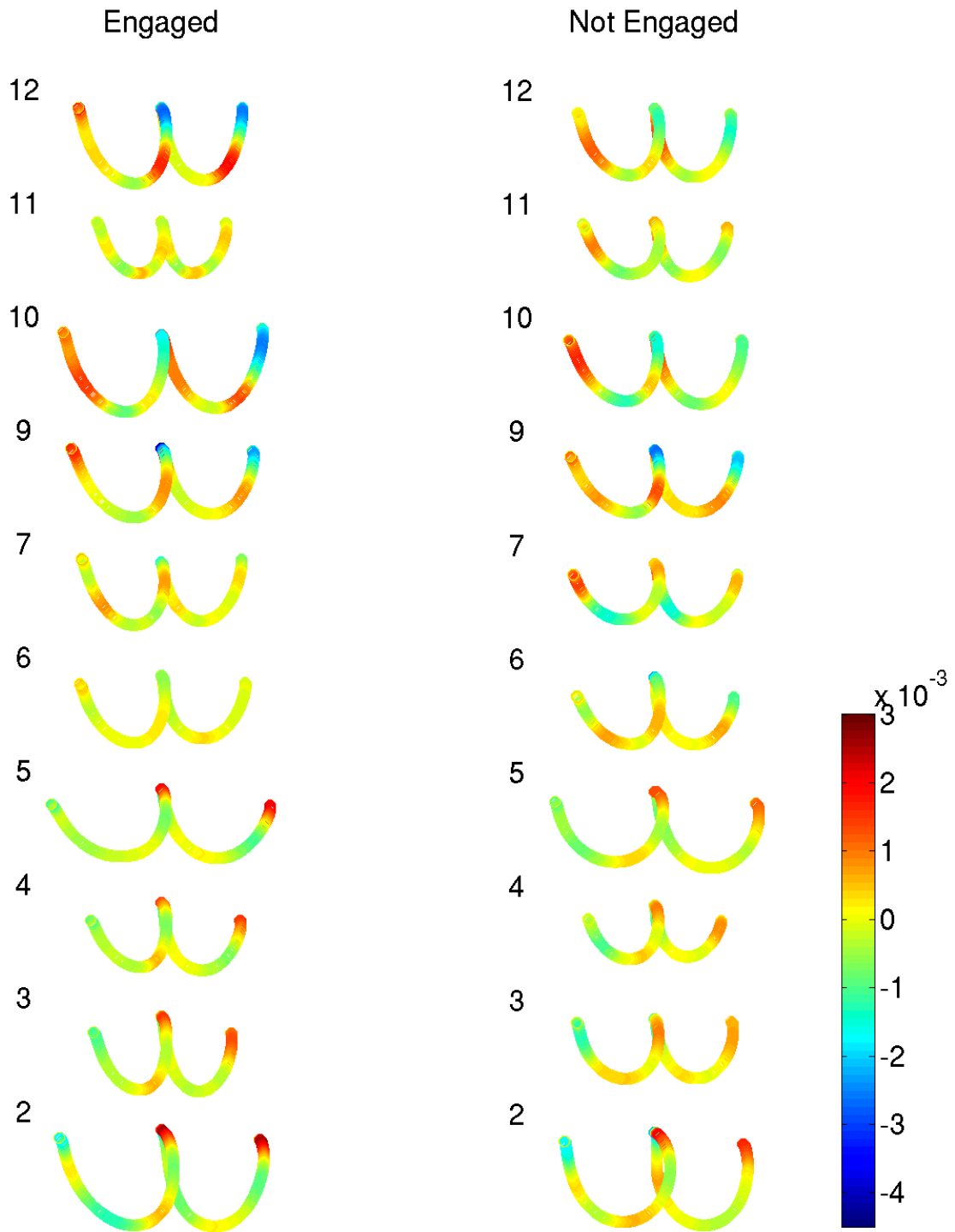


Figure B.15: Movement Trajectories, Subject 416

Bibliography

- [Becker, 2004] Becker, J. O. (2004). *Deep Listeners: Music, emotion, and trancing*, volume 1. Indiana University Press.
- [Blood et al., 1999] Blood, A. J., Zatorre, R. J., Bermudez, P., and Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat Neurosci*, 2(4):382–7.
- [Bonny and Savary, 1973] Bonny, H. L. and Savary, L. M. (1973). *Music and your mind: listening with a new consciousness*. Harper & Row, New York.
- [Bradley and Lang, 1994] Bradley, M. and Lang, P. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59.
- [Brattico et al., 2010] Brattico, E., Jacobsen, T., De Baene, W., Gleran, E., and Ter-vaniemi, M. (2010). Cognitive vs. affective listening modes and judgments of music—an ERP study. *Biological psychology*, 85(3):393–409.
- [Bregman, 1990] Bregman, A. S. (1990). *Auditory scene analysis: the perceptual organization of sound*. MIT Press, Cambridge, Mass.
- [Bregman and Campbell, 1971] Bregman, A. S. and Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *J Exp Psychol*, 89(2):244–9.
- [Brown, 2000] Brown, S. (2000). The “musilanguage” model of music evolution. *The origins of music*, 16:271–300.
- [Buccino et al., 2004] Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C. A., and Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *Journal of Cognitive Neuroscience*, 16(1):114–126.
- [Bush et al., 2000] Bush, Luu, and Posner (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci*, 4(6):215–222.

- [Carr et al., 2003] Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., and Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, 100(9):5497–5502.
- [Chanel et al., 2008] Chanel, G., Rebetez, C., Bétrancourt, M., and Pun, T. (2008). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era*, pages 13–17. ACM.
- [Clayton et al., 2005] Clayton, M., Sager, R., and Will, U. (2005). In time with the music: The concept of entrainment and its significance for ethnomusicology. In *European meetings in ethnomusicology*, volume 11, pages 3–142.
- [Clynes, 1977] Clynes, M. (1977). *Sentics: the touch of emotions*. Anchor Press, Garden City, N.Y., 1st ed edition.
- [Cohen et al., 1997] Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., and Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386(6625):604–8.
- [Coull and Nobre, 1998] Coull, J. T. and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience*, 18(18):7426–7435.
- [Critchley et al., 2000] Critchley, H., Daly, E., Phillips, M., Brammer, M., Bullmore, E., Williams, S., Van Amelsvoort, T., Robertson, D., David, A., and Murphy, D. (2000). Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. *Human brain mapping*, 9(2):93–105.
- [Csikszentmihalyi, 1990] Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. Harper & Row, New York, 1st ed edition.
- [Culham and Kanwisher, 2001] Culham, J. C. and Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Current opinion in neurobiology*, 11(2):157–163.
- [Darwin, 1872] Darwin, C. (1872). *The expression of the emotions in man and animals*. J. Murray, London.
- [Davey, 1999] Davey, H. (1999). *Brush meditation: A Japanese way to mind and body harmony*. Stone Bridge Press, Inc.
- [de Gelder, 2006] de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nat Rev Neurosci*, 7(3):242–249.

- [Decety and Lamm, 2007] Decety, J. and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13(6):580–593.
- [Delorme et al., 2010] Delorme, A., Kothe, C., Vankov, A., Bigdely-Shamlo, N., Oostenveld, R., Zander, T., and Makeig, S. (2010). MATLAB-based tools for BCI research. *Brain-Computer Interfaces*, pages 241–259.
- [Delorme and Makeig, 2004] Delorme, A. and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21.
- [Delorme et al., 2001] Delorme, A., Makeig, S., and Sejnowski, T. (2001). Automatic artifact rejection for EEG data using high-order statistics and independent component analysis. In *Proc. of the 3rd International Workshop on ICA*, volume 457, page 462.
- [Delorme et al., 2011] Delorme, A., Mullen, T., Kothe, C., Acar, Z. A., Bigdely-Shamlo, N., Vankov, A., and Makeig, S. (2011). EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. *Computational intelligence and neuroscience*, 2011:10.
- [Dempster et al., 1989] Dempster, S., Oliveros, P., and Panaiotis (1989). *Deep listening*. New Albion Records, San Francisco, Calif.
- [Deutsch, 1999] Deutsch, D. (1999). Grouping mechanisms in music. *The psychology of music*, 28.
- [Diaz, 2011] Diaz, F. (2011). Mindfulness, attention, and flow during music listening: An empirical investigation. *Psychology of Music*.
- [Eerola and Vuoskoski, 2011] Eerola, T. and Vuoskoski, J. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39(1):18–49.
- [Eerola and Vuoskoski, 2013] Eerola, T. and Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models and stimuli. *Music Perception*.
- [Egermann et al., 2009] Egermann, H., Nagel, F., Altenmüller, E., and Kopiez, R. (2009). Continuous measurement of musically-induced emotion: A web experiment. *International Journal of Internet Science*, 4(1):4–20.
- [Ekman and Oster, 1979] Ekman, P. and Oster, H. (1979). Facial expressions of emotion. *Annual review of psychology*, 30(1):527–554.
- [Fairclough et al., 2009] Fairclough, S., Ewing, K., and Roberts, J. (2009). Measuring task engagement as an input to physiological computing. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–9. IEEE.

- [Farrer et al., 2003] Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., and Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage*, 18(2):324–333.
- [Fink et al., 2009] Fink, A., Graif, B., and Neubauer, A. C. (2009). Brain correlates underlying creative thinking: EEG alpha activity in professional vs. novice dancers. *NeuroImage*, 46(3):854–862.
- [Fishman et al., 2004] Fishman, Y. I., Arezzo, J. C., and Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. *J Acoust Soc Am*, 116(3):1656–70.
- [Fredrickson and Kahneman, 1993] Fredrickson, B. L. and Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *J Pers Soc Psychol*, 65(1):45–55.
- [Frewen et al., 2008] Frewen, P., Lane, R. D., Neufeld, R. W. J., Densmore, M., Stevens, T., and Lanius, R. (2008). Neural correlates of levels of emotional awareness during trauma script-imagery in posttraumatic stress disorder. *Psychosom Med*, 70(1):27–31.
- [Friberg and Sundberg, 1999] Friberg, A. and Sundberg, J. (1999). Does music performance allude to locomotion? a model of final ritardandi derived from measurements of stopping runners. *The Journal of the Acoustical Society of America*, 105:1469.
- [Fritsch and Carlson, 1980] Fritsch, F. N. and Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246.
- [Ganesan et al., 2005] Ganesan, V., Hunter, M. D., and Spence, S. A. (2005). Schneiderian first-rank symptoms and right parietal hyperactivation: A replication using fMRI. *American Journal of Psychiatry*, 162(8):1545–1545.
- [Gomez and Danuser, 2007] Gomez, P. and Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, 7(2):377–87.
- [Goodkind et al., 2012] Goodkind, M. S., Sollberger, M., Gyurak, A., Rosen, H. J., Rankin, K. P., Miller, B., and Levenson, R. (2012). Tracking emotional valence: the role of the orbitofrontal cortex. *Hum Brain Mapp*, 33(4):753–62.
- [Gottman and Levenson, 1985] Gottman, J. M. and Levenson, R. W. (1985). A valid procedure for obtaining self-report of affect in marital interaction. *J Consult Clin Psychol*, 53(2):151–60.
- [Grindea and Menuhin, 1995] Grindea, C. and Menuhin, Y. (1995). *Tensions in the Performance of Music: a Symposium*. Kahn & Averill Publishers.

- [Halpern et al., 2008] Halpern, A., Martin, J., and Reed, T. (2008). An ERP study of major-minor classification in melodies. *Music Perception*, 25(3):181–191.
- [Hatfield et al., 1993] Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, 2(3):96–99.
- [Haueisen and Knösche, 2001] Haueisen, J. and Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of cognitive neuroscience*, 13(6):786–792.
- [Hirshkowitz et al., 1978] Hirshkowitz, M., Earle, J., and Paley, B. (1978). EEG alpha asymmetry in musicians and non-musicians: A study of hemispheric specialization. *Neuropsychologia*, 16(1):125–128.
- [Holmes et al., 1996] Holmes, A. P., Blair, R., Watson, G., and Ford, I. (1996). Non-parametric analysis of statistic images from functional mapping experiments. *Journal of Cerebral Blood Flow & Metabolism*, 16(1):7–22.
- [Huron, 2006a] Huron, D. B. (2006a). *Sweet anticipation: Music and the psychology of expectation*. MIT press.
- [Huron, 2006b] Huron, D. B. (2006b). *Sweet anticipation: music and the psychology of expectation*. MIT Press, Cambridge, Mass.
- [Hutcherson et al., 2005] Hutcherson, C. A., Goldin, P. R., Ochsner, K. N., Gabrieli, J. D., Barrett, L. F., and Gross, J. J. (2005). Attention and emotion: does rating emotion alter neural responses to amusing and sad films? *Neuroimage*, 27(3):656–668.
- [Iyer, 2002] Iyer, V. (2002). Embodied mind, situated cognition, and expressive micro-timing in African-American music. *Music Perception*, 19(3):387–414.
- [John et al., 1991] John, O., Donahue, E., and Kentle, R. (1991). The Big Five Inventory-versions 4a and 54. *Berkeley: University of California, Berkeley, Institute of Personality and Social Research*.
- [John et al., 2008] John, O., Naumann, L., and Soto, C. (2008). Paradigm shift to the integrative Big Five trait taxonomy. *Handbook of personality: Theory and research*, pages 114–158.
- [Jolliffe, 2005] Jolliffe, I. (2005). *Principal component analysis*. Wiley Online Library.
- [Jones and Nisbett, 1971] Jones, E. E. and Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior*. General Learning Press New York.

- [Jung-Beeman et al., 2004] Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J., and Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLoS biology*, 2(4):e97.
- [Juslin et al., 2008] Juslin, P., Liljeström, S., Västfjäll, D., Barradas, G., and Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, 8(5):668.
- [Juslin and Vastfjall, 2008] Juslin, P. N. and Vastfjall, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and brain sciences*, 31(5):559.
- [Kahaner et al., 1989] Kahaner, D., Moler, C., and Nash, S. (1989). *Numerical methods and software*. Prentice-Hall Englewood Cliffs, NJ.
- [Khalifa et al., 2008] Khalifa, S., Roy, M., Rainville, P., Dalla Bella, S., and Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music? *International Journal of Psychophysiology*, 68(1):17–26.
- [Khalifa et al., 2005] Khalifa, S., Schon, D., Anton, J., and Liégeois-Chauvel, C. (2005). Brain regions involved in the recognition of happiness and sadness in music. *Neuroreport*, 16(18):1981–1984.
- [Khosrowabadi et al., 2009] Khosrowabadi, R., Wahab, A., Ang, K., and Baniasad, M. (2009). Affective computation on EEG correlates of emotion from musical and vocal stimuli. In *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pages 1590–1594. IEEE.
- [Kiebel et al., 2005] Kiebel, S. J., Tallon-Baudry, C., and Friston, K. J. (2005). Parametric analysis of oscillatory activity as measured with EEG/MEG. *Human brain mapping*, 26(3):170–177.
- [Koelsch et al., 2004] Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nat Neurosci*, 7(3):302–7.
- [Koelsch and Siebel, 2005] Koelsch, S. and Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends Cogn Sci*, 9(12):578–84.
- [Kohler et al., 2002] Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848.
- [Konecni, 2008] Konecni, V. J. (2008). Does music induce emotion? a theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2):115–129.

- [Konečni, 2010] Konečni, V. J. (2010). The influence of affect on music choice. *Handbook of music and emotion. Theory, research, applications*, pages 697–723.
- [Krumhansl, 2002] Krumhansl, C. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, 11(2):45–50.
- [Krumhansl, 1996] Krumhansl, C. L. (1996). A perceptual analysis of Mozart’s Piano sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, pages 401–432.
- [Krumhansl, 1997] Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Can J Exp Psychol*, 51(4):336–53.
- [Lakey et al., 2011] Lakey, C. E., Berry, D. R., and Sellers, E. W. (2011). Manipulating attention via mindfulness induction improves P300-based brain-computer interface performance. *J Neural Eng*, 8(2):025019.
- [Lane et al., 1998] Lane, R., Reiman, E., Axelrod, B., Yun, L., Holmes, A., and Schwartz, G. (1998). Neural correlates of levels of emotional awareness: evidence of an interaction between emotion and attention in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, 10(4):525–535.
- [Laukka et al., 2005] Laukka, P., Juslin, P., and Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, 19(5):633–653.
- [Lehne et al., 2009] Lehne, M., Ihme, K., Brouwer, A., Van Erp, J., and Zander, T. (2009). Error-related EEG patterns during tactile human-machine interaction. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–9. IEEE.
- [Lerdahl and Jackendoff, 1983] Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press, Cambridge, Mass.
- [Leslie and Mullen, 2010] Leslie, G. and Mullen, T. (2010). Moodmixer: EEG-based collaborative sonification. *Proceedings of the Conference on New Interfaces for Musical Expression*.
- [Leube et al., 2003] Leube, D. T., Knoblich, G., Erb, M., and Kircher, T. T. (2003). Observing one’s hand become anarchic: An fMRI study of action identification. *Consciousness and cognition*, 12(4):597–608.
- [Levenson and Gottman, 1983] Levenson, R. W. and Gottman, J. M. (1983). Marital interaction: physiological linkage and affective exchange. *J Pers Soc Psychol*, 45(3):587–97.
- [Levitin et al., 2007] Levitin, D. J., Nuzzo, R. L., Vines, B. W., and Ramsay, J. (2007). Introduction to functional data analysis. *Canadian Psychology*, 48(3):135.

- [Lezak, 2004] Lezak, M. D. (2004). *Neuropsychological Assessment 4e*. Oxford University Press, USA.
- [Lin et al., 2010a] Lin, Y., Wang, C., Jung, T., Wu, T., Jeng, S., Duann, J., and Chen, J. (2010a). EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, 57(7):1798–1806.
- [Lin et al., 2010b] Lin, Y.-P., Duann, J.-R., Chen, J.-H., and Jung, T.-P. (2010b). Electroencephalographic dynamics of musical emotion perception revealed by independent spectral components. *Neuroreport*, 21(6):410–5.
- [Lombardo et al., 2011] Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., and Baron-Cohen, S. (2011). Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *Neuroimage*, 56(3):1832–1838.
- [Luck et al., 2010] Luck, G., Toiviainen, P., and Thompson, M. (2010). Perception of expression in conductors’ gestures: A continuous response study. *Music Perception*, 28(1):47–57.
- [Lutz et al., 2004] Lutz, A., Greischar, L. L., Rawlings, N. B., Ricard, M., and Davidson, R. J. (2004). Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. *Proc Natl Acad Sci U S A*, 101(46):16369–73.
- [Macmillan and Creelman, 2004] Macmillan, N. A. and Creelman, C. D. (2004). *Detection theory: A user’s guide*. Lawrence Erlbaum.
- [MacQueen et al., 1967] MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1.
- [Madsen and Fredrickson, 1993] Madsen, C. K. and Fredrickson, W. E. (1993). The experience of musical tension: A replication of Nielsen’s research using the continuous response digital interface. *Journal of Music Therapy*, 30:46–46.
- [Makeig, 1993] Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and clinical neurophysiology*, 86(4):283–293.
- [Makeig et al., 1996] Makeig, S., Bell, A., Jung, T., Sejnowski, T., et al. (1996). Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, pages 145–151.
- [Makeig et al., 2009] Makeig, S., Gramann, K., Jung, T., Sejnowski, T., and Poizner, H. (2009). Linking brain, mind and behavior. *International Journal of Psychophysiology*, 73(2):95–100.

- [Makeig et al., 2012] Makeig, S., Kothe, C., Mullen, T., Bigdely-Shamlo, N., Zhang, Z., and Kreutz-Delgado, K. (2012). Evolving signal processing for brain-computer interfaces. *Proceedings of the IEEE*, 100(13):1567–1584.
- [Makeig et al., 2011] Makeig, S., Onton, J., et al. (2011). ERP features and EEG dynamics: an ICA perspective. *Oxford Handbook of Event-Related Potential Components*. New York, NY: Oxford.
- [Makeig et al., 2002] Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., and Sejnowski, T. (2002). Dynamic brain sources of visual evoked responses. *Science*, 295(5555):690–694.
- [Mallat, 1999] Mallat, S. (1999). *A wavelet tour of signal processing*. Academic press.
- [Mauss et al., 2005] Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., and Gross, J. J. (2005). The tie that binds? coherence among emotion experience, behavior, and physiology. *Emotion*, 5(2):175–90.
- [Meyer, 1956] Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.
- [Miller, 1968] Miller, R. E. (1968). *The Heart Is a Lonely Hunter*. Warner Brothers.
- [Miyakoshi et al., 2010] Miyakoshi, M., Kanayama, N., Idaka, T., Ohira, H., et al. (2010). EEG evidence of face-specific visual self-representation. *NeuroImage*, 50(4):1666.
- [Mohanty et al., 2007] Mohanty, A., Engels, A. S., Herrington, J. D., Heller, W., Ringo Ho, M.-H., Banich, M. T., Webb, A. G., Warren, S. L., and Miller, G. A. (2007). Differential engagement of anterior cingulate cortex subdivisions for cognitive and emotional function. *Psychophysiology*, 44(3):343–351.
- [Molnar-Szakacs and Overy, 2006] Molnar-Szakacs, I. and Overy, K. (2006). Music and mirror neurons: from motion to ‘e’motion. *Soc Cogn Affect Neurosci*, 1(3):235–241.
- [Nagel et al., 2007] Nagel, F., Kopiez, R., Grewe, O., and Altenmüller, E. (2007). EMu-Joy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods*, 39(2):283–290.
- [Nichols and Holmes, 2001] Nichols, T. E. and Holmes, A. P. (2001). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1):1–25.
- [Nielsen, 1983] Nielsen, F. (1983). *Oplevelse af musiksk spænding (experience of musical tension)*. PhD thesis, Copenhagen, Denmark.

- [Nusseck and Wanderley, 2009] Nusseck, M. and Wanderley, M. (2009). Music and motion-how music-related ancillary body movements contribute to the experience of music. *Music Perception*, 26(4):335–353.
- [Ochsner and Gross, 2005] Ochsner, K. N. and Gross, J. J. (2005). The cognitive control of emotion. *Trends Cogn Sci*, 9(5):242–249.
- [Onton and Makeig, 2009] Onton, J. and Makeig, S. (2009). High-frequency broadband modulations of electroencephalographic spectra. *Front Hum Neurosci*, 3:61.
- [Patel, 2003] Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, 6(7):674–681.
- [Peretz and Zatorre, 2005] Peretz, I. and Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.*, 56:89–114.
- [Preston and Newport, 2008] Preston, C. and Newport, R. (2008). Misattribution of movement agency following right parietal TMS. *Social cognitive and affective neuroscience*, 3(1):26–32.
- [Preston and De Waal, 2002] Preston, S. D. and De Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and brain sciences*, 25(01):1–20.
- [Raimondo et al., 2012] Raimondo, F., Kamienkowski, J. E., Sigman, M., and Slezak, D. F. (2012). CUDAICA: GPU optimization of infomax-ICA EEG analysis. *Computational Intelligence and Neuroscience*, 2012:2.
- [Ramsay and Silverman, 1997] Ramsay, J. and Silverman, B. (1997). *Functional Data Analysis*.
- [Rentfrow and Gosling, 2003] Rentfrow, P. and 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(6):1236.
- [Rizzolatti and Craighero, 2004] Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, 27:169–192.
- [Rohenkohl and Nobre, 2011] Rohenkohl, G. and Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *The Journal of Neuroscience*, 31(40):14076–14084.
- [Romei et al., 2010] Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *The Journal of neuroscience*, 30(25):8692–8697.
- [Rosenberg, 1994] Rosenberg, H. (1994). *The tradition of the new*. Da Capo Press.

- [Ruby and Decety, 2001] Ruby, P. and Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature neuroscience*, 4(5):546–550.
- [Ruef and Levenson, 2007] Ruef, A. and Levenson, R. (2007). Continuous measurement of emotion. *Handbook of emotion elicitation and assessment*, pages 286–297.
- [Sammler et al., 2007] Sammler, D., Grigutsch, M., Fritz, T., and Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, 44(2):293–304.
- [Saxe and Kanwisher, 2005] Saxe, R. and Kanwisher, N. (2005). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *Social Neuroscience: Key Readings*, Psychology Press, New York, pages 171–182.
- [Schmidt and Trainor, 2001] Schmidt, L. and Trainor, L. (2001). Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognition & Emotion*, 15(4):487–500.
- [Schubert, 1999] Schubert, E. (1999). *Measurement and time series analysis of emotion in music*. PhD thesis, University of New South Wales.
- [Steinbeis et al., 2006] Steinbeis, N., Koelsch, S., and Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18(8):1380–1393.
- [Sussman et al., 1999] Sussman, E., Ritter, W., and Vaughan, Jr, H. G. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*, 36(1):22–34.
- [Taylor et al., 2003] Taylor, S. F., Phan, K. L., Decker, L. R., and Liberzon, I. (2003). Subjective rating of emotionally salient stimuli modulates neural activity. *Neuroimage*, 18(3):650–659.
- [Tellegen and Atkinson, 1974] Tellegen, A. and Atkinson, G. (1974). Openness to absorbing and self-altering experiences (“absorption”), a trait related to hypnotic susceptibility. *Journal of abnormal psychology*, 83(3):268.
- [Teolis, 1998] Teolis, A. (1998). *Computational signal processing with wavelets*. Birkhauser.
- [Tervaniemi and Huotilainen, 2003] Tervaniemi, M. and Huotilainen, M. (2003). The promises of change-related brain potentials in cognitive neuroscience of music. *Ann N Y Acad Sci*, 999:29–39.

- [Timmers et al., 2006] Timmers, R., Marolt, M., Camurri, A., and Volpe, G. (2006). Listeners' emotional engagement with performances of a Scriabin étude: an explorative case study. *Psychology of Music*, 34(4):481–510.
- [Uddin et al., 2006] Uddin, L. Q., Molnar-Szakacs, I., Zaidel, E., and Iacoboni, M. (2006). rTMS to the right inferior parietal lobule disrupts self-other discrimination. *Social Cognitive and Affective Neuroscience*, 1(1):65–71.
- [Vines et al., 2011] Vines, B., Krumhansl, C., Wanderley, M., Dalca, I., and Levitin, D. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, 118(2):157.
- [Vines et al., 2006] Vines, B. W., Krumhansl, C. L., Wanderley, M. M., and Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1):80–113.
- [Vollm et al., 2006] Vollm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J. F., and Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, 29(1):90–98.
- [Wanderley et al., 2005] Wanderley, M., Vines, B., Middleton, N., McKay, C., and Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1):97–113.
- [Watson et al., 1988] Watson, D., Clark, L., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6):1063.
- [Witmer and Singer, 1998] Witmer, B. and Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240.
- [Wolpaw et al., 2002] Wolpaw, J., Birbaumer, N., McFarland, D., Pfurtscheller, G., Vaughan, T., et al. (2002). Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791.
- [Yomogida et al., 2010] Yomogida, Y., Sugiura, M., Sassa, Y., Wakusawa, K., Sekiguchi, A., Fukushima, A., Takeuchi, H., Horie, K., Sato, S., and Kawashima, R. (2010). The neural basis of agency: an fMRI study. *Neuroimage*, 50(1):198–207.
- [Zander et al., 2010] Zander, T., Gaertner, M., Kothe, C., and Vilimek, R. (2010). Combining eye gaze input with a brain-computer interface for touchless human-computer interaction. *Intl. Journal of Human-Computer Interaction*, 27(1):38–51.
- [Zatorre et al., 2002] Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in cognitive sciences*, 6(1):37–46.

- [Zatorre et al., 2007a] Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007a). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7):547–558.
- [Zatorre et al., 2007b] Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007b). When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci*, 8(7):547–58.
- [Zentner and Eerola, 2010] Zentner, M. and Eerola, T. (2010). Self-report measures and models. *Handbook of music and emotion: Theory, research, applications*, pages 187–221.
- [Zentner et al., 2008] Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, 8(4):494–521.