Designing Sonic Interactions for Implicit Physiological Computing

Sebastián Mealla Cincuegrani

TESI DOCTORAL UPF / 2015

Dirigida per:

Dr. Sergi Jordà

Dr. Aleksander Väljamäe (HCI group, School of Digital Technologies, Tallinn University)

Departament de Tecnologies de la Informació i les Comunicacions

Universitat Pompeu Fabra Barcelona

To Ayşe Naz

Acknowledgements

This dissertation is written in first person of plural for a good reason. During my PhD program I had the chance to work side-by-side with very talented and inspiring people, starting with my supervisors Sergi (who backed me up when I was in Argentina, looking for a fellowship program) and Aleksander (with whom I have been working since my master degree). I would like to thank them for their dedication, advice and (especially) their patience all these years, and for making my PhD an enriching experience, both at scientific and human levels.

I am also very grateful to my colleagues at the "328", Carles F. Julià, Daniel Gallardo, Mathieu Bosi, Gianmaria Troiano, and Panos Papiotis, with whom I learnt a lot, conceived great projects, and shared memorable moments. Thanks for your support! Also, my gratitude to the master students I have worked with: Marco Zaccaria, Aditya Nandwana, Andrew Stella, Diana Mundó, Aluizio Oliveira, Xavier Marimón, and Carles Tardío Pi.

I would like to thank the MTG family for such a great working environment, in particular to Alba Rosado, the "Lunchtime Conversations" team, and the Music Hack Day team: Alba, Mohamed Sordo, Álvaro Sarasúa, Frederic Font and Panos Papiotis.

I would have never been able to carry on this PhD program without the support I received from the Centro Cultural España Córdoba (CCEC) and Universidad Blas Pascal, in Argentina. In this regard, I am extremely grateful to Pancho Marchiaro, Daniel Salzano, and Mercedes Outumuro for supporting my MAEC-AECID application back in 2008, when we were working together. My gratitude also to the CCEC family for supporting me in my "European endeavor" even when it implied to be apart for years.

A special thank goes to my lovely wife Ayse Naz, that supported me in every of my "phd madness" states, devoted long hours to proof read this document, and cheered me up for long hours.

Finally, I would like to thank my family and friends spread all around the world, for their continuous support all these years. In particular to my mother that, although she never left Argentina, is the most tireless traveler I have ever met.

This work was supported by the MAEC-AECID fellowship program, TEC2010-11599-E, and the European Union's Horizon 2020 research and innovation programme through the RAPID-MIX Project (nr. 644862).

Abstract

The field of Human-Computer Interaction (HCI) has been historically devoted to understand the interplay between people and computers. However, for the last three decades, it has been mainly based on overt and explicit control by means of peripheral devices such as the keyboard and the mouse. As devices and systems are becoming increasingly complex and powerful, this traditional approach to interface design is often lagging behind, constituting a bottleneck for seamless HCI.

In order to achieve more natural interactions with computer systems, HCI has to go beyond explicit control and incorporate the implicit subtleties of human-human interaction. This could be achieved by means of Physiological Computing, which monitors naturalistic changes in the user psychophysiological states (affective, perceptive or cognitive) for adapting system responses without explicit control. At the output level, Sonic Interaction Design (SID) appears as an excellent medium for representing implicit physiological states, as acoustic data can be processed faster than visual presentation, can be easily localized in space, it has a good temporal resolution, and account for displaying multiple data streams while releasing the visual sense.

Therefore, in this dissertation we aim to conceptualize, prototype and evaluate sonic interaction designs for implicit Physiological Computing in the context of HCI. For achieving this goal, we leverage on physiological sensing techniques, namely EEG and ECG, to estimate user's implicit states in real time, and apply diverse SID methodologies to adapt system responses according to these statuses. We incrementally develop different implicit sonic interactions (from direct audification to complex musical mappings) and evaluate them in HCI scenarios (from neurofeedback to music performance), assessing their perceptualization quality, the role of mapping complexity, and their meaningfulness in the musical domain.

Sinopsis

El campo de la interacción persona-ordenador (HCI, por si siglas en inglés) se ha dedicado históricamente a comprender la compleja relación entre usuarios y sistemas computacionales. Sin embargo, en las últimas tres décadas, esta disciplina se ha concentrado primordialmente en mecanismos de control explícitos, mediante dispositivos periféricos como el teclado o el ratón. A medida que los sistemas informáticos y sus interfaces se vuelven más complejos y potentes, esta aproximación tradicional al diseño de interfaces representa una limitación para alcanzar mecanismos de interacción más naturales e intuitivos para el usuario final.

Para superar dicha barrera, HCI debe ir más allá del control explícito e incorporar las sutilezas de lo implícito, tan características en la comunicación humana. Tal objetivo puede alcanzarse a través de las Fisiología Computacional, dedicada a monitorear sistemáticamente los cambios psicofisiológicos (e.g. afectivos, perceptivos o cognitivos) del individuo para aplicarlos al control de sistemas computacionales. Asimismo, el diseño de interacción sonora ofrece ventajas significativas para representar cambios fisiológicos implícitos en la actividad del usuario, ya que la información acústica puede ser procesada más rápidamente que las presentaciones visuales, es fácilmente localizable en el espacio, posee una resolución temporal alta, y permite representar múltiples flujos de datos al tiempo que libera la visión.

Por lo tanto, esta tesis doctoral tiene por objetivo conceptualizar, prototipar y evaluar interacciones sonoras basadas en fisiología computacional implícita, en contextos HCI. Para alcanzar dicho objetivo, se aplican métodos de medición fisiológica, específicamente EEG y ECG, para estimar estados psicofisiológicos de los usuarios en tiempo real, los cuales se utilizan para adaptar implícitamente la respuesta de interfaces sonoras interactivas. De esta manera, se presentan diferentes estrategias de interacción sonora implícita (desde audificación directa, hasta mapeos musicales complejos) las cuales se aplican en escenarios de HCI (de neurofeedback hasta performance musical) para evaluar sus cualidades de perceptualización, el rol de la complejidad de los mappings entre información fisiológica y sonido, y sus implicancias en contextos expresivos, como la creación musical.

Contents

1. Introduction **1**
1.1. Motivation . . . 1
1.2. Background Knowledge . . . 6
1.2.1. Physiological Computing . . . 6
1.2.2. Sonic Interaction Design . . . 11
1.3. Objectives and Outline of the Thesis . . . 15

2. Literature Review **19**
2.1. Introduction . . . 19
2.2. Designing Implicit Interactions . . . 19
2.2.1. Buxton's taxonomy for integrating periphery and context . . . 20
2.2.2. Calm technology . . . 23
2.2.3. Proactive computing . . . 23
2.2.4. Implicit HCI through context . . . 24
2.2.5. A domain-independent approach to implicit interaction . . . 26
2.3. Implicit interaction through physiological computing: expanding context with physicality . . . 28
2.3.1. What is a biosignal? . . . 29
2.3.2. Physiological computing systems . . . 39
2.4. Sonic interaction design . . . 43
2.4.1. Sonification . . . 44
2.4.2. New interfaces for musical expression (NIME) . . . 47
2.5. Sonic interaction design applied to physiological computing . . . 50
2.5.1. The pioneers . . . 50
2.5.2. Beyond biofeedback and towards implicit/explicit control . . . 51
2.6. State of the art of SID based on EEG . . . 56
2.6.1. Application areas . . . 58
2.6.2. Search and selection criteria . . . 60
2.6.3. Relating EEG signal processing to SID techniques . . . 61

2.6.4. EEG processing dimensions 65
2.7. A summary of learnings for this dissertation 71

3. The b-Reactable: Prototyping Sonic Interactions for Implicit Physiological Computing 75
3.1. Introduction and motivation 75
3.2. Materials 77
3.2.1. System Design 77
3.2.2. The Reactable and the physiopucks 77
3.2.3. Physiological Signal acquisition 80
3.2.4. Sound engine 80
3.2.5. EEG and ECG signal processing and sonic display 82
3.3. Methods, experimental setup and procedure 84
3.3.1. The Physio group 85
3.3.2. The Sham group 85
3.3.3. The Control group 85
3.3.4. Experimental procedure 86
3.3.5. Measures 88
3.4. Results 89
3.4.1. Comparing the four b-Reactable participant types to the Control group 90
3.4.2. Comparing Physio and Sham groups 90
3.4.3. Correlations between User-Emitter dyad 92
3.4.4. Correlations between motivational characteristics and valence-arousal ratings 94
3.5. Discussion 97
3.6. Conclusions & next steps 102

4. Perceptualization and Mapping Complexity 105
4.1. Introduction and motivation 106
4.2. Experiment 1: assessing perceptualization of sonic designs 107
4.2.1. Material & methods 108
4.2.2. Results and discussion 117
4.3. Experiment 2: mapping complexity and personalization 120
4.3.1. Material & methods 121
4.3.2. Results and Discussion 135
4.4. General discussion 136

4.5. Conclusion & next steps 140

5. A Methodological Framework for Designing and Evaluating NIME 143

5.1. Introduction and motivation 143
5.2. Designing and evaluating NIME 144
5.2.1. Learning NIME design 145
5.2.2. NIME design and evaluation frameworks 146
5.3. Framework and deployment 148
5.3.1. Context 148
5.3.2. Students profile & structure of the course 150
5.3.3. Evaluation tools 152
5.3.4. Results 154
5.4. Discussion 155
5.4.1. Analysis of the Results 155
5.4.2. Analysis of the Initial Objectives 156
5.4.3. Guidelines for future work 157
5.5. Conclusion & next steps 157

6. Enhancing NIME with Implicit Physiological Computing 159

6.1. Introduction and motivation 160
6.2. Context 161
6.2.1. Feedback design for physiology-based implicit interaction 161
6.2.2. Music, EEG and emotion estimation 163
6.2.3. Expressiveness in NIME and HCI 163
6.3. Materials and methods 165
6.3.1. System design 165
6.3.2. Physiological signal acquisition and processing 166
6.3.3. Valence and arousal estimation 167
6.3.4. Mapping design 168
6.3.5. Integration into the Reactable framework 169
6.3.6. Global and Local implicit interaction 170
6.3.7. Case study with musicians 172
6.4. Results 175
6.4.1. Validation of affective estimations 175
6.4.2. Differences between setups 175
6.4.3. Differences between tasks 175
6.4.4. Interaction between setup and task factors 176

6.4.5. Interaction between setup, task and music experience 176
6.4.6. Analysis of System and Performance aspects of the DMI 176
6.4.7. Insights from user experience interviews 180
6.5. Discussion 182
6.5.1. Performance of affective estimations and physiology-based implicit interaction setups 182
6.5.2. Meaningful integration of implicit PhyComp in music performance 183
6.5.3. The role of music knowledge in implicit interaction preference . . . 183
6.5.4. Multimodality and personalization 184
6.5.5. Limitations and future work 185
6.6. Conclusion & next steps 186

7. Bonus Track. Implicit Physiological Computing Beyond Sonic Interaction Design 189
7.1. Introduction 189
7.2. Motivation 190
7.3. Background 191
7.4. The NeuroKnitting prototype 192
7.4.1. EEG signal acquisition and toolbox for processing 192
7.4.2. Pattern generator 195
7.4.3. Deployment 1: patterns through Bach's Goldberg variations 197
7.4.4. Deployment 2: Maker Faire Rome 199
7.4.5. Discussion and conclusion 201

8. Conclusion 205
8.1. Contribution 1: understanding the context 205
8.1.1. Limitations and possible expansions 206
8.2. Contribution 2: enhancing user motivation through implicit PhyComp . . 207
8.2.1. Limitations and possible expansions 207
8.3. Contribution 3: perceptualization and mapping complexity 208
8.3.1. Limitations and possible expansions 209
8.4. Contribution 4: NIME design and evaluation 210
8.4.1. Limitation and possible expansions 210
8.5. Contribution 5: Implicit PhyComp supporting musical expression 211
8.5.1. Limitation and possible expansions 211
8.6. Contribution 6: taking implicit PhyComp beyond SID 212

8.7. Non-academic contributions 213
8.7.1. Workshops and Hackathons 213
8.8. Closing remarks 214

A. List of publications 237

List of Figures

1.1. Classification of Physiological Computing systems. 9
1.2. Sonification Diagram. 13
1.3. Graphical representation of the thesis structure. 15

2.1. Buxton's model of periphery and context. 21
2.2. Tennenhouse's model of proactive computing. 24
2.3. Ju & Leifer's implicit interaction framework. 27
2.4. The stethoscope and the registration of body sounds. 30
2.5. Model of biosignal generation, propagation, coupling, and registration. . . 31
2.6. Components of the ECG signal. 34
2.7. The EMG signal. 35
2.8. Kaniusas' biosignal classification method. 37
2.9. Diagnostic and therapeutic application of biosignals Kaniusas [2012] . . . 39
2.10. Sonification techniques: (a) audification and (b) parameter mapping [Hermann, 2008] 46
2.11. Jordà's hybrid micro + macrocontrol. 49
2.12. Alvin Lucier' Music for Solo Performer. 51
2.13. Rosenboom's *portable gold and philosophers' stones (Music from Brains in Four).* 52
2.14. Atau Tanaka performing at the Sonar Festival, Barcelona, Spain, June 1997. 53
2.15. Marco Donnarumma performing *The Moving Forest* at Transmediale.08, Berlin (2008). 54
2.16. Miranda's BCMI-Piano. System architecture [Miranda, 2006]. 55
2.17. Application areas that use sonic display of EEG data. 59
2.18. Relationship between *Data Processing* and *SID techniques.* 62
2.19. Articles corresponding to the three most popular processing techniques (a) unprocessed EEG, (b) filtered bands, and (c) spectral estimators. . . 63
2.20. Classification of articles according to time-frequency dimensions. 67

2.21. Classification of articles according to signal levels. 68
2.22. Classification of articles according to EEG channels. 70
2.23. Classification of articles according to EEG channel location. 70
2.24. Classification of articles according to the montage system. 71
2.25. Classification of articles according to high level processing features. 72

3.1. The *b-Reactable* system, first prototype. 78
3.2. A performer playing with the Reactable. 79
3.3. The Reactable architecture and components [Jordà et al., 2007]. 79
3.4. The *Physiopucks*. 81
3.5. The *b-Reactable*. Signal processing for EEG audification and ECG tempo control. 83
3.6. Confidence, Satisfaction, and Gestural Communication ratings across five participant types. 91
3.7. Significant interactions between "group" and "role-in-dyad" factors in the Physio and Sham groups. 93
3.8. Correlations between Social Affinity and task Difficulty. 96

4.1. Graphical overview of the system architecture for experiment 1. 109
4.2. Flowchart illustrating EEG signal acquisition and processing. 111
4.3. Ratings of *sound tranquility*, *reported relaxation*, and *perceived congruency* for the three sonic designs. 118
4.4. Components and data streams of the EEG sonic display system. 121
4.5. The *EEG-TOOLbox* integrated into the Matlab/Simulink library. 122
4.6. The EEG-TOOLbox. Emotiv2Simulink block. 123
4.7. The EEG-TOOLbox. Band-pass filter blocks. 124
4.8. The *EEG-TOOLbox*. Processing for envelope estimation. 125
4.9. Sonic Engine. Pd patch of the Pedal module. 127
4.10. Sonic Engine. Pd patch of the Melody module. 127
4.11. Sonic Engine. Pd patch of the Rewarding module. 128
4.12. Sonic Engine. Pd patch of the Wind module. 129
4.13. Sonic Engine. Pd patch of the Rain module 129
4.14. Sonic Engine. Graphic User Interface 130
4.15. Processing stages of EEG features and corresponding sound mappings for experiment 2. 131
4.16. Photo of the technical setup for experiment 2. 133

4.17. Experiment 2. Difference between the first and the final 3-min period of the 15-minute session. 137
4.18. Experiment 2. Difference between the first and the final 3-minute period of the 15-minute session. 138

5.1. Performances by working groups using the DMIs developed during the master course. 153

6.1. The primary music-emotion structural rules graphed on the Two Dimensional Emotion Space [Livingstone et al., 2007] 164
6.2. The updated *b-Reactable*: system architecture diagram. 166
6.3. The *b-Reactable* sound processors. 168
6.4. Mapping for indirect physiological control. 170
6.5. The second version of the *Physiopucks*. 171
6.6. Players performing with the multimodal musical system 173
6.7. Second experiment with the *b-Reactable.* Interactions between setup and task factors. 177
6.8. Second experiment with the *b-Reactable.* Interaction between setup, task, and group for subjective ratings. 178
6.9. Second experiment with the *b-Reactable.* Interaction between setup, task, and group for EEG-based valence. 179

7.1. The *NeuroKnitting* system architecture. 193
7.2. Processing stages of EEG features for the *NeuroKnitting* prototype. . . . 195
7.3. The *Knitic* PCB. 197
7.4. Scarves knitted during deployment 1. 198
7.5. Picture of one of the scarves created with *NeuroKnitting* during the second deployment at MFR. 200

List of Tables

2.1. EEG frequency bands. 36

3.1. Pucks and *Physiopucks* available in the first *b-Reactable* prototype. 86
3.2. Description of tasks involved in the first experiment with the *b-Reactable*. 87
3.3. Pearson correlation coefficients for user-emitter dyad ratings. 94
3.4. Correlations between measures of motivational and emotional experience. 95

4.1. Procedure for Experiment 1. 116
4.2. Configuration of the sonic engine for each experimental group. 132

5.1. Structure and contents of the course [Jordà and Mealla, 2014] 151
5.2. Pearson correlations for *System* and *Performance* variables. 155
5.3. Analysis of variance between 1st and 2nd Performances. 155

6.1. Protocol for the second experiment with the *b-Reactable*. 174
6.2. System and Performance aspects of *b-Reactable* Global and Local setups for the three experimental groups. 179

Nomenclature

BCI Brain-Computer Interface

CSCW Computer-Supported Cooperative Work

CSIM Cognitive Systems and Interactive Media (master course)

DMI Digital Music Instrument

ECG Electrocardiography

EDA Electrodermal Activity

EEG Electroencephalography

EMG Electromyography

EOG Electrooculogram

F-M Fixed mappings, Multiple EEG features displayed

F-S Fixed EEG-to-sound mappings, Single EEG feature displayed

GII Global Implicit Interaction

HAHV High Arousal, High Valance

HCI Human-Computer Interaction

HMK High Music Knowledge

LALV Low Arousal, Low Valence

LII Local Implicit Interaction

LMK Low Music Knowledge

MFR Maker Faire Rome

MHD Music Hack Day

MIDI Musical Instrument Device Interface

NIME New Interfaces for Muiscal Expression

OSC Open Sound Control

P-M Personalized mappings, Multiple EEG features displayed

P-S Personalized mappings, Single EEG feature displayed

Pd Pure Data

PhyComp Physiological Computing

SAM Self-Assessment Manikin pictorial scale

SID Sonic Interction Design

SMC Sound and Music Computing (master course)

V/A Valence and Arousal

VCF Voltage Controlled Filter

1. Introduction

1.1. Motivation

For more than 30 years the multidisciplinary field of Human-Computer Interaction (HCI) has been devoted to tackle technical and cognitive issues required for understanding the subtle interplay between people and computers. In this regard, a close look at the historical HCI curricula shows that the general mode of human interaction with computing systems has been mainly based on overt and explicit communication via peripheral devices such as the keyboard and mouse, a fact that has remained practically unchanged for the last 3 decades [Fairclough and Gilleade, 2014b] [Carroll, 2013]. The recent advent of mobile devices and gesture recognition algorithms represent an important shift on traditional methods of input control, but the basic interaction paradigm remains the same: a user sending explicit commands to the computer.

As devices and systems are becoming increasingly complex and powerful, this traditional approach to interface design is often lagging behind, constituting a bottleneck for seamless HCI. In the same way as people with tonal agnosia, who are unable to interpret the context, tone or intonation of the sentences, thus losing an essential component of speech to the point that they often cannot follow a conversation in a satisfactory manner, we are also limited by the constraints of current user interfaces. Communication with systems or within computer supported cooperative work (CSCW) normally takes place overtly, through explicit and conscious commands, by means such as e-mail or messaging platforms. In order achieve more natural interactions with computer systems, HCI has to be brought closer to the communication patterns of human beings [Hettinger et al., 2003], which imply to go beyond explicit control and incorporate the *implicit* subtleties that characterize human-human interaction. Through such approach we might be able to enrich user experience and contexts of interaction in the same way as archetypical conversational environments: interpreting the user body language or unconscious mental mechanisms, revealing meanings and hidden behaviors that would otherwise remain detracted from interactive systems and other users.

The affordances of implicit human-human interaction have already called the attention of HCI researchers and practitioners, as a way for making user information immediately and effortlessly available, and for relieving users from constant operation. For instance, Ju & Leifer [2008] propose a framework that helps to identify *when* and *how* to deploy implicit interactions, in order to develop less *needy* interactive systems. Buxton's taxonomy for periphery and context [1995] leverages on *background social ecology* to promote interactions that transit between foreground and background (the norm in human-human communication) to compensate the *sense of distance* that exist in computer and computer mediated interaction. Weiser and colleagues [1996] propose the concept of *calm technology,* striving for devices that engage center and periphery of user attention in a continuous tuning process. Tennenhouse's *proactive computing* [Tennenhouse et al., 1997] places implicit user behavior in the context of ubiquitous computing; and Schmidt defines implicit interaction within HCI as the result of a complex interplay between the context of use and the perceptual capabilities of computer systems [Schmidt, 2000].

Whereas most of the aforementioned approaches to implicit interaction are based on the user *situational context*, we can also incorporate implicit cues to HCI by directly inspecting user's brain and body states, as they represent a major outlet of information about her affective, perceptive and cognitive processes. In fact, the exploration of psychophysiology (i.e. the physiological bases of psychological processes) within HCI has led to the emergence of Physiological Computing (PhyComp), a field concerned with monitoring naturalistic changes in such phenomena for controlling computer systems [Fairclough, 2009]. PhyComp therefore appears as a valuable resource to expand implicit interaction beyond the *situational context,* by providing information about the *physicality* of users (what we call *subjective context*). Pragmatically speaking, PhyComp can enhance the perceptive affordances of interactive systems, making available user implicit states that can be only accessed by looking into the human body.

PhyComp has grown at a fast pace in the last decade. Moreover, with the recent appearance of wearable and non-invasive physiological devices - ranging from Brain-Computer Interfaces (BCI) to muscle and heart-rate sensors - PhyComp applications have gone beyond the lab, towards daily life environments such as gaming, sports and the quantified self [Allanson and Fairclough, 2004]. In parallel to this expansion, the field has developed alternative paradigms for interacting with computers without the need of overt forms of input, like using BCI or eye tracking systems [Tan and Nijholt, 2010][San Agustin et al., 2009]. However, the great majority of PhyComp paradigms are still designed to *read* conscious actions. They bypass body motion, looking for correlates of the user intention directly from the cortex of the brain or ocular movement, and translate these signals into actions at the interface level, but these still represent a form of input control that is explicit and intentional, just like pointing and clicking with a mouse. In these cases, the mode of input control is novel, but the mechanics of HCI remain essentially unchanged.

Nonetheless, we can use PhyComp techniques to enhance HCI with covert interactions that are neither deliberate nor volitional in any conventional sense. This is called *implicit PhyComp* [Fairclough, 2009], where the nervous system of the user is continuously *monitored*, and the resulting data is used to *characterize* her cognitive, emotional or motivational status. This dynamic representation is conveyed to the system in order to inform a real-time process of adaptation. This method of *wiretapping the user psychophysiology* [Fairclough, 2009] has been already explored by *biocybernetic* pioneers such as Alan Pope, who predicted the potential of implicit interfaces to adapt to the individual in highly personalized ways [Pope et al., 1995]. For instance, during the 1990's at NASA, Pope and colleagues developed an adaptive system for flight simulators where the pilot Electroencephalographic activity (EEG) was monitored in order to manage the status of an auto-pilot during flight time. The system aimed at sustaining the level of alertness of the pilot at an optimal level via manipulation of the auto-pilot status. Many other studies followed, such as computer games that adjusted difficulty according to the boredom of a gamer [Gilleade et al., 2005], or workload alertness systems for air-traffic control [Kaber et al., 2006].

These previous works already demonstrated how *implicit physiological input* can be used to tailor the experience to the singularities of a person (e.g. adapting parameters of a video game to the skill set of a gamer), a usage scenario (e.g. air traffic control), or a predefined agenda (e.g. mitigating frustration or promoting positive affect). However, their *output* methods are mostly direct, blatant and explicit: if the alertness of the pilot diminishes, the system prompts an alarm; if the gamer gets frustrated, the game explicitly reduces the number of enemies, if the user workload is too hight, the system cuts off tasks. But an implicit PhyComp paradigm inspired on natural human communication would probably benefit from the incorporation of *implicit outputs* as well. We humans have a wide range of explicit and implicit choices to communicate our thoughts, and a successful conversational interaction strongly depends on the hearer supplying the missing elements in the speaker's communication resource [Yus, 1999]. Therefore, implicit PhyComp has to account for implicit outputs that help the user to *interpret* the context of interaction, and not just providing unambiguous responses. In fact, human cognition is geared up for searching *relevance* in non-verbal, implicit stimuli processed in the course of a conversational interaction [Yus, 1999].

If we aim to develop seamless implicit interactions via PhyComp, our outputs have to be designed in a way that promotes relevance searching and rich interpretation as mentioned above, but in a computer-mediated context. This process of output or feedback design implies to render the user psychophysiological states by means of sound, visuals, and/or force [Jovanov et al., 1999][1]. Among different feedback techniques, sonic interaction design (SID) has shown to be an excellent medium for expressing implicit physiological states, both offline [Hermann et al., 2002] and in real time [Hinterberger and Baier, 2005, De Campo et al., 2007]. This is because acoustic data can be processed faster than visual presentation, can be easily localized in space, they can have a good temporal resolution (almost an order of magnitude better than visual [Jovanov et al., 1999]), and account for displaying multiple data streams while releasing the visual sense [Fitch and Kramer, 1994]. SID also permits a wide range of possible expressions, ranging from representative sounds for conveying information (sonification) to designs centered in expression and performance (i.e. new interfaces for musical expression - NIME) [Väljamäe et al., 2013a].

The link between sound and human physiology is long and extensive, showing that SID is an exceptional candidate for perceiving the complexity of implicit psychophysiological processes. For instance, sonification has been traditionally used for displaying physiological data in the medical practice, with the stethoscope being a remarkable example (100 years after its creation, it is still widely used by physicians). NIME researchers, on the other hand, have pioneered the use of biosignals[2] for SID in performance and creative contexts. Instead of exclusively attending at engineering problems or *functional* approaches (as in the case of sonifications) or solely looking at the user as a *receptor/listener,* NIME practitioners have mostly focused on designing PhyComp-based interfaces for supporting expressiveness. By doing so, NIME started to bridge the gap between instrument and interface design, exploring the connections of between audition, psychophysiology, touch and action. For all these reasons, it also appears as an excellent testbed for designing and evaluating physiology-based implicit sonic interactions in an expressive and relevant HCI domain.

There are, however, specific challenges that emerge from the intersection of implicit PhyComp and SID. One is related to **perceptualization**, understood as the process of associating a given sonic strategy to the psychophysiological state that acts as the input for its rendering, within an HCI context [Jovanov et al., 1999]. In short, perceptualization defines how well a sonic design represents a given implicit physiological state, aiding user perception during interaction. Naturally, any SID will always imply a level of arbitrariness. In explicit PhyComp, this arbitrariness is compensated by relying on unambiguous

[1] In fact research efforts are being put in the development of other types of feedback, such as digital smell and taste, but they are still at a very early stage of development [Ranasinghe et al., 2011].

[2] In this thesis, the term “biosignal” is used as a summarizing term for all physiological signals.

physiological inputs (for instance, a BCI based on motor imagery) and translating into explicit outputs (i.e. analogous cursor movement). In these cases, feedback design is relatively straightforward. For an implicit PhyComp system, however, the ambiguity of physiology-to-sound mappings augments, as a given sound might convey many psychophysiological states. On the other hand, the perceptualization quality of a given SID technique also depends on the unambiguity of the physiological input. Any implicit PhyComp system has to first measure, process and classify the physiological correlate (e.g. heart rate variability) of a rich and nuanced experience (e.g. emotional arousal). However, physiological streams are usually correlated with multiple user states (e.g. heart rate variability is also affected by fatigue and frustration) making psychophysiological classification a complex task.

Another challenge in the creation of sonic interactions based on implicit PhyComp is **mapping complexity**, understood as the number of physiological streams and sound parameters used in a SID strategy. Are direct physiology-to-sound mappings (e.g. audification) better than complex mappings (e.g. musical mappings) for perceptualizing implicit PhyComp in a HCI context? Even if we manage to provide an intuitive mapping to a given user, will it be perceived in the same manner by others? When screening the current literature, we see that most of the sonic designs applied to PhyComp work with a fixed mapping strategy [De Campo et al., 2007, Väljamäe et al., 2013a], meaning that users cannot change or fine-tune physiology-driven sonic interactions according to their individual perception, which can certainly affect the user experience. Moreover, when these SID strategies are applied in a HCI context (e.g. as part of a digital music instrument) studies tend to be rather elusive on determining how **meaningful** they are from an end-user perspective. In the context of NIME, this can be understood as the potential of a given physiology-driven sonic interaction for being perceived as an expressive component of the interface, through which the user can drive sound operations that, being expected or unexpected, contribute to the creative task she is committed to.

As mentioned by Fairclough and Gilleade [2014b], if implicit PhyComp is to be meaningful for users, it is important that the system *resonates* with the user experience in a seamless way. Therefore, it is the rationale of this thesis that if we:

1. systematically assess and evaluate the *perceptualization quality* and *mapping properties* of different SID techniques
2. provide *personalized* sonic interactions, tailored to the perception and subjective context of each user
3. deploy them in a relevant HCI scenario to determine whether the provided sonic

> feedback can be acknowledged (conscious or unconsciously) by the user as a *meaningful* component of the interaction

we will be able to improve user experience in implicit PhyComp systems, in manners that could hardly be obtained otherwise using pre-defined, fix display strategies. Through this approach, we aim to foster natural and seamless interactions that are closer to the communication patterns of human beings than traditional explicit HCI interfaces. Hence, analyzing end-user perception of physiology-based SIDs is the starting point of our work. In particular, in this dissertation we propose methods for designing and evaluating sonic interactions based on EEG and electrocardiography (ECG) data. The context of application is, in the first place, a prototype of an adaptive musical interface named *b-Reactable*. The outcomes from this initial study are later used for designing physiology-based sonic interactions of different mapping complexity, which are evaluated in a neurofeedback context. We also compare the performance of these designs to the one of personalized sound mappings, in a neurofeedback scenario. The results of this study are used to inform a second version of the *b-Reactable*, which is evaluated in a music performance context.

1.2. Background Knowledge

This dissertation leverages on two main domains for researching on natural interfaces based on implicit paradigms, namely physiological computing (PhyComp) and sonic interaction design (SID), the latter covering both sonification techniques and NIME. Below we provide a short summary in these fields.

1.2.1. Physiological Computing

PhyComp represents an HCI modality where system interaction is achieved by monitoring, analyzing and responding to psychophysiological activity from the user in real-time [Allanson, 2002]. Although the field has been historically nurtured by medical research, clinical practice and different cognitive sciences such as psychophysiology and neuroscience [Treacy Solovey et al., 2015], other disciplines such as music technology[3] and video game[4] research have pioneered the use of biosignals for creative and expressive purposes [Nijholt, 2009, Rosenboom, 1997a]. For this reason, interface design within

[3] for more information about the history of PhyComp in the music domain, please check the fundamental work of David Rosenboom “Extended Musical Interface with the Human Nervous System”[1997a].

[4] for a comprehensive review, please check Kivikangas et al. [2011].

PhyComp historically responded to specific demands of such fields, like data acquisition and screening (e.g. topographic maps for diagnosis and study of brain functions), rehabilitation tools (e.g. post stroke motor rehabilitation [Cincotti et al., 2012]), prosthesis design (e.g. control of a prosthetic hand through electromyography [Cipriani et al., 2008]), but also new interfaces for musical expression and gaming systems [Nacke et al., 2011].

PhyComp interfaces are already present in the consumer marketplace, shaping the creative industry. The recent emergence of off-the-shelf wearable and mobile devices (e.g. smart watches and non-invasive BCI) together with low-cost prototyping toolkits and miniaturized sensors (e.g. the BITalino[5]), have fostered the use of biosensors[6] beyond laboratory conditions, complementing explicit peripheral behavior (e.g. movement) with physiological states. Not surprisingly, the computer industry is already releasing controllers that capture detailed hand and finger motion by combining inertial measurement units (IMU) with muscle activity detection (like the wearable controller MYO[7]), or interfaces like the Microsoft Kinect 2, which detects heart-rate variability to assess gamers' responses to the context of interaction. In this way, PhyComp is rapidly taking on every-day life in the shape of activity trackers (both for wellness like the Jawbone [8] and professional sports like the Garmin Edge line[9]), cognitive trainers (like the NeuroSky MindWave headsets[10]) and alternative ways of communication (the Apple Watch hear-rate sharing system[11] being one of the most populars).

Despite biosensing became less invasive, more accessible and pervasive, the experience for many users is still frustrating since natural interaction does not yet meet their expectations. Most of the current PhyComp applications are limited to the measurement and visualization of collected physiological data, thus not exploiting the possibilities of PhyComp for real time interaction. This scenario has motivated research on different PhyComp paradigms to foster meaningful interactions beyond monitoring, as described in the following subsection.

[5] `http://www.bitalino.com/` (accessed on November, 2015).
[6] In this thesis the term "Biosensors" are used interchangeably with "physiological sensors".
[7] `https://www.myo.com/` (accessed on November, 2015).
[8] `https://jawbone.com/` (accessed on November, 2015).
[9] `https://buy.garmin.com/en-US/US/into-sports/cycling/edge-520/prod166370.html` (accessed on November, 2015).
[10] `http://store.neurosky.com/collections/eeg-headsets` (accessed on November, 2015).
[11] `https://www.apple.com/watch/new-ways-to-connect/` (accessed on November, 2015).

Interaction Paradigms in Physiological Computing

As mentioned in 1.1, there are two predominant interaction paradigms in PhyComp: direct explicit control and indirect implicit adaptation [Fairclough and Gilleade, 2014a]. The goal of **direct explicit control** is to extend the body schema through interfaces that are guided by intentionality, as in the case of BCIs or PhyComp based on peripheral signals such as electromyography (EMG) [Caramiaux et al., 2015b]. This model is analogous to command inputs (e.g. keystrokes) and has been mainly used for compensating motor disabilities, prosthesis control or rehabilitation [Wolpaw et al., 2002b]. Nowadays, explicit PhyComp applications are also designed for healthy users [Allison et al., 2007], although they still face important constraints (particularly BCIs) such as a limited control bandwidth, the need of training phases for accurate classification, and limitations of human attention [Wickens, 2002]. Previous studies have shown that direct explicit PhyComp can improve user experience and performance in goal oriented scenarios, such as video games [Nacke et al., 2011], although its seamless integration with traditional methods (i.e. gamepads or multitouch interfaces) remains a major challenge, as direct explicit PhyComp also requires intentional user commands.

Indirect implicit PhyComp, on the other hand, controls the state of the system through a *biocybernetic* loop [Pope et al., 1995] that continuously monitors psychophysiological changes for providing adaptive responses perceived as intuitive and timely by users Fairclough [2009]. It does not require consciously generated input commands or training, and can seamlessly encompass multimodality by being combined with other biopotentials and/or traditional input methods [Zander et al., 2010]. As mentioned in Section 1.1, early works as Alan Pope's adaptive piloting systems made use of this strategy for enhancing an already existent human-machine scenario by covertly informing the computer system about the user alertness state [Pope et al., 1995]. Following this paradigm, several adaptive multimodal systems have been developed using cortical oscillatory activity [Nijholt et al., 2009], ECG and electrodermal activity (EDA) [Nacke et al., 2011]. Previous studies on video games emphasize the qualities of implicit PhyComp for influencing long-term elements of the user experience, and its advantages for sustaining engagement through real-time physiological feedback [Rani et al., 2005, Gilleade et al., 2005] in affective games such as Ozen (Ubisoft)[12], which focuses on stress control.

Explicit and implicit interactions are not mutually exclusive. Stephen Fairclough and Kiel Gilleade[13] have proposed a four-group classification to better represent the ex-

[12]http://www.experience-ozen.com/ (accessed on October, 2015).
[13]For a complete description of Fairlough and Gilleade's classification, please check http://www.physiologicalcomputing.net/?page_id=227

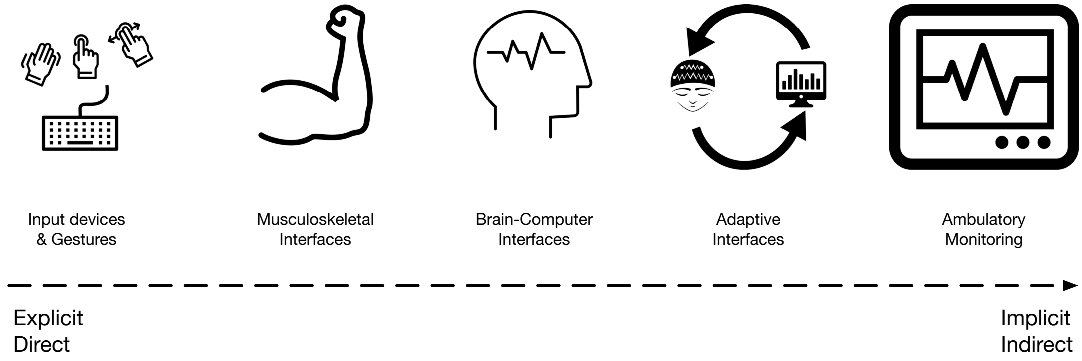


Figure 1.1.: Classification of Physiological Computing systems, based on the model proposed by Fairclough and Gilleade (`http://www.physiologicalcomputing.net/?page_id=227`).

plicit/implicit interaction continuum (see figure 1.1). It starts with *conventional input devices*, including gestures and motion detection, which are eminently explicit. Interactions driven by *Musculoskeletal activity* (e.g. eye movement) are the next in the classification, as they can be controlled volitionally but they are also subject to unconscious action (e.g. reflexes). The third group is *BCI*, where interactions bypass peripheral body movement to be directly driven by brain activity and autonomic responses from the nervous system. It is important to note that these three categories are all concerned with a degree of explicit input control at the interface level, and an analogous explicit output (e.g. left hand movement/eye gaze toward the left/imagined left hand movement → move a cursor to the left). They are therefore designed to respond to conscious actions, thus the user expects an unambiguous output as a result.

The category of *adaptive interfaces*, on the other hand, is built under an implicit indirect paradigm, thus requiring some kind of monitoring system to capture physiology in applied environments (by means of unobtrusive wearable sensors) to inform the system about spontaneous changes in psychophysiology (as in the case of Pope's adaptive piloting system). The final category of *ambulatory monitoring* refers to the use of pervasive monitoring for self-learning and self-diagnosis (i.e. tracking the progress of physical training).

Under the scope of Fairclough and Gilleade's continuum, this dissertation will mainly tackle adaptive PhyComp interfaces to design implicit sonic interactions, and explore their combination with other explicit methods to foster multimodality.

Feedback design for indirect implicit PhyComp

In PhyComp, the feedback design process is traditionally defined by its *directionality.* In this sense, research efforts have mainly focused on two modes: positive and negative feedback control[14][Carver et al., 2000], and their combination as they are not mutually exclusive. **Negative feedback** creates behavioral stability by reducing the discrepancy between the input state (e.g. physiological correlate of engagement) and a desired standard (e.g. high engagement). Negative feedback is ideal for adaptive systems designed to keep the user in the *flow* [Csikszentmihalyi and Csikzentmihaly, 1991]. Think for instance on a first-person shooter. Negative feedback would be desirable for parents that will like to keep the stress of their children low if they game becomes too challenging (i.e. more user stress, less difficulty).

Positive feedback, on the contrary, is generally designed to amplify the discrepancy between the input state and the desired state in an exponential fashion. This leads to performance instability, and it may therefore be used to adjust the desired target state upwards as the user becomes more involved with the task. If we go back to the fist-person shooter example, positive feedback would be preferred by a gamer willing to improve her game skills (i.e. more perceived difficulty will trigger more challenging obstacles).

Both types of feedback can be combined to toggle unstable episodes of skill acquisition (positive feedback) and stable moments for skill consolidation (negative feedback). Design decisions in this regard are essential, especially in the context of goal-oriented HCI domains such as video games or training applications.

In this dissertation we also explore an extra dimension of feedback design, **mapping complexity**, defined as the the number of physiological inputs and output parameters used to produce and display system adaptations. We have focused on this dimension because, whereas *directionality* is particularly important for training systems (e.g. biofeedback) or goal oriented applications (e.g. adaptive video games), *feedback complexity* plays a major role when PhyComp coexist with other input methods (e.g. tangible interaction) and in expressive HCI domains like musical performance. As pointed out by Jordà and Mealla [2014], mapping strategies are crucial for expression and for domains where content exploration or creation are as relevant as task solving. In this context, for instance, a music performer playing alone might prefer a physiology-to-sound mapping with a great *complexity* to drive several musical operations overtly and automatically, while focusing on specific musical aspects explicitly. For collaborative performance, on the other hand, a reduced *complexity* might be preferred for avoiding undesired sonic outcomes.

[14]For a comprehensive review on this topic, please refer to Fairclough [2009]

The development of seamless implicit sonic interactions that bring HCI closer to human-human communication patterns thus requires a close study on the perception of different feedback designs, taking into account the characteristics of the end user and the context of interaction. Therefore, in this thesis we will design and compare different sonic feedback strategies, stressing on the implications of user preference (through personalized mappings) and context (ranging from a canonical psychophysiological setup -neurofeedback- to expressive HCI -NIME).

1.2.2. Sonic Interaction Design

Sonic interaction design (SID) can be defined as an interdisciplinary field of research and practice that explores the ways in which sound can be used to convey information, meaning, and aesthetic and emotional qualities in interactive contexts [Franinović and Serafin, 2013]. For years, designers have been rarely aware of the extent to which sound can affect user experience, especially in a computer-mediated context. SID therefore emerged from the need to challenge predominant design approaches by considering sound as a key medium to foster novel phenomenological and social experiences with and through interactive technology. SID leverages on knowledge and methods from diverse disciplines such as interactive arts, electronic music, cultural studies, psychology, cognitive sciences, acoustics, and interaction design to explore multisensory, performative, and tactile aspects of sonic experience, and the ability of a sounding object to communicate meaning.

SID as representation: Sonification

From an HCI perspective, sound has been predominantly used to encompass screen-based interaction or to present information in form of sound. This *representational* use of sound, known as sonification or auditory display[15] [Kramer, 1993], has defined a number of functional roles for sound, such as the exploration of big databases, aid awareness through sonic alerts, or inform about action accomplishment (i.e. auditory icons) like the sound of crashing paper when moving a file to the trash bin on a graphic user interface (GUI) [Gaver, 1993]. Kramer [1993] defines sonification as "the use of non-speech audio to convey information". More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.

[15] In this thesis, auditory display is used as synonym of sonification.

Figure 1.2 shows the general structure and necessary conditions for sonification, and the different closed-loop interaction that can take place in an auditory system, as proposed by Hermann [2008]. The sonification module plays rendered sonifications to the user. Data sources feed sonification from the left side, including any system in the world that is connected to the sonification module (e.g. via sensors that measure its state), or any data under analysis that are stored separately and accessible by the sonification. The arrows describe different user interactions that can happen under this setup, with *monitoring* being the less interactive procedure, since data from a repository is continuously feeding the sonification rendering with the user being a passive listener, focusing only in parts of the sound. *Navigation*, on the other hand, implies the selection and/or browsing of chosen data. *Physiological feedback* refers to the sonification of real-time physiological data gathered through different sensors. In this case, the user produces the input data for the sonification system, making possible for her to perceive a sound that depends on her own physiological activity. Finally, *human activity supported by sonification* focuses in the physical world, and the user is driven by the goal to change a world state in a specific way.

In the recent years, research efforts have also shifted towards exploring the relations between human action and sound, introducing the notion of *interactive sonification*, where auditory signals, besides providing feedback on the data analyzed, is also refined and redirected according to the user activity (i.e. changing and adjusting parameters of the sonification module). However, in this case user action is normally channeled through traditional user interfaces (e.g. screen icons and mouse or touchscreen), or reduced hand movements. Therefore, the performative aspects of sound are not exploited to their full potential.

Despite the affordances of sound for displaying physiological activity (e.g. detecting repetitive elements, regular oscillations, discontinuities, and signal power to a degree comparable with using visual inspection of spectrograms [Pauletto and Hunt, 2005]) any sonification strategy implies a level of arbitrariness in its conversions of physiological data into sound. In this regard, most of the published work in the field do not provide sufficient details about either physiological data acquisition or applied sound synthesis [Väljamäe et al., 2013a]. Moreover, very few of these studies have conducted controlled evaluations of the chosen methods, making it difficult to replicate or validate most studies.

To overcome these aspects, this dissertation will provide a systematic identification of application domains for physiology-based implicit sonic interactions; it will also perform controlled evaluations and comparative studies to determine how different sonic strategies perform in a given HCI context, and how they differentiates/complement

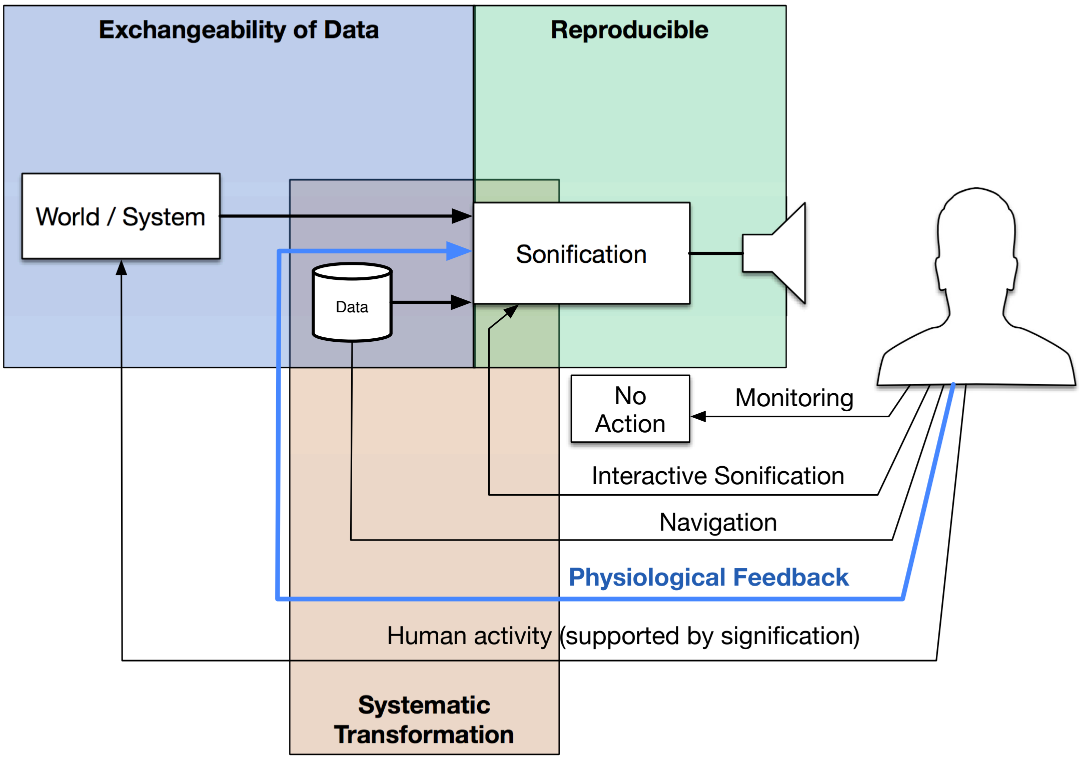


Figure 1.2.: Structure and conditions needed to achieve different types of sonifications, derived from Hermann [2008]. Physiological feedback (the sonification loop explored in this thesis) is highlighted in blue.

other interaction modalities (e.g. tangible and gesture interaction). This thesis will also investigate till what extent personalized mappings can compensate for arbitrary physiology-to-sound conversion.

SID as performance: New Interfaces for Musical Expression

Whereas sonifications allow us to represent the dynamics of physiological data by means of sound, the creation of seamless HCI experience also requires a carefully designed interface and sonic behavior. In this regard, the field of new interfaces for musical expression (NIME) has pioneered the use of physical -and more concretely physiological- resources for developing embodied, tangible interfaces where user performance is at the center of the design process. Research on digital musical instruments (DMI)[16] has widely explored the importance of body actions for creating expressive interfaces. Through this approach, NIME has placed performance and physicality [Hornecker, 2011] at the center of the design process, focusing on the body, the materiality of objects, and the context to design sonic interactions that intimately engage user behavior [Miranda and Wanderley, 2006]. NIME has expanded the scope of SID, as it does not approach the user as a mere *receiver* of auditory stimuli, but rather explores the perception-action loops that are mediated by acoustic signals [Franinović and Serafin, 2013]. In fact, music technology researchers have pioneered the use of biosignals in performance contexts. Avant-garde musicians such as Alvin Lucier, Richard Teitelbaum and David Rosenboom were among the early adopters of biofeedback for developing adaptive systems [Rosenboom, 1997a]. Musicians were among the first on experimenting with multimodal physiological input, as in the case of Spacecraft (1967), were EEG and ECG were used as a control sources for electronic synthesizers, including the nascent Moog electronic synthesizer, or Biomuse [Tanaka, 2000], that was also based on synthesized music using real time EMG data.

This dissertation aims to contribute to the SID field by designing and evaluating sonic interactions in both reception-based studies (i.e. sonification for neurofeedback) and performance-based studies, by deploying sonic interactions in a NIME scenario. Through this approach we will address individual and collaborative experiences, where sonic adaptations respond to both the implicit physiological states of the user, her physical actions (e.g. manipulation of tangible objects) and the context of the experience (i.e. playing music alone and with others).

[16]In this thesis, "Digital Music Instrument" is used to refer to devices (i.e. musical interfaces) whereas "New Interfaces for Musical Expression" is used to name the discipline and practice of musical interface design.

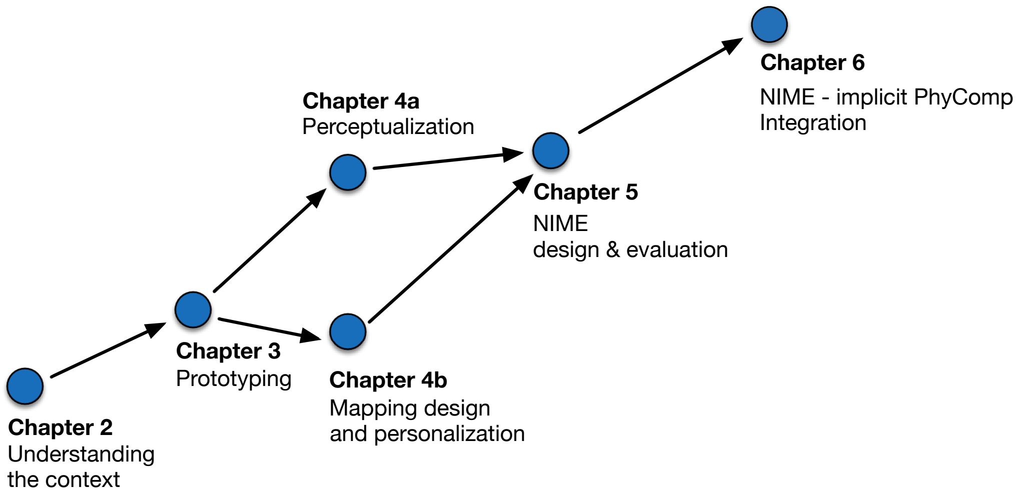


Figure 1.3.: Graphical representation of the thesis structure.

1.3. Objectives and Outline of the Thesis

In the previous subsections we have exposed the problems that motivate this thesis, described their context and how we plan to approach it in a systematic manner. Following we provide a description of our research objectives and a summary on the organization of this document.

The main goal of this thesis is to conceptualize, prototype and evaluate sonic interaction designs that incorporate the implicit cues of human psychophysiology in meaningful HCI contexts. For achieving it, we leverage on physiological computing technology to obtain inputs from user's implicit (i.e. perceptive, emotive and cognitive) states in real time. SID methodologies are then used to adapt system responses according to these implicit statuses, thus providing sonic outputs that can aid perceptualization, personalization and meaningfulness in a computer-mediated domains. We incrementally develop and evaluate implicit sonic interactions (from direct audification to complex musical mappings) and apply them in different HCI contexts (from neurofeedback to music performance). Figure 1.3 illustrates the evolution of this dissertation according to main research problems, context, and SID approach. What follows is a brief description of the structure of this document, including specific goals, a summary of methods and main results. Each chapter also includes a specific discussion on relevant results and conclusions.

In **Chapter 2** we provide a comprehensive literature review on PhyComp, implicit interaction, and SID. The goal of this section is to provide enough background for the reader to understand the state of the art on these fields, and the way they converge within HCI. For achieving this goal we analyze the different interaction models that coexist within PhyComp, stressing on implicit interaction and its differences with explicit approaches. Regarding the SID literature, we review different methodologies and design strategies for data representation (sonification) and performance (NIME), also describing their approaches to PhyComp (both offline and in real time) from an historical perspective.

Chapter 3 presents our first conceptual, technical and methodological approach to SID for implicit PhyComp. The goal of this chapter is to study how implicit physiology-based interactions driven by sound can affect user experience in a meaningful HCI context: digital musical instruments (DMI). To do so, we start with a simple and straightforward SID strategy: audification of EEG activity and temporal control of sound by means of ECG. We use these SIDs to create a DMI based on the Reactable, a tabletop interface for music performance [Jordà et al., 2007]. The prototype that emerges from this combination is called *b-Reactable*, and being based on a previous tabletop system, it allows explicit gestural interaction (through tangible objects) and implicit interaction (through EEG and ECG) for sound generation and control. For a coherent integration with the Reactable framework, EEG and ECG SIDs are embodied in tangible objects named *physiopucks*, which add a level of physicality to the above mentioned implicit interactions. We evaluate the effects of this approach in user motivation, and compare them to the original Reactable. The experiments involve dyads collaborating in three experimental groups. The results of this chapter show that motivation dimensions are significantly higher in *b-Reactable* than in the Reactable, stressing on the positive effects of physiology-based implicit sonic interaction, and its combination with other inputs methods even in multi-user HCI scenarios.

Chapter 4 presents a set of studies meant to address two SID aspects that emerge from our first experiment with *b-Reactable*: *perceptualization* and *mapping complexity* in physiology-based implicit sonic interaction. The goals of this chapter are to determine (i) what types of sonic designs perform best in representing a given implicit physiological state (e.g. relaxation) according to end-user perception (*perceptualization*) and (ii) whether *mapping complexity* and mapping personalization by end-users play a role in the perceptualization of implicit physiological states through sound. We address these issues separately by means of two experiments based on neurofeedback training. The first one assesses the *perceptualization* quality of the most used sonic designs for displaying EEG activity (as suggested by our literature review). The evaluation is based on end-user perception (both subjective and through physiological measures) of own relaxation states estimated from EEG alpha activity. The second study leverages on the findings of the first one to implement a sound engine (i) capable of generating designs of different *mapping complexity* (both in terms of physiological streams and sound parameters) and (ii) able to be personalized by end-users through sliders. Both *mapping complexity* and personalization are tested in alpha/theta neurofeedback training, collecting subjective and objective (EEG) measures of relaxation. Results from the first experiment suggest that parameter mapping sonification and musical mappings are good candidates for *perceptualizing* implicit physiological states, whereas the second experiment provides insights about the positive effect of *mapping complexity* and end-user personalization in the *perceptualization* of sonic designs for physiology-based implicit interaction. The studies also empirically demonstrate that personalization becomes less instrumental when multiple physiological features are displayed through sound.

One of the main SID aspects to be explored in this dissertation is the *meaningful* integration of implicit PhyComp in a relevant HCI context, namely music performance. Since this task requires systematic and specific evaluation methods, **Chapter 5** presents a framework for designing and evaluating NIME, with special focus on *expressiveness,* a fundamental aspect to determine the *meaningful* contribution of implicit PhyComp in the NIME field. The objectives of this Chapter are (i) to analyze the most relevant NIME design and evaluation frameworks available in the literature, (ii) to identify how they tackle different stakeholders, and (iii) to propose and test a framework focused on *expressiveness*, on the *mapping* component in the NIME creation chain, on different stakeholders (i.e. *designers*, *performers* and *listeners*), and considering how previous music knowledge would affect each of these roles. This framework is deployed in a one-trimester NIME master course where groups of participants (i.e. students) prototype DMIs within a restrictive setup, consisting of smart-phones controllers and the Pure Data (Pd) programming language, and perform with them in front of the rest of the class, which in turn evaluates the performances as *listeners*, in an iterative process. The insights gathered during the study suggest that students with different backgrounds were able to effectively engage in the NIME design processes; that the assessment tools proved to be consistent for the evaluation of *systems* and *performances* aspects of NIME; and that the outcome of the evaluation translated into a traceable progress in the students' DMIs.

Chapter 6 addresses the issue of *meaningfulness* that, together with *perceptualization* and *mapping complexity*, constitutes one of the main aspects of physiology-based implicit sonic interaction in the context of this dissertation. The goal of this Chapter is to systematically explore how implicit PhyComp contributes to the design of a digital musical instrument (DMI) in a *meaningful* way, implying that it will be perceived as an expressive component of the interface, through which the player is able to produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to. To do so, we create a new version of the *b-Reactable*, that incorporates a number of features informed by the results of our previous experiments in *perceptualization* and *physiology-to-sound mapping*, namely a parameter mapping approach, end-user personalization, and a more complex psychophysiological input (valence and arousal estimated through EEG) operated through two different setups (Global and Local). We test this new incarnation of the *b-Reactable* in an expressive context (i.e. music performance) involving 15 participants with different levels of music experience, who perform musical improvisation exercises under two conditions (Global and Local implicit interaction). The main results show that our affective estimations are valid for the context of music performance, and that participants use these implicit sonic interactions (both Global and Local) in a distinctive and meaningful manner. Subjective, behavioral and psychophysiological data show that Global and Local implicit interaction is perceived in significantly different ways according to participants' previous musical experience, with preference for the latter.

Chapter 7 can be seen as a *bonus track* that aims to explore implicit PhyComp beyond SID and the musical domain. In order to achieve this goal, we approach the field of personal fabrication and present *NeuroKnitting*, a system that can be used to create knitted garments according to the users' affective responses estimated from EEG. We deploy this system in two recording sessions, from which we extract preliminary insights and design guidelines. The tests show that *Neuroknitting* can be used for embodying implicit psychophysiological data into unique, customized physical objects. As every human being reacts differently to a given experience, the knitted patterns change according to the user and her context. *NeuroKnitting* thus opens the door to further structured analyses on aspects such as the perception of implicitly generated fabrication patterns, and the use of different stimuli to trigger meaningful user experiences during fabrication precesses.

The document concludes with **Chapter 8**, which provides a summary of the work done, list of main findings, and discussion on the future perspectives of physiological computing and sonic interaction design.

2. Literature Review

2.1. Introduction

In this chapter we provide a comprehensive review on the topics and fields that conform the corpus of this thesis. The chapter starts with a review on different approaches to implicit interaction for HCI perspectives (theoretical, technological and design standpoints). Then we discuss how implicit interaction is tackled by physiological computing (PhyComp), providing an introduction to biosignals and physiological computing systems, with special focus on the way they foster new forms of implicit interaction. We then move towards sonic interaction design (SID), as sound and music are the main display techniques employed in this dissertation. We put special emphasis on representational (i.e. sonification) and aesthetic, performance-based approaches (more concretely, new interfaces for musical expression). Finally, we offer a detailed analysis on the state of the art of SID applied to electroencephalography (EEG), the main biosignal applied in this dissertation. The goal of this section is to provide enough background for the reader to understand the framing of this thesis and its contribution to the current state of the art in the related fields.

2.2. Designing Implicit Interactions

When observing communication between humans we can see that great amounts of information are only exchanged implicitly through gestures, body language and voice [Schmidt, 2000]. The way people interact with each other and the situation in which this happens, carries information that is often exploited implicitly to improve human-human communication by, for instance, disambiguating information. We people rely on implicit cues to adapt our actions with considerable ease in a great sort of circumstances, adapting the way we speak or performing actions that we believe are expected in a given situation. And we do all this without conscious thought. As mentioned by Ju and colleagues [2008], these accommodations do much to *smooth* our day-to-day interactions with one another. Without the need of explicit commands we can therefore manage attention and cognitive load, and measure expectations.

The interest of HCI for leveraging on our great experience in implicit interaction was rooted in the dawn of interactive computing. In 1960 J.C.R. Licklider envisioned a *human-computer symbiosis* based on cooperative interaction between humans and electronic computers, involving a close coupling between the members of the partnership [Licklider, 1960]. This symbiosis aimed to enable humans and computers to make decisions and cope with complex situations *together* in a cooperation that does not rely on inflexible predetermined programs. This would make computer systems *sensitive* to the interaction context and *proactive* beyond human supervision.

Approaches like Licklider's led to the advent of more natural ways to interact with computer systems (concretely, graphic user interfaces -GUI- and Windows, Icons, Mouse and Pointer -WIMP- interactions) which were eminently overt, based on conscious inputs from the user, and explicit outputs from the system. Commands therefore shifted from symbolic (i.e. the command line) to object-oriented metaphors directly manipulated through interfaces such as the mouse, but HCI research and practice were mainly producing interfaces that constantly required human attention. This approach has governed desktop computing for more than 30 years, and it works well when the user has attentional and physical availability, but it is far from ideal when we think about current interaction scenarios, characterized by devices integrated in our daily life in the shape of wearable, mobile and embedded interfaces. As mentioned by Neil Gershenfeld [1999]: "We are expending more and more time responding to the demands of machines".

As a response to this situation, different disciplines within HCI have strived to create more natural, seamless interactions inspired on the implicit nature of human-human communication. It mainly started in the late 1990's, fostered by the concurrence of emerging embedded computing, the tremendous leap on wireless connectivity and network infrastructures, and the popularization of mobile and *smart* devices [Tennenhouse et al., 1997]. Following we describe different approaches that sought to incorporate the implicit within interactive computing, from theoretical, technological, and design perspectives.

2.2.1. Buxton's taxonomy for integrating periphery and context

In 1995, Bill Buxton proposed an human-centric model for incorporating non-intentional interaction based on *foreground* and *background* [Buxton, 1995]. The model was created to aid the development of new technology by looking at both human-computer and

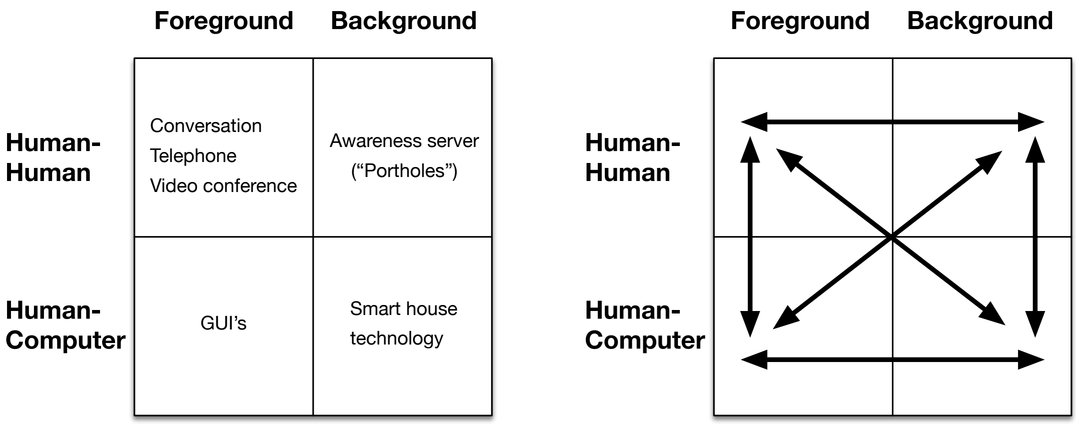


Figure 2.1.: Buxton's basic model for integrating periphery and context in HCI, and its possible transitions among quadrants [Buxton, 1995]

computer supported interaction. Buxton's model is based on a 2 x 2 matrix illustrated in Figure 2.1. The two dimensions are the *ground* (horizontal axis) and the *object of communication* (vertical axis). The former imply activities that move from the fore of human consciousness (foreground) such as speaking on a telephone or typing into a computer, and tasks that take place in the periphery of human attention (background) like being aware of the presence of another person by hearing her typing from another room, or a light switching on automatically when we enter home. The second dimension (*object of communication*) refers to who or what the user is communicating with, always in a technology mediated scenario. In this way, a phone call between two people is an example of human-human communication, whereas operating a computer through a GUI fall into human-computer communication.

Back in the mid 90's telematic technologies such as chatting, telephone and videoconferences were a reality, but Buxton questioned the *sense of distance* that existed in technology mediated communication. He associated it to the lack of key affordances that occur naturally when people collaborate and interact in close physical proximity. Interacting in a shared space, even without being in the same room, makes background information immediately and effortlessly available (i.e. a light on in the office next to ours might indicate that someone is working at late hours). This background information fosters what Buxton calls *peripheral awareness,* and the lack of it is what makes the right hand side of Buxton's model so difficult to populate, "the *real sweet spot* for interactive applications" [Buxton, 1995].

Buxton's taxonomy leverages on this *background social ecology* to develop interfaces that share this periphery among users, and combine it with explicit, foreground technology to improve the sense of presence and quality of communication. For instance, the "awareness server" *Portholes,* developed by Duorish & Bly [1992], offers information about the surrounding space (who is next door? what is she/he doing? is she/he available?) by sending video snapshots every 5 minutes to a desktop application that remains in the background of our attention. These types of *background interfaces* aimed to improve computer supported collaborative work by letting us know when a colleague is in the building, if he/she is engaged in a given task, if it is a good moment to interrupt him/her, etc. Under Buxton's perspective, a *smart lighting system* that automatically switch on the lights when we arrive home can be considered a background human-computer interaction.

Further elaborations can be made from this bi-dimensional framework by promoting seamless transitions between quadrants as illustrated in Figure 2.1. Buxton emphasizes on how background tools like *Portholes* can help to seamless go from top right to the top left quadrant if we, for instance, use it to check if a given person is available (by looking at his/her last frame) and if so, use *Portholes* to make a call. If, for instance, the person we want to contact is not available, we can setup this background tool to automatically detect when he or she is back and automatically deploy an alert to proceed with a videoconference. In that case, we move from the bottom right quadrant (when the system looks at the portholes images while we are focusing in another task) to the bottom left (when the alert appears).

Transitions between background and foreground are norm in human-human communication, and a significant amount of the complexity in humans dealing with technology is due to having to explicitly sustain foreground activity. Everyday communication is completely different, and Buxton's approach shows how we can significantly reduce complexity of interactive technology if, likewise, information is pushed to the background through a context sensitive approach. Buxton's model, however, also presents limitations. On the one hand, implicit cues are exclusively modeled from contextual information (i.e. user perception is not directly addressed). On the other hand, as stated by Ju and Leifer [2008], Buxton's model of attentional foreground only considers user-initiated interactions (e.g. typing on a keyboard) therefore conflating attention with intention. This approach is therefore inadequate for device-initiated interactions (e.g. a cellphone ringing). These interactions clearly take place in the foreground, but are not intentionally driven by the user.

2.2.2. Calm technology

In a similar key as Buxton's taxonomy, Mark Weiser and John Seely Brown also stressed on the need of designing for the periphery [1996]. This approach, called *calm technology*, aimed to create devices capable or fostering user comfort by engaging both the center and periphery of our attention in a continuous tuning process. Periphery is anything but unimportant, as what is in the periphery at one moment might well be at the center in the next. As periphery is informing without overburdening, calm technology should be capable of seamlessly move from the background of our attention to the foreground, and back. Calm technology shows three main empowering properties: (i) it facilitates the *motion* between center and periphery (and the power of the user for controlling such tuning action), (ii) enhances our *peripheral reach* by bringing more details into the background, and (iii) promotes *locatedness* by putting us in a familiar place where we know what is happening around us.

Like in the case of Buxton's taxonomy, Weiser and Brown look for the implicit in the context without looking at user perception or user generated information. Likewise, their approach is device-oriented, as it overlooks user-generated actions and human-human interaction.

2.2.3. Proactive computing

In the early 2000's, and from a ubiquitous computing perspective, David Tennenhouse proposed a shift in the HCI agenda from human-center to human-supervised computing [2000]. In the same line as Buxton's and Weiser's approach, this was also motivated by the limitations of *human-in-the-loop* computing, as networking computers start to outnumber human beings. In order to craft a research agenda for this *excess* of networked and embedded interactive systems, Tennenhouse proposed a different mode of HCI operation, denominated *proactive computing*, by rethinking the boundaries between the physical and the digital world, and the time at which computation happens, in order to foster unsupervised systems that can relief human from constant operation demands, thus increasing productivity and quality of life.

Proactive computing builds implicit interactions by being intimately connected to the physical world by means of pervasive sensors and actuators. As show in Figure 2.2, proactive computing aims at greater system autonomy in complex daily life environments, as opposite to traditional human-centered HCI approaches. Rather than being in direct contact with the user, proactive computing interfaces with its physical context, monitoring and modeling the surroundings of interaction.

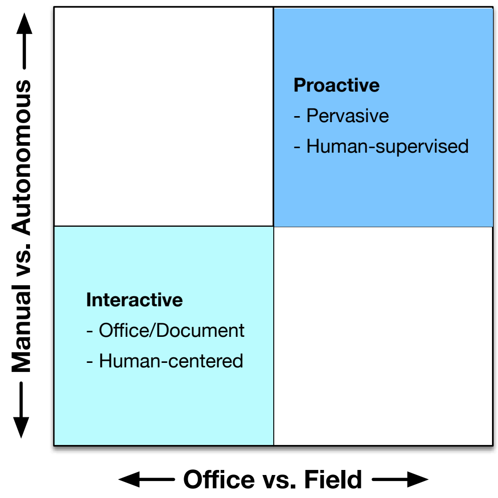


Figure 2.2.: Tennenhouse's proactive approach within the four quadrants of ubiquitous computing [2000].

Proactive computing is an approach mainly based on pervasive technology for context modeling. Therefore it focuses on gathering implicit inputs from the context and providing explicit outputs to the user. Therefore, it does not consider aspects related to user perception and cognition, like the multiple manners in which context can be perceived, and its effect on human behavior. Although Tennenhouse discuss the challenge of bringing software closer to the physical environment, methods for displaying systems responses (both implicitly and explicitly) are not tapped in this model.

2.2.4. Implicit HCI through context

Albrecht Schmidt proposed an approach that directly tackled implicit HCI based on the interplay between the context of use and the perceptional capabilities of computer devices [2000]. His vision was motivated by the shift from desktop to mobile interaction that started at the beginning of this millennium, which promoted interaction periods often shorter than in the case of traditional desktop computing. This implied new contexts of interaction, where applications are mainly used while doing something else, or to carry out a certain task in the real world (like looking for a place in a navigation app). These contexts called for a reduction of explicit human-machine interaction to promote more seamless and natural interfaces. From Schmidt's perspective, devices with perceptional capabilities would be able to adapt applications to highly changing contexts by looking at aspects such as time (an specific hour, morning or night...), number of people concurrently using the application, conditions of the physical environment (location, type, temperature, etc.) social settings, etc.

From Schmidt's perspective, implicit interaction happens when the users performs an action that is not primarily aimed to interact with a computer system, but the system understands it as an input. Under this premise, if a system has a level of knowledge about *the context in which the interaction takes place*, implicit interaction can happen. As overt action from the user her/himself is not required, it can easily be combined with other concurrent applications that require explicit inputs.

As Buxton's taxonomy, Schmidt's approach is mainly focused on how the system directly *perceives* and *interprets* the world, without directly addressing how the user perceives or interprets the same context of interaction. It also discusses how to provide implicit inputs to the systems (those that are not consciously produced by the users) without considering the ways in which the system could provide implicit outputs to happen in the periphery of user's attention, or initiate interactions in an implicit or explicit manner. This context-centric approach to implicit HCI translates to 3 main building blocks:

1. The ability to *perceive* use and environment
2. Mechanisms to *understand* what is perceived
3. Applications to make a *meaningful use* of this information

(1) and (2) define what Schmidt calls *situational context*, which includes location, the surrounding environment and the state of the device being use for interaction. (3) refers to applications that are *context enabled*. In order to equip devices with the required perceptional capabilities, Schmidt consider 4 basic approaches: databases (like calendars, address books, etc.), explicit inputs to applications running in the device (e.g. taking notes, adding an event in the calendar), active environments (cameras, audio, etc.) and sensors (GPS, accelerometer, etc.). This perceptual capabilities can be built-in the device itself, in the surrounding environment, or in another device that shares this context over a network. Technology designed under this approach can therefore benefit at the output level by:

- Adapt outputs to a current situation (i.e. volume, brightness, privacy settings, etc.)
- Find the most suitable time for interruption
- Reduce the need for interruption

And at the input level:

- Adapting the input to the current situation (i.e. applying audio filters and recognition algorithms)

- Limiting the need for inputs (i.e. by looking to information that has already captured from the context of use)
- Reducing the selection space (by offering options tailored to the current context).

2.2.5. A domain-independent approach to implicit interaction

One of the main limitations of the above mentioned approaches to implicit interaction is the difficulty for generalizing and extrapolating the emergent designs to different contexts, as they are based on *domain specific* knowledge. In order to overcome this limitation, Ju and Leifer [2008] have proposed a framework that provides a *domain independent* design methodology for creating a broad class of interactions (both implicit and explicit) focusing on identifying when implicit interactions are useful. Instead of looking for situation-specific intelligence, this model proposes *interaction patterns* to help designers to *model behaviors* based on everyday interactions for reaching the appropriate user experience.

Ju & Leifer's model characterizes implicit interaction as communication without explicit input and/or output, both from the human and the system side. Following this approach, an interaction can be implicit if it occurs on the *background* of the user's attention (e.g. when the computer autosaves or backups files) or if the exchange is initiated by the system rather than by the user, like when the computer displays a sound to alert us about new email. In this case, although the output is clearly explicit (as it appeals to our attention) the interaction is based on an *implied* demand of information (new emails). This approach translates to a classification of interactions according to *attentional demand* and *initiative*, as shown in Figure 2.3. The former ranges from foreground interactions (which require user attention) to those background interactions that elude user attention. *Initiative*, on the other hand, refers to who (and till what extent) initiates the interaction: reactive interactions are initiated by the user, whereas proactive interactions are fostered by the computer system. Ju & Leifer's framework models implicit interaction not only by considering implicit inputs, but also implicit outputs; something that is generally overlooked by the approaches exposed before. The implicit does not depends exclusively in the inputs generated by user and the context, also the system can promote implicit interaction through its outputs, in a reactive or proactive fashion.

All the models presented before also tapped *attention* as one of the main components of implicit interaction. But by introducing the dimension of *initiative*, Ju & Leifer decouple attention from intention. *Initiative* basically determines who initiates an interaction and

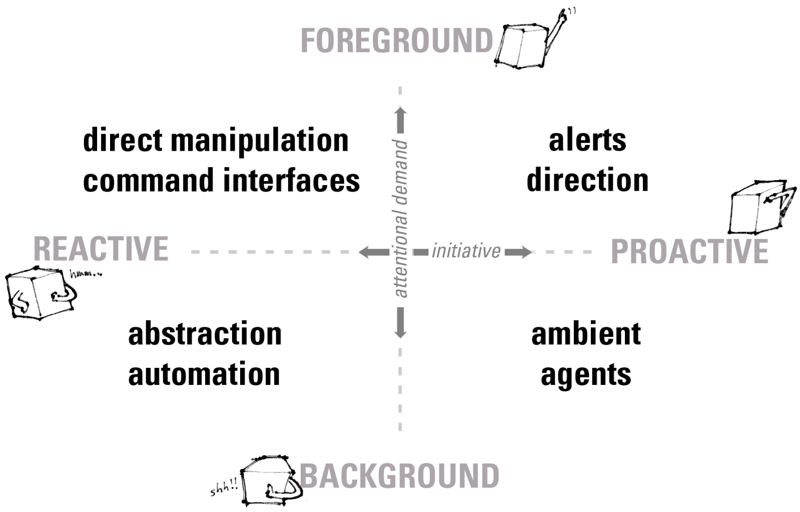


Figure 2.3.: Ju & Leifer's implicit interaction framework showing the range of interactive system behaviors [2008]

how, and this is a fundamental aspect in every day human-human interaction, when working together, when engaging in activities in which users have to be coordinated for succeeding, etc. And as *attention* operates under perceptual focalization, *initiative* requires proactivity. This bring us directly to Tennenhouse proactive computing (see Figure 2.2) and his assumption of proactivity operating under presumption. Proactive computer systems that operate without constant user supervision require a presumption of the world. That is why Tennenhouse, Weiser and Schmidt were mainly concerned with issues such as sensing, context modeling and data aggregation. But Ju & Leifer argue that proactive interaction cannot rely in technology alone. By referring to the (in)famous case of *Clippy*, the Microsoft Office Helper agent, Ju & Leifer expose the problem of using technologies that lack *facework* [Goffman, 2005], the social training required to accurately guess what the users want or what are they trying to do. *Initiative* therefore accounts for this social aspect by describing a mixed interplay between reactive and proactive actions, from both the user and the system, to foster a greater understanding of the situation of interaction, as it naturally happens in daily life.

This framework should therefore be seen as a tool for problem representation ("less as a hammer, more as a lens" in words of the authors) that helps designers to consider implicit interactions as a possible design solution by focusing on human-human interactions and translating them to HCI.

2.3. Implicit interaction through physiological computing: expanding context with physicality

The above mentioned approaches to implicit interaction were majorly motivated by the gap between traditional HCI (human-centric, command-dependent, supervised) and contemporary interaction scenarios characterized by mobility, pervasiveness, ubiquity, and concurrently happening with daily life activities. This situation calls for more sophisticated interaction designs that leverage on the implicit component of human-human interaction for coping with complexity, without the user constantly being *in-the-loop*. And while such approaches bring valuable information on how to promote the implicit though contextual intelligence or behavior modeling (of both humans and machines) they do not directly discuss one of the main sources of implicit information in human communication: the body.

For researchers working in the field of physiological computing (PhyComp), to look into the human body for depicting implicit perceptive, affective and cognitive states was the way to fill in this gap. As stated by Hettinger and colleagues, the standard mode of HCI is *asymmetrical* with respect to information exchange: the computer is capable of providing great amounts of information regarding the system state and the context of interaction, but very few about the user psychological state (motivations, emotions) is available for the system [Hettinger et al., 2003]. Therefore, there is little space to leverage on the implicit nuances of the human body to enrich and adapt HCI. To overcome this asymmetry, PhyComp aims at expanding HCI by monitoring, analyzing, and responding to user's physiological activity in real time [Allanson and Fairclough, 2004]. In these type of interactive systems, physiological data (also known by the summarizing term of *biosignals*) are used as an input for control, even without requiring overt response from the user. Through this approach, PhyComp is capable of capturing spontaneous and subconscious facets of user state (from affective responses to attention and cognitive load) to augment the bandwidth of HCI, opening new implicit communication channels with computer systems [Hettinger et al., 2003].

Differently from the design approaches presented in section (2.2), PhyComp provides *user's context* to the system [Fairclough, 2009] that could seamlessly encompass the *situational context* of interaction, as defined by Schmidt (see 2.2.4). In this regard, PhyComp not only is an excellent candidate to foster adaptive and proactive behaviors from the system side, but also for assessing and quantifying the impact of the system's adaptive response on the user. This is what Fairclough calls the *reflexive quality* of PhyComp, as the system not only can fine-tune responses according the the preferences and states of the user, but it can also *learn* about the user's response in the same way as it can learn form her/his physical or digital context. Similar to the aforementioned design approaches to implicit interaction, PhyComp seeks for a seamless symbiotic relationship between user and system, one that reliefs human for constant overt control while empowering the system for adapting autonomously and meaningfully.

Whereas PhyComp has been mainly applied in fields such as rehabilitation, neuroscience, and prosthesis control, the recent emergence of wearable non-invasive physiological sensors have promoted the use of PhyComp in daily life environments. *Out of the lab* PhyComp technology is nowadays more common than ever, and it is rapidly taking on the market in the shape of activity trackers (both for wellness like the Jawbone[1] and professional sports like the Garmin Edge line[2]), EEG based cognitive trainers (like the NeuroSky MindWave headsets[3]) and alternative ways of communication (the Apple Watch hear-rate sharing system[4] being one of the most populars).

To better understand how PhyComp works and leads to different types of interaction paradigms (both implicit and explicit), we provide an overview on fundamental concepts such as biosignals, sensing techniques, kinds of PhyComp systems that can be developed from them, and methods for displaying physiological data in real time.

2.3.1. What is a biosignal?

Within the scope of biomedical sensing, a biosignal can be defined as a description of a physiological phenomenon, irrespective of the nature of this description [Kaniusas, 2012]. Given the broad sense of this definition, variety of biosignals can span from visual inspection up to signals recorded from the human body using different types of sensors, like electrocardiography (ECG) for heart rate activity, electroencephalography (EEG) for brain cortical activity, or electromyography (EMG) for muscle activity. A good example for understanding the nature of a biosignal - from its generation till its registration- are the the acoustic biosignals used for the assessment of cardiorespiratory pathologies, and the mechanism for assessing them: the stethoscope (see Figure 2.4). In the heart, the biosignal *source* is the periodic closure of heart valves, which conveys sound. Additionally, the lungs also generate sounds by air turbulences in the branching airways of the lung, whereas the snoring sounds arise in the upper airways, due to elastic

[1]https://jawbone.com/ (accessed on November, 2015)
[2]https://buy.garmin.com/en-US/US/into-sports/cycling/edge-520/prod166370.html (accessed on November, 2015).
[3]http://store.neurosky.com/collections/eeg-headsets (accessed on November, 2015)
[4]https://www.apple.com/watch/new-ways-to-connect/ (accessed on November, 2015).

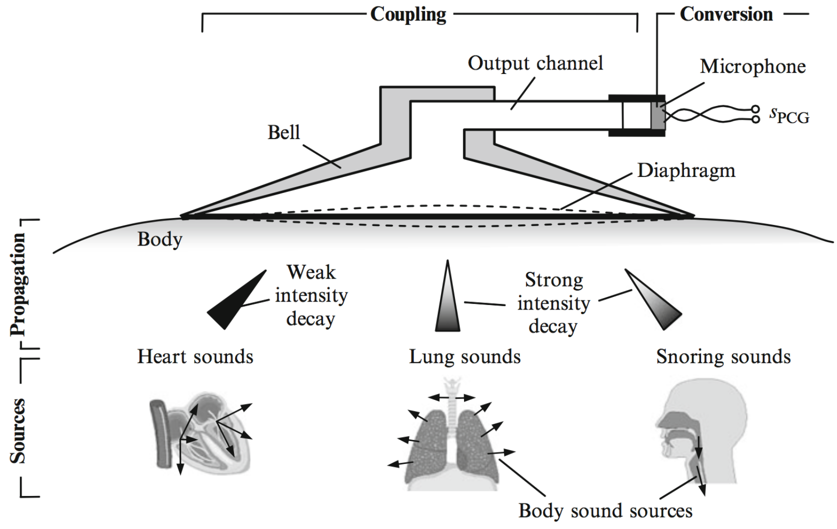


Figure 2.4.: The stethoscope and the registration of body sounds. The sources of the acoustic biosignals are depicted, along biosignal's propagation, coupling, and conversion for later registration [Kaniusas, 2012]

oscillations in the pharyngeal walls [Kaniusas, 2012].

The sounds *propagate* throughout the tissue and undergo attenuation due to increasing distance from the source and damping by the medium itself. As indicated in Figure 2.4 by intensity decay, the attenuation is different for different sounds, since their spectral components differ: the attenuation is less for the heart sounds than for the lung and snoring sounds, since the latter sounds exhibit more high-frequency components facing a stronger damping. The *coupling* (and amplification) of sounds is performed by the stethoscope chest piece, with an oscillating diaphragm and a resonating volume. Finally, the *conversion* of the acoustical pressure vibrations into an electric signal is carried out by an electroacoustic transducer (a microphone).

The principle behavior in the formation of an arbitrary biosignal can be modeled as an equivalent circuit as shown in Figure 2.5. The source of the biosignal is represented by a sinusoidal voltage source $u(t) = U \cdot \cos(\omega t + \varphi_U)$ with complex amplitude

$$\underline{U} = U \cdot e^{j\varphi U} \tag{2.1}$$

magnitude U , angular frequency $\omega(= 2\pi \cdot f$ with f as oscillating frequency), and phase φ_U , satisfying $u(t) = Re[\underline{U} \cdot e^{j\omega t}]$. The propagation losses are represented by a series impedance

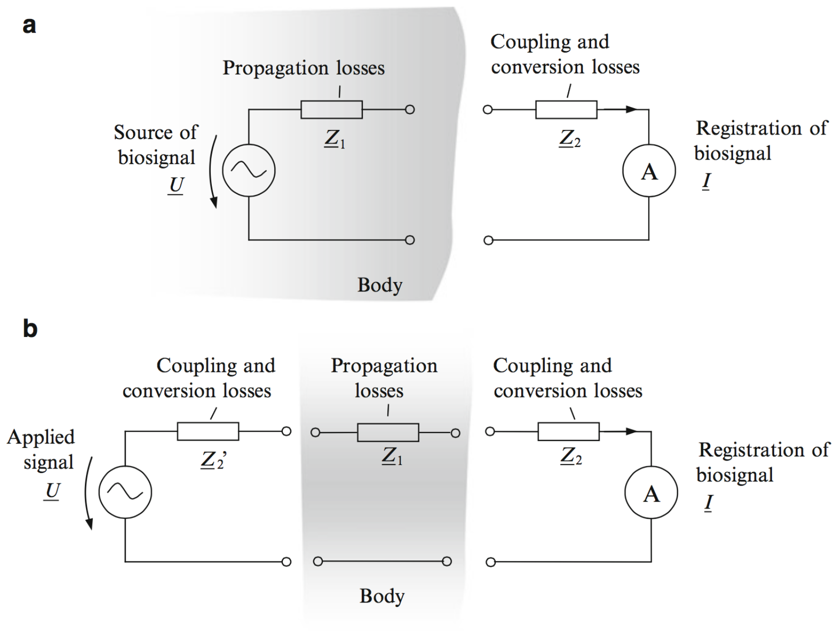


Figure 2.5.: Model of biosignal generation, propagation, coupling, and registration. (a) Permanent biosignal. (b) Induced biosignal [Kaniusas, 2012]

$$\underline{Z_1} = Z_1 \cdot e^{j\varphi 1} \tag{2.2}$$

the coupling and conversion losses by another series impedance

$$Z_2 = Z_1 \cdot e^{j\varphi 2} \tag{2.3}$$

and the registered biosignal by the resulting current $i(t) = I \cdot \cos(\omega t + \varphi_I$ with complex amplitude

$$\underline{I} = I \cdot e^{j\varphi I} \tag{2.4}$$

satisfying $i(t) = Re[\underline{I} \cdot e^{j\omega t}]$. According to Ohm's law

$$\underline{I} = \frac{\underline{U}}{\underline{Z_1} + \underline{Z_2}} \tag{2.5}$$

In other words, the higher the losses, e.g., the magnitudes $Z_1 (\neq 0)$ and of $Z_2 (\neq 0)$ usually capacitive-resistive losses, the weaker the registered biosignal will be, i.e., the magnitude I. In general, $\varphi_1 \neq \varphi_2$ provided that $\varphi_1 \neq 0$ or $\varphi_2 \neq 0$; likewise, if all losses can be modeled by real resistances then $\varphi_1 = \varphi_U$ and $I = U/(Z_1 + Z_2)$. It should be noted that physiological phenomena of interest are hidden not only in $\underline{U}$ but also in $\underline{Z_1}$, for the propagation may influence the resulting $\underline{I}$ in a significant and even advantageous way.

If the acoustic biosignal showed in Figure 2.4 is considered in the light of the above model (Figure 2.5), the temporal behavior of an acoustical source can be described by $u(t)$ and its intensity by U. The strength of the propagation losses of the body sounds can be given as Z_1 (2.2) while the capacitive behavior of the propagating medium can be described by the corresponding phase angle $\varphi_1(\neq 0)$. Alternatively, the strength of the coupling and conversion losses in the acoustical sensor can be defined as Z_2, whereas the corresponding $\varphi_2(\neq 0)$ can describe the time delay in the chest piece and the conversion delay in the microphone (2.3). The output $S_{PCG}(t)$ of the microphone—as schematically shown later in Figure 2.8—corresponds then to $i(t)$.

While Figure 2.5(a) applies to biosignals with their source already inside the body, Figure 2.5(b) depicts a model of an induced biosignal. Here, the biosignal is generated outside the body with an artificial signal source with its complex amplitude U. After coupling and conversion losses $\underline{Z'}_2$ on the input side, the induced signal undergoes propagating losses $\underline{Z}_1$ in the body, which are modulated by a physiologic phenomena of interest. On the output side, the coupling and conversion losses $\underline{Z}_2$ co-determine the resulting induced biosignal $\underline{I}$ according to

$$\underline{I} = \frac{\underline{U}}{\underline{Z}_1 + \underline{Z}_2 + \underline{Z}_{2'}} \tag{2.6}$$

To give an example, $\underline{U}$ could characterize an incident artificial light beam coupled into a finger, whereas $\underline{Z}_1$ varies by the changing light absorption due to pulsating blood volume [Kaniusas, 2012]. Since blood pulsations carry cardiac and respiratory information, the transmitted light characterized by $\underline{I}$ reflects cardio-respiratory activity, as depicted later in Figure 2.8(c).

Sensing techniques

There are a number of sensing techniques commonly used for acquiring physiological data. Following we summarize the most common ones within PhyComp (and the most relevant for this dissertation).

Electrodermal activity (EDA) Also known as galvanic skin response, refers to the change of the skin's electrical conductance properties caused by stress and/or changes in emotional states [Boucsein, 2012]. It reflects the activity of sweat glands and the changes in the sympathetic nervous system, being a direct indicator of overall arousal state. The signal is normally measured at the palm of the hands or the soles of the feet using two electrodes between which a small, fixed voltage is applied and measured. Changes in the skin's resistance are caused by activity of the sweat glands. In this way, when a person is presented with a stress-inducing stimulus; his/her skin conductivity will increase as the perspiratory glands secrete more sweat. The EDA signal is easy to measure and reliable; it is one of the main components of the original polygraph or *lie detector* [Clark and Tifft, 1966], and is one of the most common signals used in both psychophysiological research and in the field of affective computing [Picard and Picard, 1997].

Electrocardiogram (ECG) The ECG is a measurement of the electrical activity of the heart as it progresses through the stages of contraction [Haag et al., 2004]. Figure 2.6 shows the components of an ideal ECG signal. In HCI systems for non-clinical applications, the heart rate (HR) and heart rate variability (HRV) are the most common features measured. For example, low and high HRs can be indicative of physical effort. In affective computing research, if physical activity is constant, a low HRV is commonly correlated to a state of relaxation, whereas an increased HRV is common to states of stress or anxiety [Haag et al., 2004].

Electrooculogram (EOG) EOG is the measurement of the Corneal-Retinal Potentials (CRP) across the eye using electrodes. In most cases, electrodes are placed in pairs to the sides or above/below the eyes. The EOG is traditionally used in HCI to assess eye-gaze and is normally used for interaction and communication by people that suffer from physical impairments that hinder their motor skills [Knapp and Lusted, 1990].

Electromyogram (EMG) Electromyography is a method for measuring the electrical signal that activates the contraction of muscle tissue [Kaniusas, 2012]. It measures the isometric muscle activity generated by the firing of motor neurons. Motor Unit Action Potentials (MUAPs) are the individual components of the EMG signal that regulate our ability to control the skeletal muscles. Figure 2.7 illustrates a typical EMG signal and its amplitude envelope.

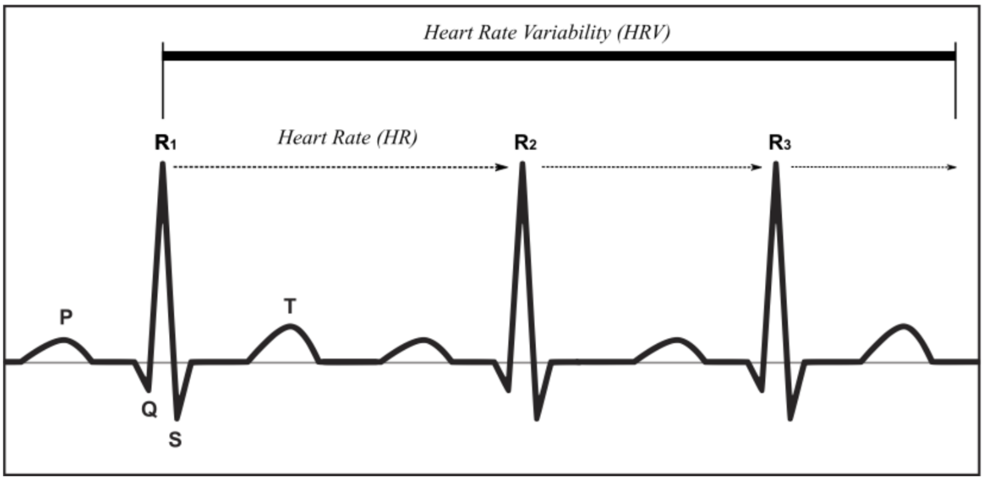


Figure 2.6.: Components of an ideal ECG signal, depicting heart rate (HR) and heart rate variability (HRV). *P wave* indicates atrial depolarization, or contraction of the atrium; the *QRS complex* indicates ventricular depolarization, or contraction of the ventricles (source: http://www.rnceus.com/ekg/ekgnorm.html accessed on September, 2015).

EMG-based interfaces can recognize motionless gestures across users with different muscle volumes without calibration, measuring only overall muscular tension regardless of movement or specific coordinated gestures [Caramiaux et al., 2015b]. They are commonly used in the fields of prosthesis control and functional neuromuscular stimulation. For HCI applications, EMG-driven interfaces have traditionally been used as continuous controllers, mapping amplitude envelope to control various system parameters [Tanaka, 2000].

Electroencephalogram (EEG) The Electroencephalogram monitors the electrical activity caused by the firing of cortical neurons across the brain's surface [Kropotov, 2010]. In 1924, German neurologist Hans Berger measured these electrical signals in the human brain for the first time and provided the first systematic description of what he called the electroencephalogram. In his research, Berger noticed spontaneous oscillations in the EEG signals, and identified rhythmic changes that varied as the subject shifted his/her state of consciousness. These variations, which would later be given the name of alpha waves, were originally known as Berger rhythms [Berger, 1934].

Brainwaves are complex signals. In non-invasive EEG monitoring, any given electrode picks up waves pertaining to a large number of firing neurons, each with different characteristics indicating different processes in the brain. The resulting large amount of

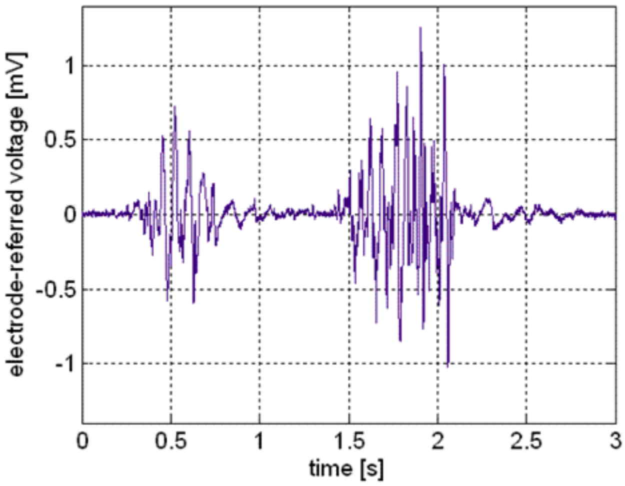


Figure 2.7.: EMG signal recorded with Ag/AgCl electrodes over bicep during two brief muscle contractions. The amplifiers were configured with a lower cut-off frequency of 10 Hz and an upper cutoff frequency of 1 kHz (source: http://www.intantech.com/signals_RHA2000.html accessed on September, 2015)

data that represents brain activity creates a difficult job for physicians and researchers attempting to extract meaningful information. Brainwaves have been categorized into four basic groups or bands of activity related to frequency content in the signals: Alpha, Beta, Theta and Delta [Kropotov, 2010].

- *Delta* waves are slow periodic oscillations in the brain that lie within the range of 0.5 to $4Hz$ and appear when the subject is in deep sleep or under the influence of anesthesia.
- *Theta* waves lie within the range of 4 to $7Hz$ and appear as consciousness slips toward drowsiness. It has been associated with access to unconscious material, creative inspiration and deep meditation.
- *Alpha rhythm* has a frequency range that lies between 8 and $12Hz$. Alpha waves have been thought to indicate both a relaxed awareness and the lack of a specific focus of attention. In holistic terms, it has been often described as a state of relaxation and awareness.
- *Beta* refers to all brainwave activity above $12Hz$ and is further subdivided into 3 categories:
  - Slow beta waves ($13 - 20Hz$) are the usual waking rhythms of the brain associated with active thinking, active attention, focus on the outside world

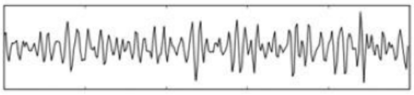
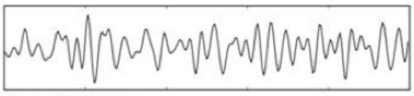
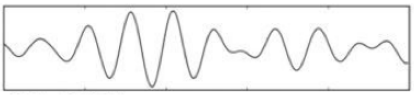
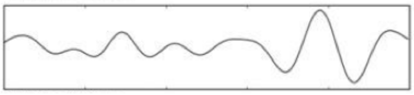
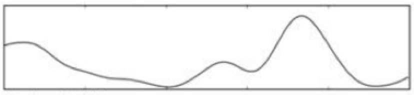
Name	Frequency	Signal
Gamma	30-100+ Hz	
Beta	12-30 Hz	
Alpha	8-12 Hz	
Theta	4-7 Hz	
Delta	0-4 Hz	

Table 2.1.: EEG frequency bands.

or solving concrete problems

- Medium beta waves ($20 - 30Hz$) occur when the subject is undertaking complex cognitive tasks, such as making logical conclusions, calculations, observations or insights.
- Fast beta waves (Over $30Hz$) often called Gamma, is defined as a state of hyper-alertness, stress and anxiety. It is found when performing a reaction-time motor task

Table 2.1 shows each of the frequency bands as displayed by an EEG monitoring system. This categorization however, is the source of certain controversy as some researchers recognize up to six different frequency bands [Väljamäe et al., 2013b].

Classification of biosignals

Given the great variety of biosignals, there is no an unique way for classifying them. In this regard, Kaniusas has proposed three methods of classification based on *existence, dynamic nature,* and *origin* of the biosignal. Figure 2.8 offers a description of these methods.

Regarding their existence, *permanent* biosignals are those that exist without the need of any artificial trigger or excitation from outside the body, as their source is the body

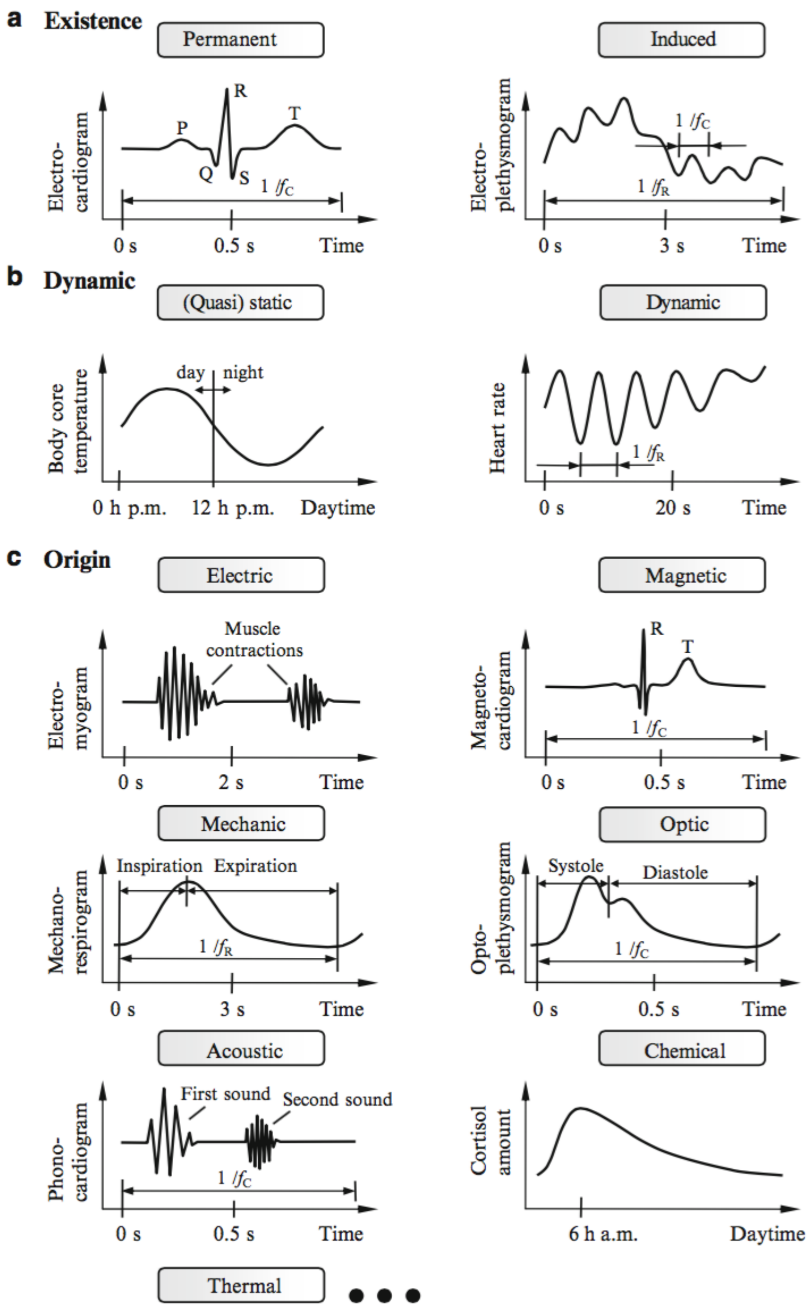


Figure 2.8.: Kaniusas' biosignal classification method based on (a)existence (b)dynamic and (c) origin, with indicated heart rate f_C and respiratory rate f_R [2012]

itself and are available at any time (see Figure 2.5a). For instance, an electrocardiographic signal (ECG) induced by electrical heart muscle excitation with the typical peaks P–Q–R–S–T (Figure 2.8a) and the aforementioned acoustic biosignal induced by the consecutive heart valve closures and captured by the stethoscope, with the typical first and second heart sounds (Figure 2.8c) belong to the group of permanent biosignals. *Induced* biosignals, on the other hand, are artificially triggered, excited, or induced (see Figure 2.5b).

In contrast to permanent biosignals, induced biosignals exist roughly for the duration of the excitation. As soon as the artificial impact is over, the induced biosignal decays with a certain time constant determined by the body properties. The interaction of the tissue with the induced stimulus, irrespective of the stimulus nature, is then recorded as an induced biosignal. A corresponding example could be given by electric plethysmography, in which an artificial current is induced in the tissue and a voltage along the current path reflects tissue impedance changes. The voltage is then registered as an induced biosignal (electroplethysmogram) with discernible cardiac and respiratory components (see Figure 2.8a). Alternatively, optical oximetry uses artificially induced light while the transmitted light intensity is mainly governed by light absorption through local pulsatile blood volume. The transmitted light is detected as an induced biosignal, showing a steep systolic increase and a slow diastolic decrease (see Figure 2.8c). In general, the origin of the induced stimulus, e.g., magnetic field from coils above the head for magnetic stimulation, may be different from that of the registered biosignal, e.g., generated electric potentials from electrodes on the head.

Whereas Kaniusas' classification is mainly based on the physiological nature of the biosignals themselves and from a biomedical standpoint, other PhyComp practitioners attempt to classify them according to high-level affordances to describe their possible integration into computer systems and user interfaces. In this regard, Hugo Silva [2015] proposes a classification based on:

1. *Controllability*: the degree of control that the user has over the source. Three types of control are considered, namely: voluntary, indirect (or mixed), and involuntary.
2. *Acceptability*: the extent to which the interface will be cause excessive disruption or annoyance to the user. For this property three classes are considered, namely: invasive, wearable, off-the person.
3. *Observability*: the degree to which data can be recurrently acquired from a given source within a certain period of time if needed. For this property the following classes can be considered: pervasive, momentary, and controlled.

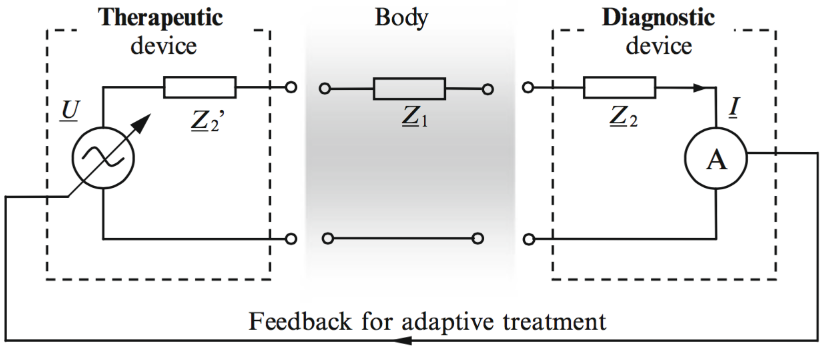


Figure 2.9.: Diagnostic and therapeutic application of biosignals Kaniusas [2012]

4. *Stability*: Even if a biosignal source is highly controllable, acceptable, and observable, another property that needs to be taken into account in the design of HCI applications is how stable is the source over time. Sources can be characterized as persistent, repeatable, and sporadic.

This model is particularly useful to characterize different types of biosignals in a PhyComp context, where they act as input for interactive systems.

2.3.2. Physiological computing systems

Historically, biosignals were mainly used within medical practice for diagnosis and therapy. The former is concerned with the assessment of health status through biomedical sensing, whereas therapy utilizes biosignals as an objective feedback for selecting appropriate therapeutic measures, continuously monitoring their impact, and improving their efficiency. As shown in Figure 2.9, the biosignal registered by a diagnostic device and presented by $\underline{I}$ controls a therapeutic device by adjusting its stimulus given by $\underline{U}$.

However, as biomedical sensing evolved toward portable, wearable, and inexpensive equipment, it expanded beyond its traditional application domains. This is how computer scientists and HCI practitioners embraced physiological data as a type of input for interactive systems. As mentioned by Fairclough and Gilleade, many types of interactions can be facilitated by biosignal input, ranging from intentional control to implicit software adaptation [Fairclough and Gilleade, 2014b]. It represents a new type of human-computer interaction that directly interfaces with the human brain and body, and as it is capable of offering quantitative correlates of the user affective, cognitive and perceptual states, it has opened a door for a wide range of implicit and explicit interactions.

Application domains and types of PhyComp systems

Beyond traditional diagnosis and therapy applications, other PhyComp systems have been designed for a wide range of purposes, like improving performance or maximize positive user experiences by looking to cognitive and/or affective aspects of the user behavior. For instance, Wilson and colleagues used EEG in combination with ECG and respiration rate to characterize the mental workload of users. By means of an artificial neural net, they categorized the level of mental workload to automate elements of the undergoing task when the operator was mentally overloaded [Wilson and Russell, 2007, Pope et al., 1995]. This study reported substantial improvements of performance when adaptive automation was controlled through physiological responses. Affective computing, on the other hand, has leveraged on PhyComp for exploring the psychophysiology correlate of emotions in laboratory conditions, inducing positive emotions (as happiness or surprise) and negative (as anger and sadness) [Picard and Picard, 1997, Scheirer et al., 2002]. In this line, PhyComp has been successfully used to evaluate user experience in scenarios such as video games by looking at cognitive-emotional responses [Mandryk and Atkins, 2007], but also using these responses for altering elements of the game in real time to improve the experience or prevent frustration [Nacke et al., 2011].

Although biosignals can be use for designing a great variety of user interfaces, whose interaction can range from overtly explicit to covertly implicit (without being mutually exclusive), Fairclough and Gilleade [2014a] proposed to distribute PhyComp systems in three main categories:

Input control A PhyComp system designed to communicate *intentional actions* to a computer program by transforming physiological activity (e.g. in the cortex, muscles or eye moment) in discrete or continuous input commands, as if using a keyboard, mouse, touchscreen on gesture-based interface. Besides monitoring and therapy, direct input control was one of the first application domains of PhyComp, mainly developed for users with physical disabilities in order to allow them to use desktop software (i.e. word processors) [Li et al., 2008] or for controlling a robotic prosthesis [Bitzer and van der Smagt, 2006]. One salient example in this regard are Brain Computer Interfaces (BCI) [Wolpaw et al., 2002a] which offer an alternative way of input control, but rather than tapping the final psychomotor stage (i.e. the muscle) BCIs are designed to capture electro cortical activity at source: the intention that precedes movement, the spark of activation in response to a particular stimulus, the localization of visual attention, etc. In this manner, BCIs offer hands-free interaction capable of communicating with conventional

GUI-based technologies as well as prostheses. It is important to mention that BCIs, like muscle controlled interfaces, are still a direct form of input control designed to *emulate* standard devices such as the keyboard and mouse. As mentioned by Fairclough and Gilleade, issues like novelty, ease of use, bandwidth and speed play an utter role for end-user acceptance (especially in the case of healthy users) [2014a].

Biocybernetic adaptation Inspired by Pope's concept of biocybernetic loop [1995] these systems monitor spontaneous activity from the brain and the body in order to capture psychological states linked to user's emotion, cognition and perception. This information is later used to inform adaptive systems that react to user states (e.g. fatigue) to compensate such states (e.g. reducing anxiety) by providing a tailored feedback. This type of biocybernetic adaptation encompasses a wide array of software applications, with implicit interaction among the most interesting use cases, as HCI practitioners can leverage on the continuos signals of the human body to create adaptive interfaces that *understand* and *respond* to the user in a highly personalized manner. As stated by Fairclough and Gilleade "regardless of precise context, biocybernetic systems are fundamentally designed to deliver software adaptations that will be perceived as timely, intuitive and *intelligent* by the user" [Fairclough and Gilleade, 2014a, p4].

Ambulatory monitoring This category is related to the *surveillance function* of PhyComp, as its goal is feedback and data visualization -for the individual and other connected people- rather than real-time adaptive control. Biosignal monitoring thus cannot be consider an interactive modality by itself. However, it is reasonable to assume that PhyComp systems will rely more and more on lightweight and unobtrusive *wearable* sensors, as in the case of smartwatches. Users may therefore carry these devices unobtrusively every day, and the flow of physiological data from person to system is the lifeblood of all PhyComp. These data (which is mostly acquired implicitly and without user's direct intention) drive the algorithms used to facilitate software adaptation, but they may be recorded for other purposes as well. Obvious candidates for continuous physiological monitoring are users with chronic health problems who are being treated as out-patients. Basic autonomic functions, such as heart rate, blood pressure and respiration patterns, could be recorded wirelessly and made available to qualified medical staff who wish to monitor those individuals outside of a medical facility. Alternatively, a social network of carers, close friends and family members may be granted access to real-time data feeds from patients for purposes of monitoring or reassurance [Fairclough and Gilleade, 2014b]. However, this approach can be useful for healthy users as well, which are engaged in activity tracking and quantified self communities.

As described in 1.2, these three categories proposed by Fairclough and Gilleade are intended to represent a continuum rather than a hard distinction between different types of systems (see 1.1). In fact, all these types of PhyComp are already being used in combination with conventional modes of input control (as in the case of the MYO[5] armband, that combines EMG and IMU for controlling interactive systems and applications). It is therefore easy to envision implicit adaptive PhyComp systems seamlessly complementing and enhancing conventional input control. They will not require extra attentional or motor effort from users, but rather enhance their control bandwidth without penalizing explicit input methods.

Implicit interaction and PhyComp

Although most PhyComp literature do not tackle directly the concept of *implicit interaction*, their goals and motivations are indeed very similar to the design approaches exposed in Section 2.2: enhancing conventional modes of HCI, not to replace them; creating more autonomous and proactive systems that will not require constant input from the user for knowing and learning about what she/he wants; systems that will foster more sophisticated interactions that will seamlessly transit from the background to the foreground of our attention, and back. The aforementioned design approaches looked at *device intelligence*, *situational context* and everyday behavior. What PhyComp brings to the table is *user context,* mechanisms to implicitly inform systems about the user perceptive, emotive and cognitive state at the moment in which the interaction happens, with the chance of feeding back a highly tailored response, making systems even more proactive.

For incorporating implicit user information, PhyComp systems have to passively monitor psychophysiological changes in order to inform interface adaptations in real-time. Physiological data has to be autonomously collected as the user performs a related or unrelated task, and the system has to subsequently use this information to activate implicit or explicit software adaptations if certain triggering conditions are met. Clearly, these systems would operate outside the direct, intentional control of the user. The system might have, however, a specific agenda (e.g. to achieve a specific target state in terms of human performance or psychological state) [Fairclough and Gilleade, 2014b] so feedback is designed to explicitly or implicitly influence the psychophysiology of the user in order to establish/sustain a desired state.

[5] `https://www.myo.com` (accessed on November, 2015).

One of the earliest biocybernetic adaptive systems was developed by NASA to be used in flight simulators [Pope et al., 1995]. The physiology of the pilot (i.e. EEG) was monitored in order to manage the status of an auto-pilot facility during flight time. The agenda of the system was to sustain the level of alertness of the pilot at an optimal level via manipulation of the auto-pilot status (i.e. alertness), which tended to decline during auto-pilot activation and to increase when the pilot manually controlled the craft. Computer games, on the other hand, are designed for a particular skill set that may not accurately reflect the skill set of the individual player [Gilleade and Dix, 2004]. There are multiple measures of cognitive workload (e.g. frontal theta [Klimesch, 1999]), which can be used to infer perceived difficulty during game play. In this manner, an implicit PhyComp system can make use of these measures to dynamically adjust the level of difficulty in order to match the ability of the player in real-time.

The concept of *biofeedback loop* is at the core of several PhyComp systems. It refers to the process that enables an individual to learn about her/his physiological state by means of perceptualization techniques such as visualizations, sonic displays, or haptic feedback [Kropotov, 2010]. The correct perception of such sates permits the individual to change his/her physiological activity with the purposes of improving a given experience or performance[6]. Thus, feedback strategies are of utter importance for designing adaptive PhyComp systems. As in this dissertation sound will be the main medium for displaying implicit physiological data, following we provide a summary of sonic interaction design applied to biosignals and PhyComp for both aesthetic (i.e. musical) and representational purposes.

2.4. Sonic interaction design

From Edison's phonograph to the mp3 and the current boom of streaming services, sound and music has been a driving factor for technology [Serra et al., 2007]. In this context, sonic interaction design (SID) allow us to think about sound -not only music- as one of the main design dimensions of the environments in which we live and work [Rocchesso et al., 2008], overcoming the sound-as-noise cultural barrier and promoting a sound-as-information attitude. SID thus can be defined as an interdisciplinary field of research and practice that explores ways in which sound can be used to convey information, meaning, aesthetic and emotional qualities in interactive contexts [Franinović and Serafin, 2013]. This trend has existed within the performing arts and computer science for many years,

[6] Association for Applied Psychophysiology and Biofeedback: `http://www.aapb.org/i4a/pages/index.cfm?pageid=1`. Accessed on November, 2015.

as exposed by the New Interfaces for Musical Expression (NIME)[7] series of conferences or the International Community for Auditory Display (ICAD)[8].

Designing the *sonic appearance* of interactive systems is thus a main topic within HCI, which depends on knowledge that has grown in different interrelated fields of research and practice. As described by Rocchesso and colleagues:

> "the relevant areas of SID research have grown in significance, both as a result of the design needs associated to improving the sonic jungles we are increasingly confronted with, and as economies of scale and miniaturization have contributed to a widening array of interactive artifacts and systems that are embedded with ever more sophisticated computing, sensing and actuating capabilities" [2008, p3970].

SID emerges from the desire to challenge traditional design approaches that limited sound as a functional, signaling, or iconic element of HCI, by considering sound as an active medium that can enable new experiences with and through interactive technology [Franinović and Serafin, 2013]. Therefore, the goal of SID is not only to expand the existing research on interactive sound, but also to promote new applications on a wide range of domains, such as interaction design, architecture, product and service design, the arts, and of course PhyComp. In the context of this dissertation two main approaches to SID (sonification and NIME) are explored to develop adaptive PhyComp systems that leverage on sound to foster implicit interaction. Whereas **sonification** mostly relies on representation-based design strategies, **NIME** is more concerned with the embodied and performance aspects of SID. This allow us to spam from low-level studies on the auditory perception of implicit user activity to high-level HCI approaches (digital musical instruments and music performance) for studying the impact of sound-based implicit PhyComp in the user experience. Below we offer a summary on both SID trends.

2.4.1. Sonification

Sonification describes "the use of non-speech audio to convey information (...) it refers to the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation" [Kramer et al., 2010]. In order to better define which acoustic renders can be considered as sonification, Grond and Berger [2011] proposed four conditions:

- The sound has to *reflect properties/relations* in the input data.

[7] `www.nime.org` (accessed on November, 2015).
[8] `http://www.icad.org` (accessed on November, 2015).

- The transformation has to be *systematic*. This means that there is a precise definition of how interactions and data cause the sound to change.
- The sonification has to be *reproducible*. Given the same data and identical interactions/triggers the resulting sound has to be structurally identical.
- The system can intentionally be *used with different data*, and also be used in repetition with the same data.

Sonification research has historically gravitated around topics such as the auditory display of datasets (interactive or otherwise), auditory feedback in computing displays, auditory icons, earcons, signaling, mobile communication and computing applications, bringing together interests from the areas of data mining, exploratory data analysis, human–computer interfaces, and computer music [Rocchesso et al., 2008]. By contrast with visualization, sonification inherently develops in time and exploits the fastest of human senses.

As exemplified by Hermann and Hunt [2005] the simplest auditory display conceptually speaking is the auditory *event marker*, a sound that's played to signal something (e.g. a telephone ringing, pressing a button in an ATM or smartphone). *Auditory icons* and *earcons* have been developed for this purpose [Barrass and Zehner, 2000] and are frequently used as direct feedback to an activity, and as their feedback usually consists of discrete events, they're rarely used to display larger data sets (which would require continuous and dynamic auditory renders).

Other sonification techniques such as *audification* are used to convert data series to samples of a sound signal (see Figure 2.10). Many of the resulting sounds are played back without interruption, and there's no interaction with the sound. However, the high number of acoustic attributes makes sonification a high-dimensional data display. Therefore, in almost every sonification a mapping strategy is applied. This is the driving feature of another type of sonification, *parameter mapping*, where data (or data-driven) features are mapped to acoustic attributes such as pitch, timbre or brilliance [Hermann and Hunt, 2005] (see Figure 2.10b).

Another strategy for examining data using sound is *model-based sonification* [Hermann and Ritter, 2005]. Whereas in other techniques data attributes relate to sound parameters, in model based sonification the data is used for a dynamic system setup, what Hermann and Ritter call a *virtual data-driven object*, or sonification model: "think, for instance, of data-driven points forming a solid capable of vibration. Excitation, achieved by the user interacting with the model, is required to move the system from its state of equilibrium. Damping and other energy loss mechanisms naturally cause the sonification

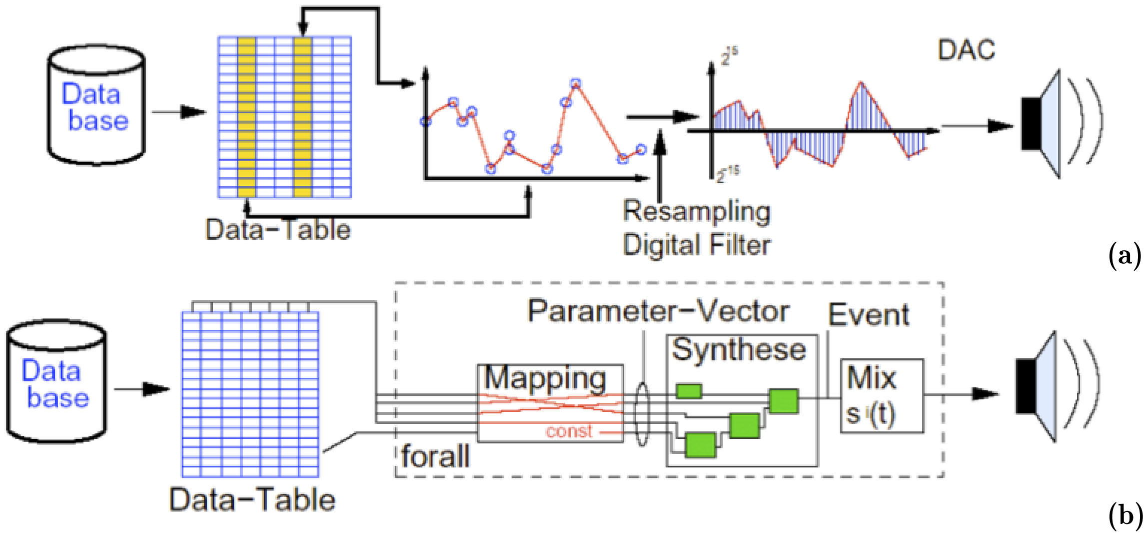


Figure 2.10.: Sonification techniques: (a) audification and (b) parameter mapping [Hermann, 2008]

to become silent without continuing interaction" [2005]. In this manner, interacting with sonification models has similar characteristics to interacting with physical objects such as musical instruments, and thus hopefully inheriting their advantageous properties.

As sonification techniques make more use of acoustic dimensions, or are applied to setup a dynamic model, they become critically dependent on interaction (moreover if the intention is to aid users to improve their experience or performance in a given activity). This approach has led to *interactive sonifications*. For instance, audification can be turn into an interactive sonification technique by letting the user move freely back and forth in the sound file, giving the user instantaneous and accurate control and portrayal of the signal characteristics at any desired point in the data set. Concerning parameter mapping, interactive control can play several roles: navigating through the data, adjusting the mapping on prerecorded data, or molding the sonification of data in real time. We can therefore increase the interactivity of sonification by including interactive controls and input devices to continuously move through the data set and control its transformation into sound.

In the same way as implicit interaction and PhyComp (see section 2.2) sonification has also evolved driven by the need of multimodal human computer interfaces for coping with complex HCI scenarios, where visual display may not be effective, or may not be an option. Think, for instance, in a user equipped with a mobile device in a daily life and mobile scenario [Gaye et al., 2006] (e.g. walking on the streets, working out, or cooking). In these cases, displaying information visually might not be the best option, as the user is clearly engaged in another activity and taking out the smartphone to check the screen can compromise her/his attention. In other scenarios such as driving a car, the inconvenience of relying on visual attention becomes even more evident.

As demonstrated by sonification research and practice, sonic displays offer great potentiality for encompassing multimodal interfaces and complement visual displays [Kramer et al., 2010, Hermann and Hunt, 2005]. Human perception is tuned to process complex combined experiences that involve the five senses, and that change instantaneously as we perform actions. Thus we can leverage on this human attribute by using different modalities for data representation, where sound plays a fundamental role. As mentioned by Hermann and Hunt, "the more we understand the interaction of different modalities in the context of human activity in the real world, the more we learn what conditions are best for using them to present and interact with high-dimensional data" [2005].

2.4.2. New interfaces for musical expression (NIME)

The domain of new musical instruments -both research and artistic practice- is also one of the main components of the SID interdisciplinarity. Unlike sonification approaches (which focus on reception based auditory studies) or interactive sonification techniques (whose behavior depends on the data under investigation and representation-based design strategies), the field of new musical instruments focuses on the embodied aspects of the human *perception-action loop* mediated by acoustic signals, where physical interaction yields continuous sonic feedback, but also tactile and visual feedback [Leman, 2008]. Scholars in this field have thus expanded the scope of SID by considering the human as something more than a receiver of auditory stimuli, leveraging on tangible, embedded and embodied computing for creating expressive HCI, consolidated in venues such as the New Interfaces for Musical Expression (NIME) and the Sound and Music Computing (SMC)[9] communities.

It is therefore natural to place the *performative* aspects of sonic interaction at the core of the NIME agenda. By combining the knowledge of fields such as interactive arts, electronic music, cognitive sciences, cultural studies and interaction design, NIME has developed a rich and ecological framework to explore new design principles that tightly connect audition, touch and movement, creating adapting system that sonically respond to the physicality of users to foster expression. Since the first NIME workshop in CHI

[9] `http://smcnetwork.org` (accessed on November, 2015).

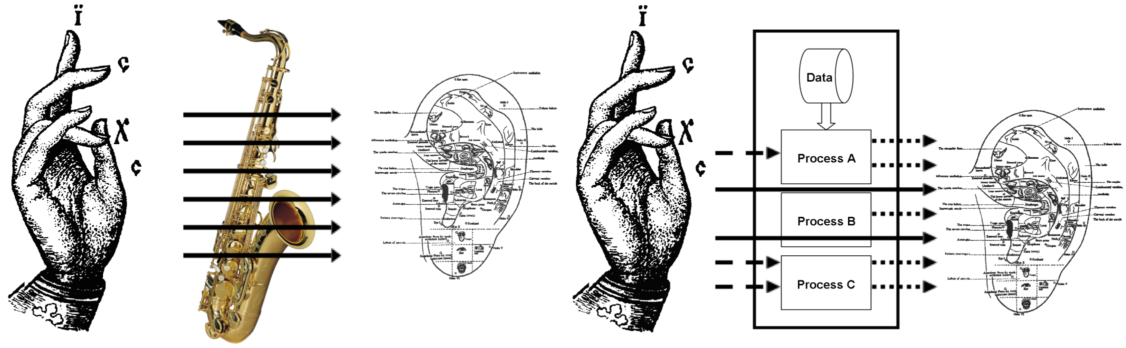
2001[10], the community has done a major contribution to HCI by transferring the embodied and performance aspects of musical instruments to interface design. As pointed out by Herman & Hunt:

> "The development of electronic musical instruments can shed light on the design process for human–machine interfaces. Producing an electronic instrument requires designing both the interface and its relationship to the sound source. This input-to-output mapping is a key attribute in determining the success of the interaction. In fact, Hunt, Paradis, and Wanderley 2003 have shown that the form of this mapping determines whether the users consider their machine to be an instrument (...) Another important aspect to consider is naturalness. In any interaction with the physical world, the resulting sound fed back to the user is natural in the sense that it reflects a coherent image of the temporal evolution of the physical system. " [Hermann and Hunt, 2005].

NIME leverages on the physical and performative aspects of acoustic instruments, and on the automation of computer systems to create interfaces that foster continuous and complex interactions, but where the user's attention is free to concentrate on higher goals rather than on streams of low-level control needed to operate the system. This differs from acoustic instruments, which require a continuous input to drive the sound source (thus continuously engaging the player in the interaction loop) provoking constant modulation of all the available sound parameters given the complex cross-couplings that takes place in the physical instrument. NIME also takes distance from GUI-based computer systems, which are mostly driven by choice-based inputs (menus, icons, etc) and rely on language or symbolic processing rather than physical interaction. Jordà has also discussed the *hybrid micro + macrocontrol* that can emerge from interactive music systems (see Figure 2.11):

> "(...) in traditional instruments the performer is responsible for controlling every smallest detail, leaving nothing to the instrument responsibility. In new interactive music instruments, the instrument's 'intelligence' may be partially responsible for one or more musical processes, the control of which may be entirely left to the instrument's responsibility or may be shared in different ways with the performer (...) Any combination of micro and macro-meta-control is now conceivable. Different layers or levels may even exist. The dialog (...) may be more or less present or implicit, but what is clear

[10] http://www.nime.org/2001/ (accessed on November, 2015).



(a) (b)

Figure 2.11.: (a) the traditional instrument microcontrol and (b) The 'interactive instrument' hybrid micro + macrocontrol by Jorda [2005]

> is that within this approach, instruments tend to surpass the one-action → one-event and the note-to-note playing paradigms, thus allowing them to work at different musical levels (from sound to form), and forcing performers to make higher level (i.e. more compositional) decisions on-the-fly" [Jorda, 2005].

To summarize, SID (and HCI in general) has gained from the NIME field not only guidelines and techniques that were traditionally found in the sound and music computing communities, but also valuable design principles and methodologies that guide HCI in the creation of socially meaningful, physically engaging and aesthetically pleasant sonic interactions through situated and explorative approaches [Franinović and Serafin, 2013]. Studies such as *Form Follows Sound* by Caramiaux and colleagues [2015a] and *Musical Metaphors for Interface Design by* Bau and colleagues [2008] are excellent examples of how action-sound relationships are exploited in the HCI domain, embracing techniques such as participatory design, workshopping and prototyping for understanding the natural and ecological connection that lay behind intuitive sonic interfaces. Sonification systems, on the other hand, can be also seen as *sonic instruments,* where acoustic behavior depends on the data under investigation, although they are designed for learning more about the data, rather than for driving musical expression.

2.5. Sonic interaction design applied to physiological computing

2.5.1. The pioneers

The idea of making biosignals audible began in parallel with visual body inspection, without the use of any instrument (e.g. using percussion or auscultation), and it moved towards tools for signal augmentation and registration [Kaniusas, 2012]. A concrete example is Laennec's stethoscope (1819) and its advantages for conducting the sounds generated inside the body. In the early 1930s. Prof. Edgar Adrian listened to his own EEG signal while replicating the first EEG experiments by Hans Berger [1934]. Indeed, sonification appeared to be well-suited for applications based on real time EEG, since sound can readily represent the complexity and fast temporal dynamics of brain signals.

Musicians and sound computing researchers pioneered several aspects of PhyComp, from biosignal sensing to processing and, more importantly, control and display paradigms. In fact, avant-garde musicians were among the first on exploring *biofeedback*[11] for developing adaptive electrical and electronic systems. In this regard, Rosenboom's book *"Extended Musical Interface with the Human Nervous System"* [1997a] represents a fundamental document about the history of physiology-driven sonic interaction. The document offers valuable insights on the early efforts of researchers and artists for bridging the gap between sound and human physiology, like the 1965 work *Music for Solo Performer* by Alvin Lucier, that is considered the first application of biofeedback in the arts, that achieved a direct mapping of a soloist's EEG alpha rhythms onto the orchestrational palette of a percussion ensemble (see Figure 2.12). A few years later, Teitelbaum developed *Spacecraft* (1967), *Organ Music* and *In Tune* [1976] to explore physiological multimodality by adding heart rate and breathing measures to EEG signals in the creation of electronic music textures.

In the 1970's, Rosenboom began his own research on physiology-driven music, under the hypothesis that it might be possible to detect aspects of the musical experience in cortical brain activity (EEG). In his attempt to go beyond direct unidimensional mappings (audification) he developed *The Performing Brain* and *Portable Gold and Philosophers' Stones* (Figure 2.13), in which he introduced a musical system whose parameters were driven by the EEG, galvanic skin response (GSR) and temperature of 4 performers. The EEG signal was processed to extract cortical correlates of human selective attention, although no study or empirical evidence corroborate this claim. Similar to Teitelbaum,

[11] Check Section 2.3.2 for a definition on biofeedback

Figure 2.12.: John Cage adjusting the electrodes on Alvin Lucier's head for a 1988 performance of *Music for Solo Performer* at Wesleyan University (source: http://www.statesofawareness.net)

Rosenboom worked on an audiovisual *happening* entitled *Ecology of the Skin,* that involved neurofeedback and heart rate from performers and the audience to drive musical textures.

2.5.2. Beyond biofeedback and towards implicit/explicit control

Before the 1990s most endeavors combining PhyComp and sound were based on biofeedback paradigms, and used brain waves amplitude (mostly alpha waves) or other direct parameter to drive arbitrary music compositions. But as biosensing technology evolved, musicians and sound designers explored the expressive possibilities of human physiology beyond biofeedback. They were among the first on developing physiology-based systems for direct control, as in the case of the *Music Activated Dance-Directed Music* (MADDM) [Gillett, 1985] where dancers' motion was captured through myoelectric signals (EMG) and classified for synthesized sound control. In 1997, Rosenboom presented a system using music generating rules, based on digital filtering and coherence analysis of EEG signals [Wu et al., 2009]. EMG pioneer Atau Tanaka (Figure 2.14) made an extensive use of the Biomuse system [Knapp and Lusted, 1990] for multimodal, multichannel control of

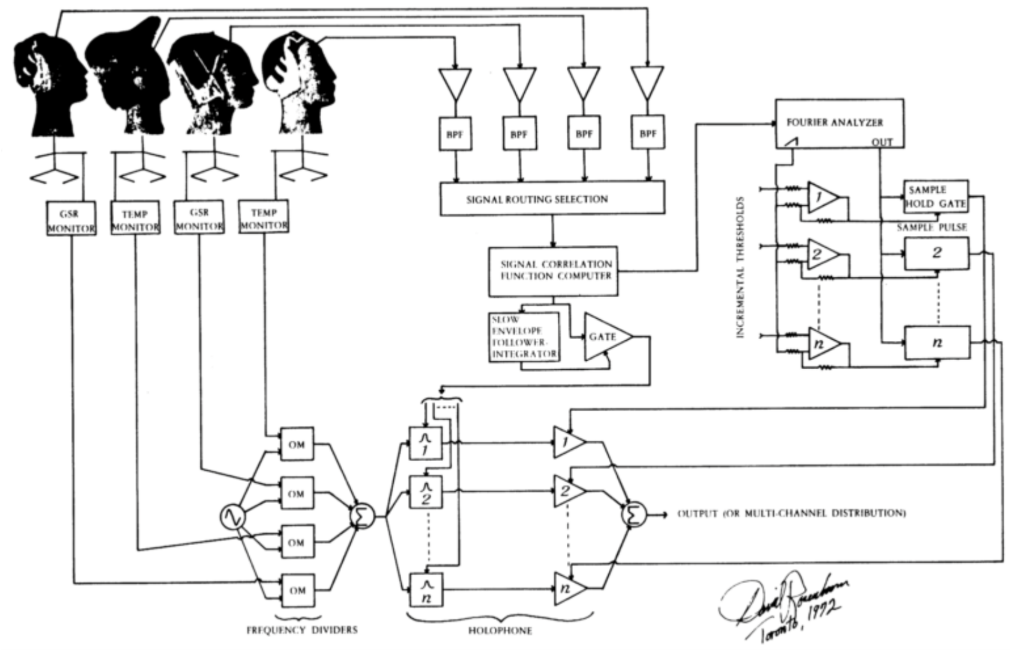


Figure 2.13.: System diagram of Rosenboom's *portable gold and philosophers' stones (Music from Brains in Four)* (1972). Score of a musical composition that includes measurement and analysis of EEG signals, GSR and body temperature changes from a quartet of performers Rosenboom [1997b, 24]

Figure 2.14.: Atau Tanaka performing at the Sonar Festival, Barcelona, Spain, June 1997 (Source: Daniel Langlois Foundation, http://www.fondation-langlois.org/html/e/page.php?NumPage=285)

electronic musical devices using both EMG and relative position sensing[12] [2000]. More recently, researcher and performance artist Marco Donnarumma [2012] also explored the affordances of direct control mappings in a series of works related to *biophysical music*[13] (Figure 2.15) using mechanomyography (MMG), the mechanical signal observable from the surface of a muscle when it is contracted (producing subsequent vibrations due to oscillations of the muscle fibers at the resonance frequency of the muscle).

The further development of Brain-Computer Interfaces (BCI) allowed the use of brain-driven technologies for music making, a topic of especial interest for this dissertation as the implicit PhyComp systems implemented in this thesis are based on brain activity. Following we describe the recent developments on this field.

Eduardo Miranda's *Brain-Computer Music Interfaces* (BCMI), as the *BCMI-Piano* [2006], looks for information in the EEG signal and match the findings with assigned generative music processes. The BCMI-Piano architecture is defined by 4 main modules, as shown in Figure 2.16: (1) Sensing: 7 pairs of EEG electrodes (bipolar montage) that sense the whole surface of the cortex. (2) Analysis: generates two streams of control parameters: (a) prominent frequency band in the signal, used by the music engine to

[12] A video of Atau Tanaka performing with the Biomuse can be found at http://vimeo.com/2483259 (accessed on November, 2015)

[13] A video of the piece *The Moving Forest* can be found at http://marcodonnarumma.com/works/the-moving-forest-act-1/ (accessed on November, 2015)

Figure 2.15.: Marco Donnarumma performing *The Moving Forest* at Transmediale.08, Berlin (2008). Source: http://marcodonnarumma.com/

generate two styles of sound, depending on whether the EEG indicates salient alpha levels ($8 - 13Hz$) or beta levels ($14 - 33Hz$); (b) complexity of the signal, using Hjorth signal complexity analysis. The music engine uses this information to control the tempo and the loudness of the music. (3) Music engine: contains the generative music rules. Each rule produces a musical bar or half-bar. (4) Performance module: plays the music using a MIDI-enable acoustic piano. It is important to highlight that Miranda's system *interprets* the meaning of the user's EEG instead of being explicitly controlled by the user. Still, the authors acknowledge the possibility of biofeedback with their system. The BCMI-Piano is programmed to search for information within the EEG signal and match what it finds to various generative musical processes in different musical styles. In terms of sonic display, spectral information in the EEG is used to activate the generative music rules, and the complexity of the signal is used to control the tempo of the music.

Miranda et al. also introduced neurogranular sampling, which is a sound synthesis method based on spiking neural networks (SNN) to control the triggering of sound grains from a certain sampled output.

The Georgia State University BrainLab has developed The Neural Music Software [Moore, 2003] to translate brain signal and brain-signal patterns directly to Musical Instrument Device Interface (MIDI), allowing a tonal representation of the signal. The software has also been ported to the BCI2000[14] framework for brain-computer interface (BCI) research used for data acquisition, stimulus presentation, and brain monitoring appli-

[14]http://www.schalklab.org/research/bci2000 (accessed on September, 2015).

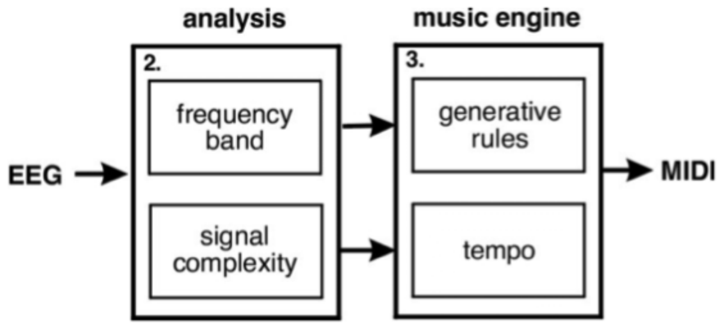


Figure 2.16.: Miranda's BCMI-Piano. System architecture [Miranda, 2006].

cations. More recently, Wu and colleagues presented a method for representing mental states through music [Wu et al., 2010]. Arousal levels, based on EEG features extracted using wavelet analysis, were mapped to music parameters such as pitch, tempo and rhythm. After the extraction of the EEG signal features, musical segments based on the extracted features were defined. Next, bars of music are generated, and notes are fixed based on bar parameters. Finally, a melody is constructed using MAX/MSP[15], and a MIDI file is generated. Their results suggest that mental states can be identified by listening to the corresponding music composed using the system. Wu and colleagues also presented a mapping of EEG waveform amplitude to pitch based on the scale-free phenomenon. The change of EEG energy was mapped to note volume and the period of EEG signal was mapped to the duration of notes [2009].

Other musical BCI systems have been developed using P300 features, an event related potential (ERP) -or brain response- component elicited in the process of decision making. P300 is considered to be an endogenous potential, as its occurrence is not linked to the physical attributes of a stimulus, but to a person's reaction to it [van Dinteren et al., 2014]. When recorded by EEG, it surfaces as a positive deflection in voltage with a latency of roughly 300ms after the stimulus appeared. Differently from the BCMI approaches mentioned before, P300 allows to create user interfaces that accounts for explicit control of the computer system using user selective attention. P300 based interfaces have been mainly built using visual stimuli (i.e. grids of numbers or letters flashing at different frequencies). This is the case of Hamadicharef and colleagues [2010] who used a P300 interface for allowing the user to select notes, rest, delete or play in the creation of short melodies.

[15] `https://cycling74.com/products/max/` (accessed on November, 2015).

Vamvakousis and colleagues [2014] went one step further by developing a BCMI (*P300 Harmonies*) based on auditory P300, using a low-cost EEG device (Emotiv EPOC[16]). In this case the user could voluntarily change the harmony of an arpeggio by focusing and mentally counting the occurrences of each note. The arpeggio consisted of 6 notes separated by an interval of 175ms. The notes of the arpeggio were controlled through 6 switches, where each switch has two possible states: up and down. When a switch was in the up-state, the note produced by this switch was one tone or semitone -depending on the switch- higher than when in the down-state. By focusing on each of the notes of the arpeggio, the user could change -after 12 repetitions- the state of the corresponding switch. The notes of the arpeggio appeared in a random order, and the state of each switch was shown on a screen and flashed when the corresponding note was heard. An interest aspect of this BCMI is its multimodality, as the user could either focus exclusively on the auditory presentation or make use of the visual stimuli as well. In the same line, Grierson and colleagues [2008] developed a computer music device also controlled by P300 events, through which users could control a synthesizer or sequencer remotely without moving, by making a subjective decision to focus on a particular choice offered to them on a display.

2.6. State of the art of SID based on EEG

The previous sections helped us to illustrate the significant contribution that sound and music research had throughout the past 50 years in the development of implicit and explicit interfaces based on human physiology. From all the possible approaches that we have described before, the sonic interaction designs based on human EEG are of particular interest for this dissertation, as it is the biosignal we have mostly used in the design of sonic interactions for implicit physiological computing systems. The decision of leveraging on EEG above other biopotentials was motivated by:

- **Background knowledge**: EEG is a well known and well studied technique that has been widely applied in different domains such as medical diagnosis, rehabilitation, cognitive science, neuroscience, among others. The research corpus around EEG offer widely accepted methods for acquiring, processing and interpreting EEG data, thus providing access to cortical correlated of sensory, cognitive and motor activity.
- **Affordances for implicit interaction**: as shown in previous sections, EEG-based interfaces can rely on the continuos monitoring of the user cognitive, emotive and

[16] `https://emotiv.com/epoc.php` (accessed on September, 2015)

perceptive states for implicitly inform adaptive system, without the need of direct user attention or intention. This is a main aspect of the *neurofeedback* loop (biofeedback based on brain activity) where the systems not only implicitly adapts to the user's states (e.g. relaxation), but also utilizes this information to trigger personalized (implicit or explicit) outputs (e.g. sound, images, audiovisual) that help the user to learn about her/his current physiological state, and even train for reaching a given goal state (e.g. stress control).

- **Affordances for SID**: as shown by Rosenboom's review, EEG has been one of the most used biosignals for SID [1997b]. The idea of making EEG signals audible accompanied brain imaging development from the very first steps in the early 1930s, when Prof. Edgar Adrian listened to his own EEG signal while replicating Hans Berger's experiments [Adrian and Yamagiwa, 1935]. On the other hand, SID appears as an excellent candidate for displaying EEG, since sound can readily represent the complexity and fast temporal dynamics of brain signals, and human auditory perception provides the highest temporal resolution among the sensory modalities.
- **Ergonomics & multimodality**: the recent emergence of portable, wireless and wearable EEG sensors such as the Emotiv EPOC or the Enobio have made possible to incorporate EEG into HCI without restricting user mobility and behavior. Moreover, as non-invasive EEG sensors are directly attached to the user's scalp using a cap, it can be easily combined with other implicit or explicit inputs methods such as motion sensing, tangible and touch surfaces, and other biopotentials such as heat-rate activity, etc.
- **Richness**: differently from other biosignals such as ECG or EDA, a great variate of high level features can be extracted from EEG data, from affective states (looking at brain oscillatory activity) motor actions or motor imagery (the mental execution of a movement without any overt action or peripheral (i.e. muscle) activation [Mulder, 2007]) or directly addressing visual or auditory perception (as in the case of P300).

On the negative side, many of the research that combine EEG and SID use rather arbitrary conversions of EEG data into sound. In addition, the associated publications often do not provide sufficient details about either physiological data acquisition or applied sound synthesis. Very few of these studies have conducted any kind of controlled evaluation of their chosen methods, making it impossible to replicate or validate most studies. Given these widespread limitations, a critical review of the current state of this emerging field may help to facilitate its future healthy development. Recent technical developments in both PhyComp and sound technology are making practical new ways of exploiting the real-time representation of brain activity using sound, for functional representation (sonification) and expressive performance (NIME). For this reason, following we provide a systematic review on the recent SID research based on EEG (published as a conference paper in Väljamäe et al. [2013b] and updated for this dissertation).

We first present a synoptic summary spanning over fifty different research projects and published articles. This is followed by the analysis of SID approaches and EEG data dimensions such as time-frequency filtering, signal amplitude, and location, before going through higher order EEG features. By following this approach, we are able to address wider questions, namely:

- *What application domains* have employed sound for representing EEG activity?
- *What SID techniques* have been applied to real-time EEG activity?
- *What EEG features* are mostly expressed through sound, and with what temporal resolutions?
- *What experimental, methodological and validation techniques* have been used in SID for EEG?

2.6.1. Application areas

There are diverse application areas that rely on sound for displaying EEG activity, but for various different reasons. Six distinct application areas can be differentiated in terms of their use of data (real-time or off-line), and on a continuum between functional (sonification) and aesthetic (NIME) as shown in Figure 2.17. By following this approach, six main application areas can be identified:

- **Monitoring:** designed to inform a third person about the user's EEG in real time (e.g. an anesthetist during surgery [Glen, 2010]) or to inform the user her/himself (e.g. an air traffic controller being warned of her critical fatigue level [Pope et al., 1995]).
- **Diagnosis:** sonic display of recorded (offline) EEG data for diagnostic purposes, where brain imaging data is normally speeded up (usually by a factor of 50-200 times) to allow fast identification of prominent changes (e.g. different sleep states [Olivan et al., 2004] or epileptic seizures [Khamis et al., 2012]).
- **Neurofeedback:** these applications target learning about one's own brain states, aiming at altering this state by training (e.g. for post-stroke rehabilitation or stress

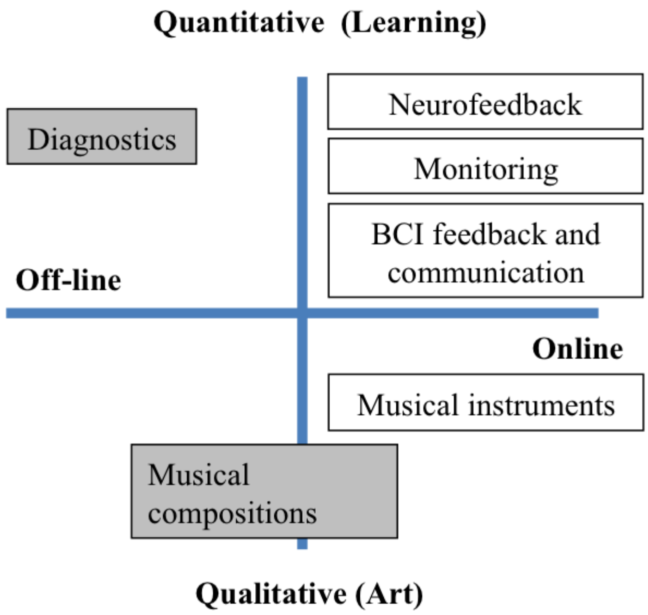


Figure 2.17.: Application areas that use sonic display of EEG data. Applications areas can be distinguished in terms of *EEG data processing* (real-time or off-line), and on a continuum between *functional and aesthetic*. [Väljamäe et al., 2013b]

control).

- **Brain Computer Interface and communication**: in this case, SID is applied for representing EEG activity. which is used as direct input for a system (e.g. the above mentioned motor imagery to control, for instance, a computer application or a prosthesis [Hinterberger et al., 2004, McCreadie et al., 2013]).
- **Musical instruments**: This domain includes the aforementioned BCI for direct control of musical parameters [Vamvakousis and Ramirez, 2014, Grierson, 2008, Wu et al., 2010]
- **Musical compositions**: EEG patterns converted to music in an indirect manner (without direct explicit user control) like in the case of Miranda's BCMI technology, which detects specific brain patterns and turn them directly into musical composition rules [Miranda and Wanderley, 2006].

All of these application domains have different objectives, different constraints and different validation methods[17]. One of the dimensions that helps to differentiate between

[17] See Barrass et al. [2006] for a thorough review on this topic.

SID approaches is to contrast the ways in which one would expect the end result to be judged (i.e. quantitative vs qualitative). In the case of EEG-driven musical compositions, just listening to the sonic display may be sufficient for demonstrating that the system works. In the case of BCI based musical instruments, the player can judge the extent to which the interface accounts for expression and control. Both domains are however more concerned with the aesthetics of the resulting sonic composition than any other kind of validation. The situation is very different for diagnostic and neurofeedback, where the informational and perceptual value of produced sounds is of primary importance and determines the functionality of the application.

2.6.2. Search and selection criteria

To get an overview of the range of published works we searched a number of databases including *Web of Science*, *Pubmed* and *Google Scholar*. The search terms included "sonification, audio, sound, auditory display, EEG, neurofeedback, biofeedback". A clear trend of growing activity in EEG sonification was observed. For example, when searching Google Scholar using "sonification + EEG", only 25 publications are returned for 2002 but already 140 for 2012. Furthermore, approximately 70% are conference publications. We decided to report on a selection of these publications, according to the following selection criteria:

- **Use of sound**: we followed Grond's four conditions for sonification (see Section 2.4.1) and complemented it by also considering NIMEs. This approach allowed us to include a wide range of sonic interactions ranging from auditory icons and earcons (that can be used to represent discrete events of brain imaging data using ecological or symbolic sounds) to musical BCI or EEG-driven musical compositions. Auditory BCIs used in speech applications has been omitted[18].
- **Real-time**: several works in the field deal with auditory display of pre-recorded, time compressed EEG data with the purpose of prescreening specific events (the above mentioned diagnosis category). These works were not included in this review. However, we included works that did off-line analyses but could potentially lead to on-line SID at the later stage.
- **EEG only**: We did not include papers in our review that would deal with other types of brain imaging technology such as fMRI [Schmele and Gomez, 2012, Weinberg and Thatcher, 2006] or ECoG [Potes et al., 2012].

[18] See Wagner et al. [2013] for a comprehensive review on auditory and multi-sensory BCI's.

- **Same study, multiple papers**: When a journal paper revisited the same content than an earlier conference publication, the journal publication was used. In cases where the overlap was only partial, multiple publications were used as sources, but only counted as one in our statistics and table entries.
- **Contemporaneity**: the original publications of some early works on EEG sonification and music making were not founded. We did not include these works in this review.

2.6.3. Relating EEG signal processing to SID techniques

Figure 2.18 offers a first visual overview of the assessed research, where the labels corresponds to *SID strategies* and the X-axis to *data processing* techniques. Within these two dimensions, all selected publications were classified into one or more categories (i.e. a single article may cover more than one SID technique and/or data processing strategy).

SID techniques

As exposed in Section 2.4.1, we considered 4 main SID strategies that cover both representational (sonification) and performative (NIME) approaches: audification, parameter mapping sonification, model-based sonification and generative music:

- **Audification**: represents the simplest and oldest approach for displaying physiological data through sound [Adrian and Yamagiwa, 1935]. In this technique, variations in EEG data values are directly treated as a sound wave. It is often applied when off-line EEG data is time-compressed by a factor of 50-200, shifting EEG frequencies to audible spectra. This offline approach is excluded from Figure 2.18. Besides very early works, none of the reviewed papers have used audification.
- **Parameter mapping sonification**: this technique is currently the broadest and most used strategy for displaying real time EEG data through sound. The simplest example would be mapping activity in the EEG alpha band to the intensity level of a sound. This technique encompasses many mapping methods as described by De Campo et al. [2007], Hermann et al. [2002] (e.g. continuous, event-based, or Spectral mappings; distance Matrix method; induced waves/spikes; judging correlation; vocal sonification, etc.).
- **Model-Based Sonification:** this approach relies on mathematical models that generate sound according to the EEG data input. For example, a sound synthesis model for a bell sound might be changed by the amplitude of the EEG alpha

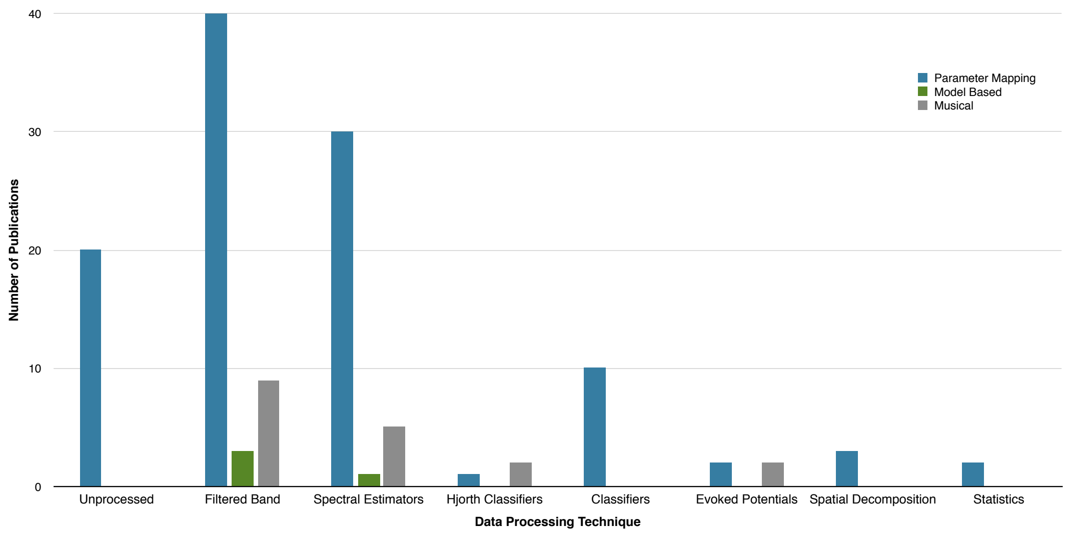


Figure 2.18.: Relationship between *Data Processing* and *SID techniques*. This figure shows several trends. *Simple processing of EEG data* has been gaining in popularity. Remarkably, only *Filtered Bands* and *Spectral Estimators* processing categories cover all 4 SID techniques. Among different frequency bands, alpha and beta activities have been the most displayed through sound. Interestingly, many works that applied more complex processing such as *Classifiers* tended to use simpler display techniques based on parameter mapping, while most of the EEG-based music approaches relied on simple signal processing methods. This may illustrate a difference between computer science and computer music communities, reflecting their different purposes. Audification is excluded as there were no works matching this SID category [Väljamäe et al., 2013b].

rhythm. This has an analogy to real-world sound generation phenomena. For example, one could shake a box in order to find out what is inside by means of the produced sounds. This indirect sonification approach has been gaining increasing attention over the last decades due its suitability for using different data sets as input [Hermann and Ritter, 2005, Halim et al., 2007].

- **Generative music:** this category describe systems that use musical rules and structures to create sound output driven by EEG data as input signal, like BCMIs[Miranda, 2006] and other performance oriented works that are mainly concerned with music expressiveness [Mann et al., 2007, Le Groux et al., 2010].

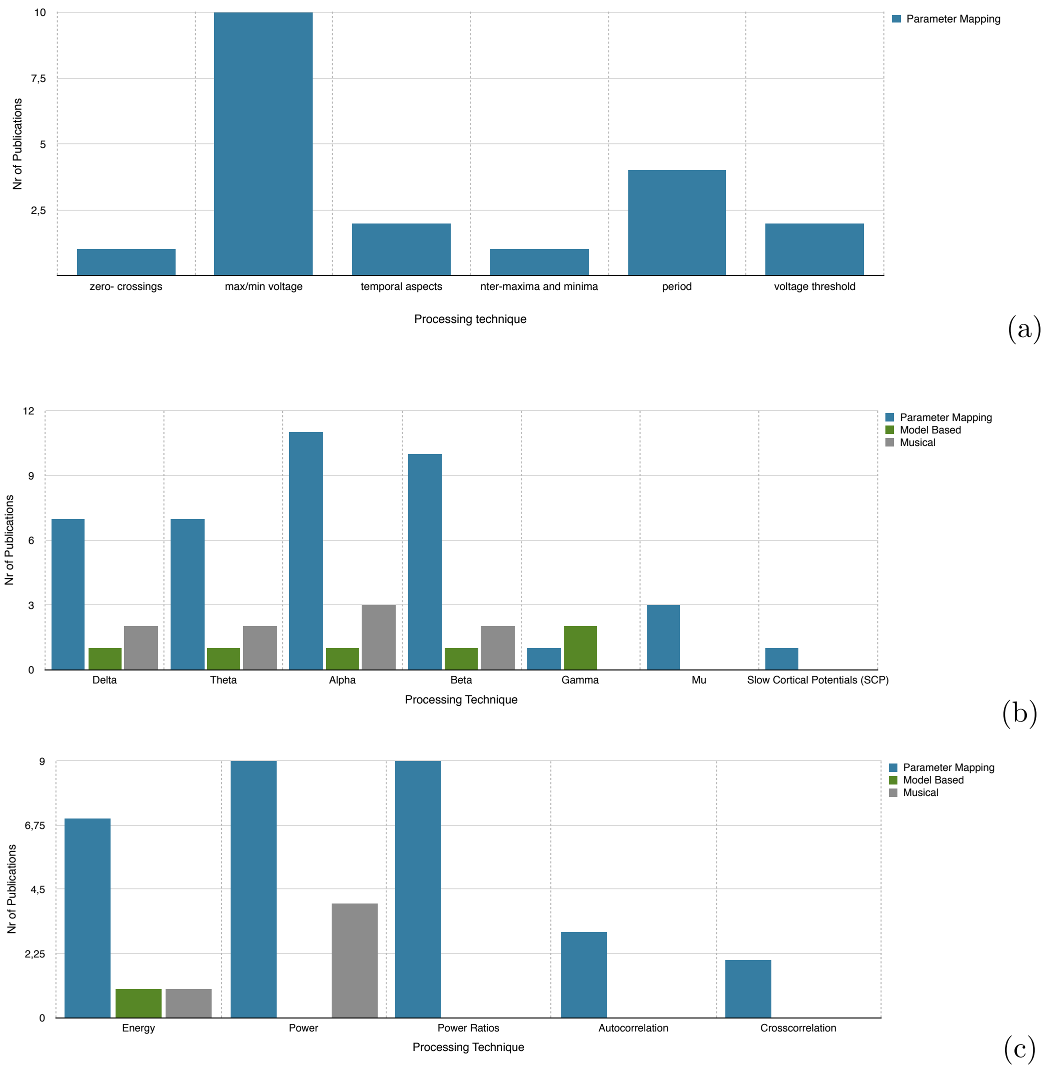


Figure 2.19.: Articles corresponding to the three most popular processing techniques (a) unprocessed EEG, (b) filtered bands, and (c) spectral estimators.

EEG Processing techniques

After analyzing the literature, we found eight main data processing categories (X-axis in Figure 2.18):

1. **Unprocessed**: sonic displays that use the raw EEG signal and its features such as registered zero- crossings, max/min voltage, temporal aspects of max/min voltage change, inter-maxima and minima values, period, and voltage threshold values (see Figure 2.19a).
2. **Filtered Bands**: refers to the auditory display of real time EEG data filtered in the time domain for the following bands: Delta, Theta, Alpha, Beta, Gamma, Mu rhythm, and Slow Cortical Potentials (SCP) (see Figure 2.19b).
3. **Spectral Estimators:** block-wise frequency conversion using FFT or other transforms in order to estimate the EEG features such as energy, power, power ratios including index of symmetry (SI), asymmetry ratio or affective states estimation, autocorrelation, and crosscorrelation (see Figure 2.19c).
4. **Hjorth's parameters:** refers to projects that use this signal complexity measure, like Baier et al. [2006] for parameter based sonification and Trevisan and Jones [2011], Miranda [2006] for music generation.
5. **Miscellaneous Classifiers:** this category includes Linear Discriminant Analysis (LDA) (McCreadie et al. [2012, 2013] for parameter mapping sonification), Principal Component Analysis (PCA) (Hermann et al. [2006] for parameter mapping sonification), Independent Component Analysis (ICA) (Vialatte and Cichocki [2006], Rutkowski et al. [2006] applied to parameter mapping sonification) or artificial neural networks (McCreadie et al. [2012, 2013], Arslan et al. [2005], Filatriau et al. [2006] applied to parameter mapping sonification, and Miranda [2006] applied to music generation).
6. **Evoked potentials:** mostly used for sound production [Rutkowski et al., 2006, Hamadicharef et al., 2010] and generative music [Le Groux et al., 2010, Miranda, 2006].
7. **Spatial Decomposition:** includes Common Spatial Patterns (CSP) (McCreadie et al. [2013], Arslan et al. [2005] applied to parameter mapping sonification), and Common Spatial Subspace Decomposition (CSSD) (Filatriau et al. [2006] applied to parameter mapping sonification).
8. **Statistical Analysis:** based on features such as Spectral Entropy (Filatriau et al. [2006] applied to parameter mapping sonification) and Gaussian Kernel 1 (Her-

mann et al. [2008] applied to parameter mapping sonification).

2.6.4. EEG processing dimensions

To have a closer look at the SID techniques used by different articles, we first identified the main EEG dimensions described by authors when converting brain data into sound. This resulted in four main dimensions:

Time-frequency dimension

One aspect of EEG processing to consider is to see how its temporal aspects are addressed, particularly if latencies are introduced. As mentioned is Section 2.4.1, the simplest and most straight-forward sonic display technique is the direct conversion of instantaneous values of EEG signals to sound (audification). This approach, despite being vulnerable to signal transients, can be useful for purposes such as locating outliers or detecting periodic patterns in the EEG [Hermann et al., 2002]. In this case latency, when present, is mostly caused by hardware limitations. A sliding window technique is often used to smooth the data (e.g. for reducing muscle artifacts by computing a moving average). This windowing approach can also introduce a delay, caused by the size of the window (typically varying between 50ms and several seconds).

Another source of delay is signal filtering. Time domain based approaches normally use finite or infinite impulse response filters to look at the signal in specific frequency bands. Frequency domain approaches, on the other hand, use windowing and block-based strategies to select a number of samples to convert into frequency domain, typically by means of FFT. In such systems, window size majorly determines the system delay.

The biggest latencies in the sonic display of EEG appear in event-based sonification approaches, where EEG data is buffered till a significant event occurs. Figure 2.20 shows the distribution of publications according to these time-frequency features. Please note that the list of works that use wavelet transforms is given in Figure 2.25, as typically these transforms are used as a first stage of higher order EEG processing.

Several trends become apparent when comparing the analyzed articles. First, the most popular technique for filtering is a block-based (typically FFT) conversion of the time signal into the frequency domain. This approach normally leads to latency, reflecting the used window size. Several works deliberately increase this, since their auditory display strategies are based on musical structures and event-based mappings. Second, a considerable number of works (25) still rely on straightforward sonification of instantaneous values with the objective of detecting outliers or periodical patterns in the EEG data. Many artistic works (that do not apply heavy signal processing) also fall into this category.

System latency is perhaps the most unreported (and more critical) issue for many applications. For instance, in sonic interactions for neurofeedback, any delay in the feedback will reduce the contingency of the sound signal to the brain activity, greatly diminishing training. In real-time EEG visualization, it is common to apply some averaging to the signal, in order to reduce any eye strain caused by the rapid flickering of the display, but this is often done without any appreciation of the detrimental consequences. Given the excellent temporal resolution of both the human auditory system and EEG, this is probably a key advantage of real-time EEG sonic display over visualization techniques.

The works using transform or filtering-based conversion to frequency domain are depicted in Figure 2.20. The majority of these works display certain frequency power or energy as described in next section and Figure 2.21. It should be noted, however, that a few works use frequency domain conversion to trace spectral dynamics of EEG and use shifts of a certain frequency or a certain band maximum frequency as features to be sonified [Hinterberger et al., 2004, Hermann et al., 2002, Wu et al., 2010, Hinterberger, 2011].

Level-based dimension

The level of amplitude of EEG is one of the fundamental properties of the signal as it reflects the firing rate of the neurons, which in turn mirrors the activation level of the underlying area of the brain and its information processing. A fast and accurate representation of this parameter is critical for estimating the rapid fluctuation of the underlying cognitive, affective and/or perceptual states of the user. In both neurofeedback and EEG monitoring, the level of particular sets of parameters is of primary interest (e.g. frontal alpha power reflecting the alertness of the user) Figure 2.21 summarizes the sonic interaction and display papers based on their treatment of various scalar features of the EEG signal, both in temporal and frequency domains. As it can be seen, EEG power is the most used parameter for sonic display. The second most used parameter is voltage amplitude. Surprisingly, relative power, which is commonly applied in EEG fields like neurology or neurofeedback, is rarely used for sonic display or interaction. One possible reason for this issue concerns the availability of baseline levels in the two contexts. In experimental settings, baseline levels (e.g. rest conditions) are routinely available to normalize the data for later comparison across users or sessions. By contrast, in SID, baseline recordings are rarely available. Ideally, both relative and absolute level values

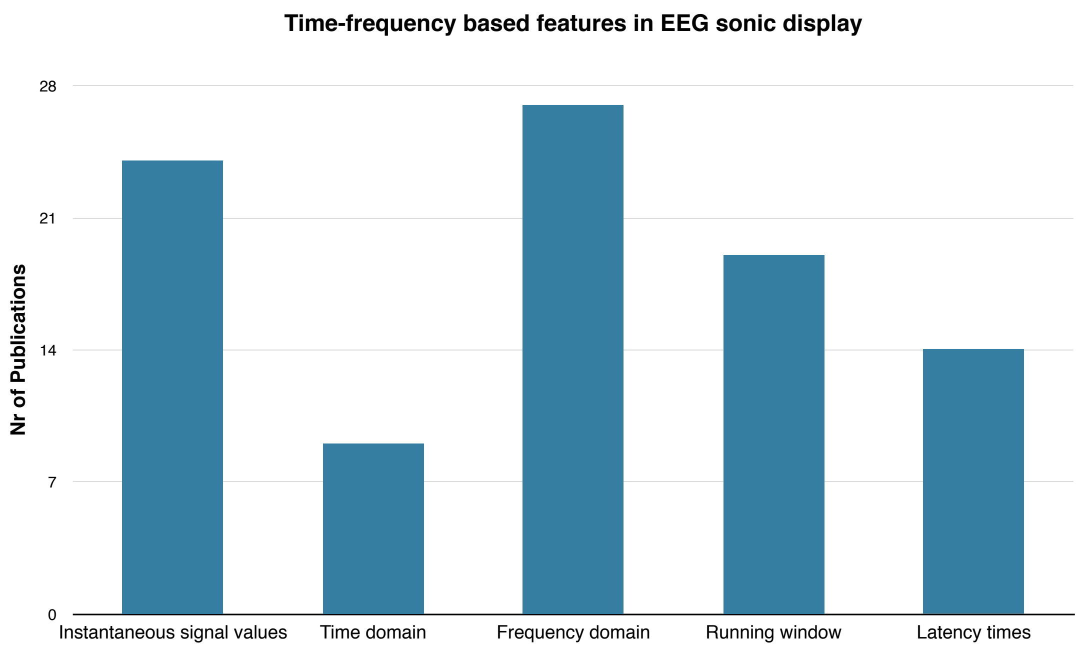


Figure 2.20.: Main EEG processing dimensions and number of articles falling in each category: time-frequency dimension [Väljamäe et al., 2013b]

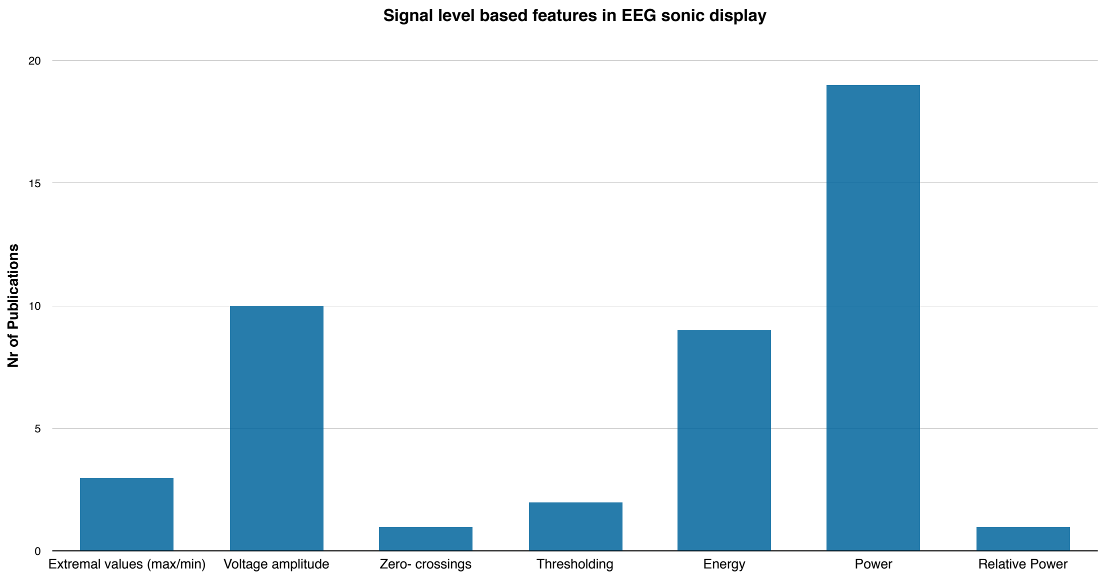


Figure 2.21.: Main EEG processing dimensions and number of articles falling in each category: signal level [Väljamäe et al., 2013b]

should be monitored. Because of the complexity and constant activity of the brain, there is always a high amounts of background noise to deal with when inferring activation of a particular area or network of the brain (e.g. onset of epileptic attack [Baier et al., 2007a, 2006, 2007b]). One of the main methods for addressing this issue is to detect extreme values, maxima values and link these to sonic events. Another method commonly used in neurofeedback is to detect when the parameter exceeds a given threshold or zero-crossing value. In this case, threshold selection is dependent on the application and can be critical for its success.

Location-based dimensions

The brain has specialized regions for different tasks with the sensorimotor cortex being the most prominent for EEG-based applications such as BCI and neurofeedback. However, many mental operations rely on networks of neurons working in concert. Therefore, despite the relatively poor spatial resolution of EEG, the location of electrodes is an important factor for measuring a specific cognitive operation. Furthermore, the usage of multiple electrodes allows detecting the neuronal activity from specific networks, permitting a number of space-related features for sonic display.

Figure 2.22 shows the distribution of papers regarding their use of multi-channel EEG systems. Most of the reviewed works use up to 20 EEG channels, with a smaller group of authors working with higher spatial resolution. The systems that used more than 32 channels are quite recent, reflecting the growing interest on multi-channel EEG systems. The amount of channels used also reflects the type of application. While practical and wearable systems tend to have fewer channels, research and diagnostic applications normally rely on more spatial resolution. Many high-order statistical methods for brain activity localization, like ICA and LORETA (see next section), require a minimum of approximately 20 channels.

Regarding the location of the electrodes in the scalp (see Figure 2.23) few works use sound features such as timbre or pitch transformation to represent electrode location. Spatialization and panning are common mappings, mostly designed to display sound according to the hemispheric locations of the brain activity. Interestingly, there are a number of emerging works that make use of channel correlation as an input for sonic display, such as timbre or pitch.

Details of electrodes location depend on their placement or, in other words, montage (see Figure 2.24). Standard electrode placements encourage the replication of SID techniques. Figure 2.24 shows that the so called 10-20 system [Homan et al., 1987] is the most common in the reviewed works, with a few specific exceptions that made use of custom placement methods, mostly due to hardware design (i.e. do-it-yourself devices or low-cost headsets). Surprisingly, a considerable number of publications (around 40%) neither specified placement system nor EEG sensor positions.

Features based on higher level processing of EEG

A number of projects applied computational techniques before converting data into sound. Indeed, it seems that any reliable correlate of human brain activity that represents some cognitive, emotional or perceptual process is likely to come from higher order processing of EEG data. An intermediate step towards this goal is to use higher-order processing of EEG to represent different brain region activities, as in the case of the biologically inspired method of bump sonification by Vialatte and colleagues [2006] that reveals time-frequency structures (oscillatory patterns) of brain activity through sound in real time. Many higher-order processing techniques depend on multiple electrode systems, which limits their application with less invasive and low-cost devices, but it might gain wider application as multichannel EEG sensing gets more affordable in the near future. On the other hand, it has to be said that many of the computational techniques used in EEG analysis are still not well suited for real-time applications. Figure

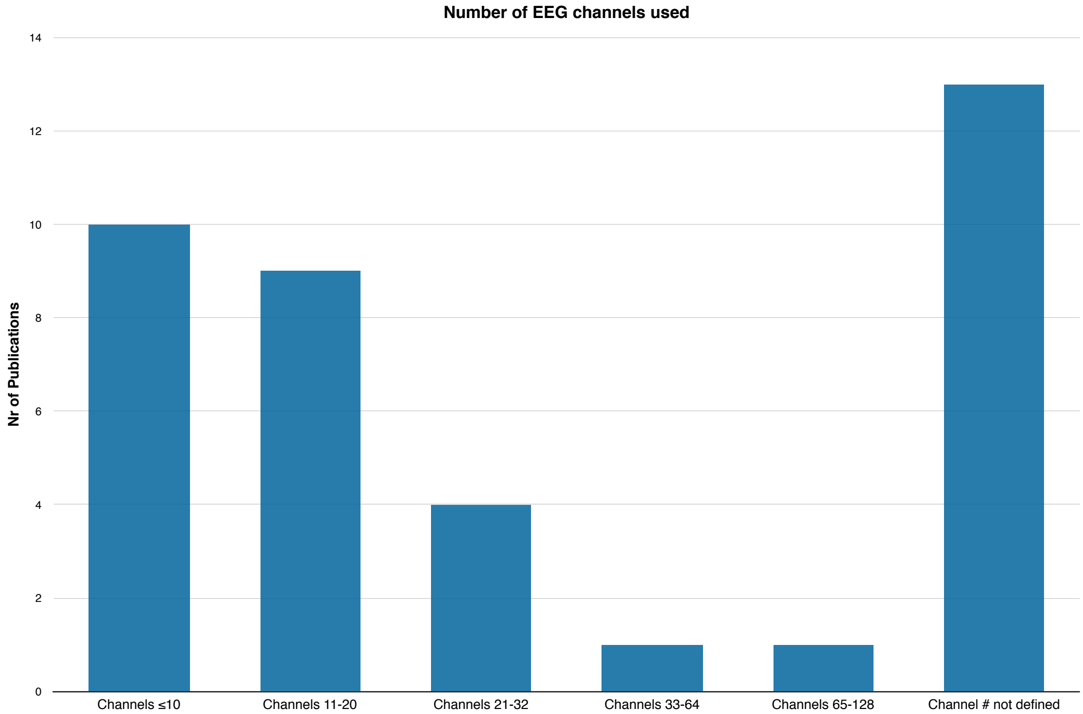


Figure 2.22.: Main EEG processing dimensions and number of articles falling in each category: channel number [Väljamäe et al., 2013b]

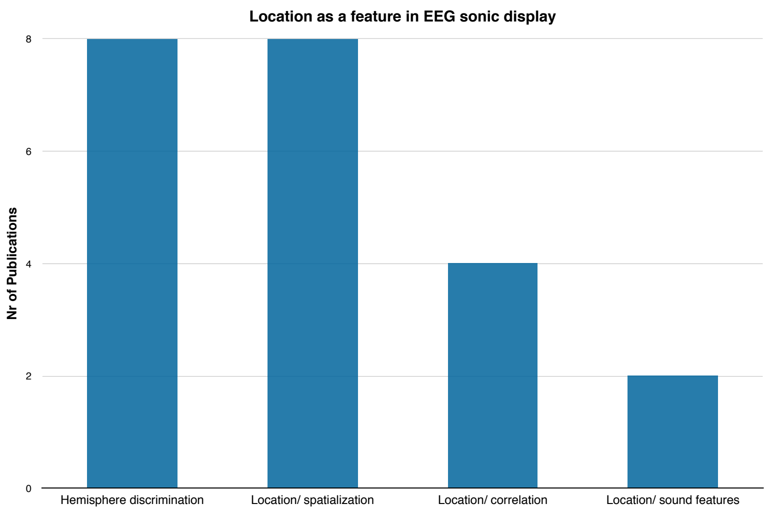


Figure 2.23.: Main EEG processing dimensions and number of articles falling in each category: channel location [Väljamäe et al., 2013b]

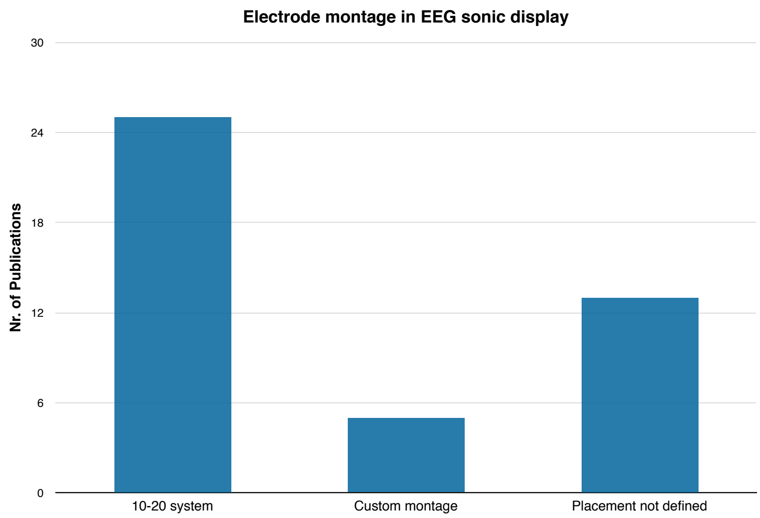


Figure 2.24.: Main EEG processing dimensions and number of articles falling in each category: montage system [Väljamäe et al., 2013b]

2.25 below demonstrates the divergence in tool sets used by different research groups.

Hjorth and wavelet based analysis are most commonly used among the presented techniques. However, it should be noted that it is still a small number of works (< 10 for each processing method) compared to more simple power band approaches reviewed in Figure 2.20. The Hjorth's EEG descriptors are mainly used in BCMI works. A recent and promising trend is to sonically display affective states of the user, since basic acoustic features are closely linked to emotional responses (e.g. rising/falling intensity influences emotional processing [Tajadura-Jiménez et al., 2008]). Real-time emotional state sonifications can thus be directly applied both for emotion regulation and for basic research on affective chronometry.

2.7. A summary of learnings for this dissertation

In this chapter we have provided a comprehensive review on the topics that conforms the background of this dissertation. We have shown that the interest for exploiting the implicit repertoire, rooted in human communication, is at the core of the HCI agenda, and has been tackled from theoretical, technical, and design standpoints as a way to create more natural, seamless interactions. We have also demonstrated that Physiological Computing (PhyComp) is an excellent candidate to foster implicit interaction through

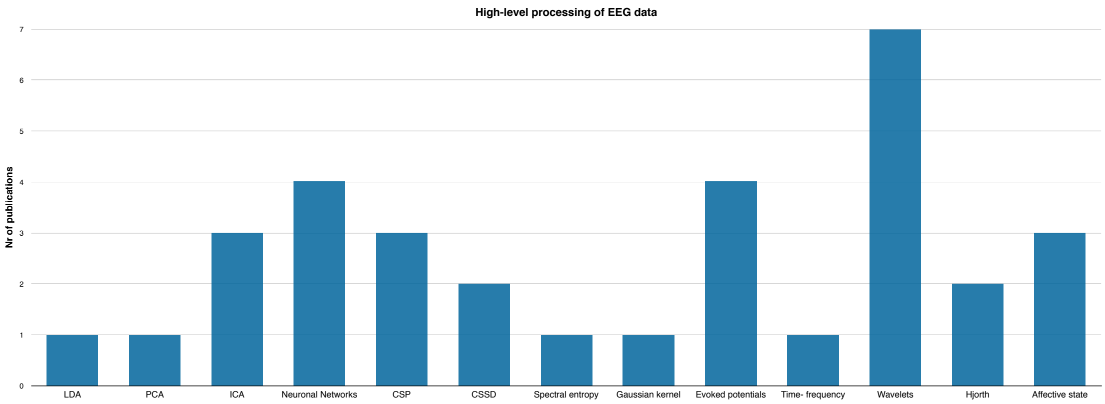


Figure 2.25.: Main EEG processing dimensions and number of articles falling in each category: high level processing features [Väljamäe et al., 2013b]

adaptive systems that can directly address the human body to access to the implicit psychophysiological states of the user (cognitive, perceptive, and emotional) for promoting personalized responses that can be delivered implicitly or explicitly to the user to improve her/his experience and performance.

We have also introduced sonic interaction design (SID) as a method for displaying such implicit psychophysiological states, analyzing the affordances of both representational, knowledge-driven approaches (i.e. sonifications) and aesthetic, performance-driven approaches (i.e. NIME). Finally, we have provided a systematic analysis on the state of the art of SID applied to the electroencephalogram (EEG), which is the main biosignal used in this dissertation.

Theoretical, technical and design driven approaches to implicit interaction have mainly agreed on the utter importance of understanding (or better said, of creating *systems that understand*) the situation and context where HCI takes place. Buxton's approach leverages on *background social ecology* to promote interactions that transit between foreground and background (as it is the norm in human-human communication) to compensate the *sense of distance* that exist in computer and computer mediated interaction. Weiser strived for devices that are capable of fostering user comfort by engaging center and periphery of our attention in a continuous tuning process. Tennenhouse's proactive computing emerged from the need of rethinking the boundaries between the physical and the digital world, strictly looking at the physical context; and Schmidt defined implicit interaction within HCI as the result of a complex interplay between the context of use and the perceptual capabilities of computer systems. For these contextual aspects, PhyComp appears as a valuable resource to expand our notion of interaction context by providing information about the *physicality* of users (what we call *subjective context*) that can easily encompass the *situational context.* In a pragmatical sense, PhyComp enhances the perceptive affordances of interactive systems, making available implicit information that can be only accessed by looking into the human body.

Another aspect of implicit interaction that can be significantly assisted by PhyComp is *proactivity* and *initiative.* These are fundamental aspects if we want to make background information immediately and effortlessly available, and relief humans from constant operation. Buxton associated the complexity in humans dealing with technology to having to explicitly sustain foreground activity; Weiser claimed for make computer systems proactive beyond human supervision; Ju & Leifer presented implicit interaction as a way to develop more sophisticated and less *needy* interactive systems, and their proposed framework helps us to identify *when* and *how* to deploy implicit interactions. In this regard, PhyComp can operate at both input and output levels: by informing about the user's perceptual, emotive and cognitive state in real time, through implicit channels (i.e. interpreting biosignals) without the need of user attention and intention; and by promoting system adaptation based on *subjective context* to decide *when* and *how* to be proactive and take the lead on the interaction.

As many other researchers and practitioners for the last 50 years, in this dissertation we will make use of sound for developing implicit PhyComp interactions. As has been shown in this Chapter, sound can readily represent the complexity and fast temporal dynamics of physiological signals, and human auditory perception provides the highest temporal resolution among the sensory modalities. We will also cover both representational and performance-based approaches to sonic interaction design (SID) to systematically study how these different modalities perform in an implicit interaction context. As it is expected, we will also pay special attention to the challenges and problems we have found when reviewing the current literature on EEG SID.

As shown by our review, many previous works lack technical details on the SID strategies used, the EEG recording equipment and the processing techniques. Our SID methodologies for implicit PhyComp interaction will therefore describe:

- The physiological features (objective properties or relations) that are sonically displayed (i.e. signal level, temporal, spectral or spatial patterns)
- The sonic parameters that are used for transforming/generating the auditory content, and what are the precise mappings established between these parameters and the extracted physiological features

Almost none of the analyzed works have carried on a systematic validation of their SID methods. In other words, it is still uncertain what types of SID strategies are most efficient for representing implicit physiological data in real time, and in different scenarios. In this regard, validation currently differs according to the application domain. Traditional PhyComp fields, such as medical diagnosis and neurofeedback therapies, mostly focus on assessing the effectiveness of a given display techniques (sound, visual or haptic) for *representing* a given physiological feature (e.g. heart-rate variability) or psychophysiological state (e.g. relaxation) through techniques such as randomized and double blind control studies. The NIME works reviewed for this dissertation, on the other hand, mostly either lack evaluation or, when it exists, it is mostly focused on *unstructured tests* in a musical context (i.e. live performance). In this dissertation we aim at filling in this gap, designing and comparing different SID techniques, in different context (including NIME) with fully described methodologies that facilitate cross-validation and replication.

Finally, very few PhyComp articles have explored multimodality within SID (e.g. combining sonic interactions with visualization). Given that human perception is multi sensorial [Calvert et al., 2004], combining auditory, visual and tactile information is likely to produce enhanced PhyComp interfaces that will benefit from different input methods for improving both functional and aesthetic purposes.

3. The b-Reactable: Prototyping Sonic Interactions for Implicit Physiological Computing

This Chapter presents our first conceptual, technical and methodological approach to Sonic Interaction Design (SID) for implicit Physiological Computing. The goal of this chapter is to study how implicit physiology-based interactions driven by sound can affect user experience in a meaningful HCI context: digital musical instruments (DMI). In order to do so, we start with a simple and straightforward SID strategy: audification of EEG activity and temporal control of sound by means of ECG. We use this SID to create a prototype called *b-Reactable*, based on a previous tabletop system (the Reactable). The *b-Reactable* allows implicit interaction (through EEG and ECG) and explicit gestural interaction (through tangible objects) for sound generation and control. Physiology-driven SIDs are embodied in tangible objects named *physiopucks*, which add a level of physicality to the above mentioned implicit interactions. We evaluate the effects of this approach in user motivation, and compared it to the original Reactable. The experiments involve dyads collaborating in three experimental groups (N = 56). The results of this chapter show that motivation dimensions are significantly higher in *b-Reactable* than in the Reactable, stressing on the positive effects of physiology-based implicit sonic interaction, and its combination with other inputs methods even in multi-user HCI scenarios.

3.1. Introduction and motivation

In the previous Chapter we have shown how Physiological Computing (PhyComp) can be used to achieve implicit and indirect interaction by continuously monitoring the users physiological activity to inform system adaptations and provide user-tailored feedback (implicit or explicit). Under this paradigm, user cognitive, perceptive and emotional states are classified and subsequently embedded into interactive processes without her explicit intention.

Although physiology-based implicit interaction has been already explored by different disciplines such as cognitive psychology, affective computing, and enactive media (see Chapter 2.3.2 for a review) most of these applications are restricted to activity tracking and monitoring, informing users about their performance in a given task (e.g. running or cycling). In this context, we are still in the need of implicit PhyComp systems designed for meaningful HCI scenarios, thoroughly evaluated and compared to existing interaction paradigms, in order to determine to what extent they might enhance human-computer and computer mediated interaction.

This chapter therefore presents our first sonic designs for physiology-based implicit interaction, deployed in a meaningful HCI domain: music performance. This first approach is based on the combination of explicit conscious control (i.e. tangible input) and implicit interaction based on two biosignals: EEG and ECG. To perform an empirical evaluation on how such implicit sonic interactions could affect HCI, we created a Digital Musical Instrument (DMI) named *b-Reactable*, which is based on the Reactable, a renowned musical tabletop interface [Jordà et al., 2007]. Our DMI can be therefore operated explicitly (through tangible objects called *pucks*) and implicitly by means of participant's brain activity (EEG) and (ECG). Through these physiological measures we are able to estimate low-level physiological features in real-time: user frontal EEG activity and heart rate. Both explicit and implicit interaction modes are used for triggering different types of sonic (musical) interactions. In the case of the physiological data, we use it for sound synthesis (audification of EEG) and for controlling the tempo of the DMI (beats per minute - BPM). These implicit physiological controllers are embodied in two physical objects, named *physiopucks*, allowing users to combine EEG and ECG features with other Reactable pucks, such as filters and controllers.

To compare experimental groups we use a Computer-Supported Cooperative Learning approach, which operationalizes user experience in multiple dimensions of motivation [Jones and Issroff, 2005]. The main research goal of this study is to determine whether the use of multimodal control through gestures and implicit interaction (physiopucks) can influence user motivation in a musical task, compared to the use of a gesture-only tabletop system, the Reactable. To this end, we hypothesize that:

- **Hypothesis 1**: Motivation of participants working with the *b-Reactable* will be higher than for participants working with the standard Reactable when performing the same musical task. Specifically, participants working with PhyComp (*Emitters*) will have stronger motivation than regular, gesture-only participants (*Users*) due to the addition of implicit interaction.

- **Hypothesis 2**: Experimental groups will differ on motivation scales. Specifically, the Sham group (where *physiopucks* are driven by pre-recorded EEG and ECG signals) will show lower motivation ratings compared to participants in the Physio group, were *physiopucks* respond to real-time EEG and ECG signals.

3.2. Materials

3.2.1. System Design

In this section we describe the *b-Reactable*, a DMI that introduces PhyComp to the Reactable system by means of real time EEG and ECG measures that are associated to tangible objects (physiopucks)[1]. Figure 3.1 offers an overview of the system architecture, and following we describe each component in depth: the Reactable, the physiopucks and their operation, the physiological signal acquisition and treatment, and the implemented sonic designs.

3.2.2. The Reactable and the physiopucks

The Reactable is a DMI based on a tabletop interface where tangibles objects (pucks) and hand gestures are used for controlling musical operations [Jordà et al., 2007] (see Figure 3.2). This is done by means of computer vision techniques (reacTIVision) that track both fiducial markers and finger gestures on the surface of the interface [Kaltenbrunner and Bencina, 2007] (see Figure 3.3). The Reactable sound synthesis and control methods follow a modular approach, a prevalent model in electronic music, which is based on the interconnection of sound generators and sound processors units. In the Reactable this is achieved by relating pucks on the surface of the table, where each puck has a dedicated function for the generation, modification or control of sound. The Reactable's objects can be categorized into several functional groups such as audio generators, processors (i.e filters and effects), controllers (which affect the behavior of the generators or processors they are connected to) and global objects (which affect the behavior of all objects within their area of influence). Each of these families is associated with a different puck shape and can have many different members, each with a distinct and human-readable symbol on its surface (see Figure 3.2).

[1] A video of the *b-Reactable* can be found at `http://www.dtic.upf.edu/~smealla/PhD_Material/videos.html`

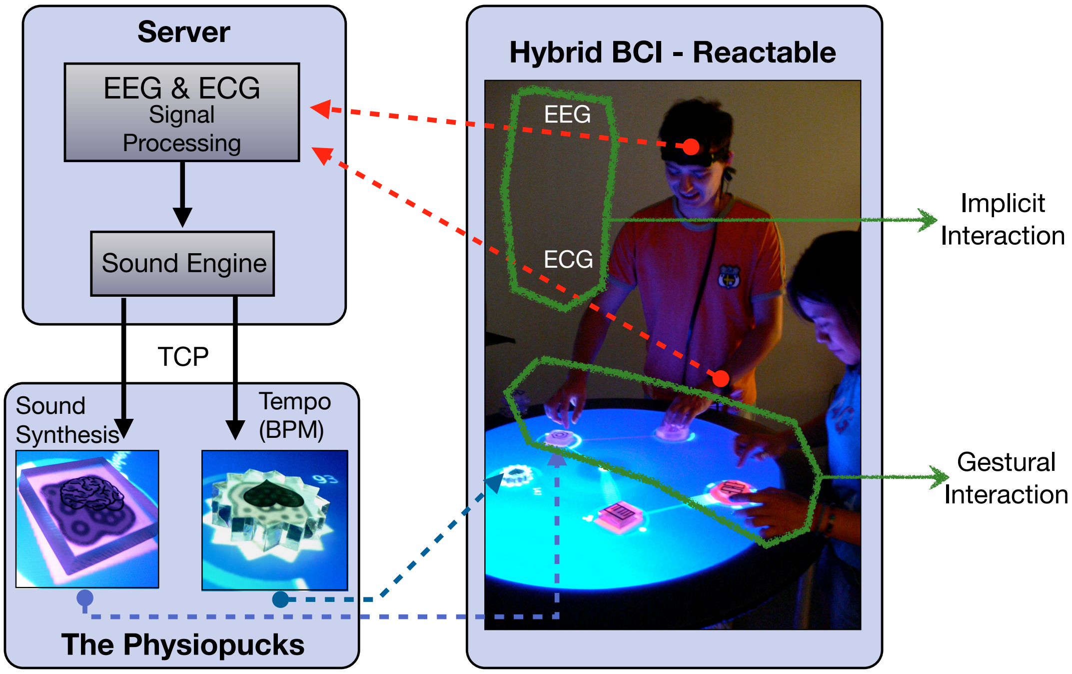


Figure 3.1.: The *b-Reactable* system. Physiological signals (red dotted arrows) are wirelessly streamed to a server that applies a signal processing and sonication. EEG-based sound synthesis and tempo control through heart rate are integrated in the Reactable framework, and presented to performers as physiopucks (blue dotted arrows).

Figure 3.2.: A performer playing with the Reactable.

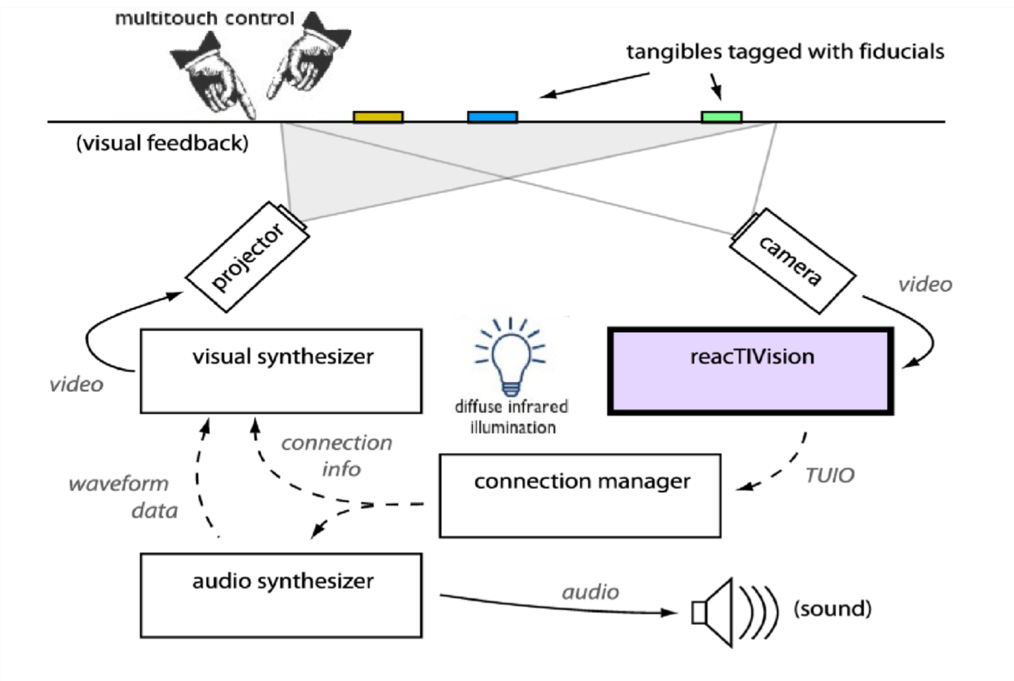


Figure 3.3.: The Reactable architecture and components [Jordà et al., 2007].

Because of this modular approach, the integration of physiological computing features into the Reactable is straightforward. We added two new pucks (*physiopucks*) to the current Reactable framework. In this new version of the Reactable, *physiopucks* allow performers to use their physiological signals (namely EEG and ECG) to generate sound (audification) and control the tempo (BPM) of compositions, in the same manner as using standard Reactable objects (see Figure 3.4).

3.2.3. Physiological Signal acquisition

The *b-Reactable* allows input from different physiological equipment. In this study we use a Starlab *Enobio*[2], a wearable, wireless electro-physiology sensing system that captures three biopotentials: EEG, ECG, and EOG. The Enobio features four channels connected to dry active electrodes at a sample rate of 250 Hz. For acquiring physiological activity to feed the *b-Reactable,* an electrode is placed on the frontal midline (Fz) lobe of participants (for EEG recording), and another electrode is placed in the participant's wrist for ECG detection. Physiological signals are acquired, amplified and streamed wirelessly to the Enobio software suite that applied a band pass filter (centered between 50 and 60 Hz) for noise reduction, and sent physiological data to a sound engine via TCP/IP.

3.2.4. Sound engine

In this study, the choice of SID strategies had two main motivations. First, we wanted to provide direct sonic feedback on the changes of EEG frequency bands with a minimum of latency. Second, we aimed at a simple, distinctive sonic interaction that would stand out from other sounds generated by Reactable. These guidelines led us to design a sonic interaction based on audification (see Section 2.4.1) as it allows directly translating data waveform into sound. This is normally achieved by resampling and digitally filtering input values to make them audible. Audification is particularly applicable to large datasets with periodic components, as in the case of the EEG and ECG.

The *b-Reactable* leverages on the already existing Reactable sound engine to generate a direct mapping between EEG frequency bands and the audible sound frequency spectrum. For performers, this sonification appears as a sound generator puck (brain-labeled physiopuck) on the b-Reactable. On the other hand, ECG activity is used to control the tempo (beats per minute - BPM) of the interface, appearing as a heart-labeled physiopuck (see Figure 3.4). To make it easier for participants to understand the tempo of the

[2] `www.neuroelectrics.com` (accessed on October, 2015).

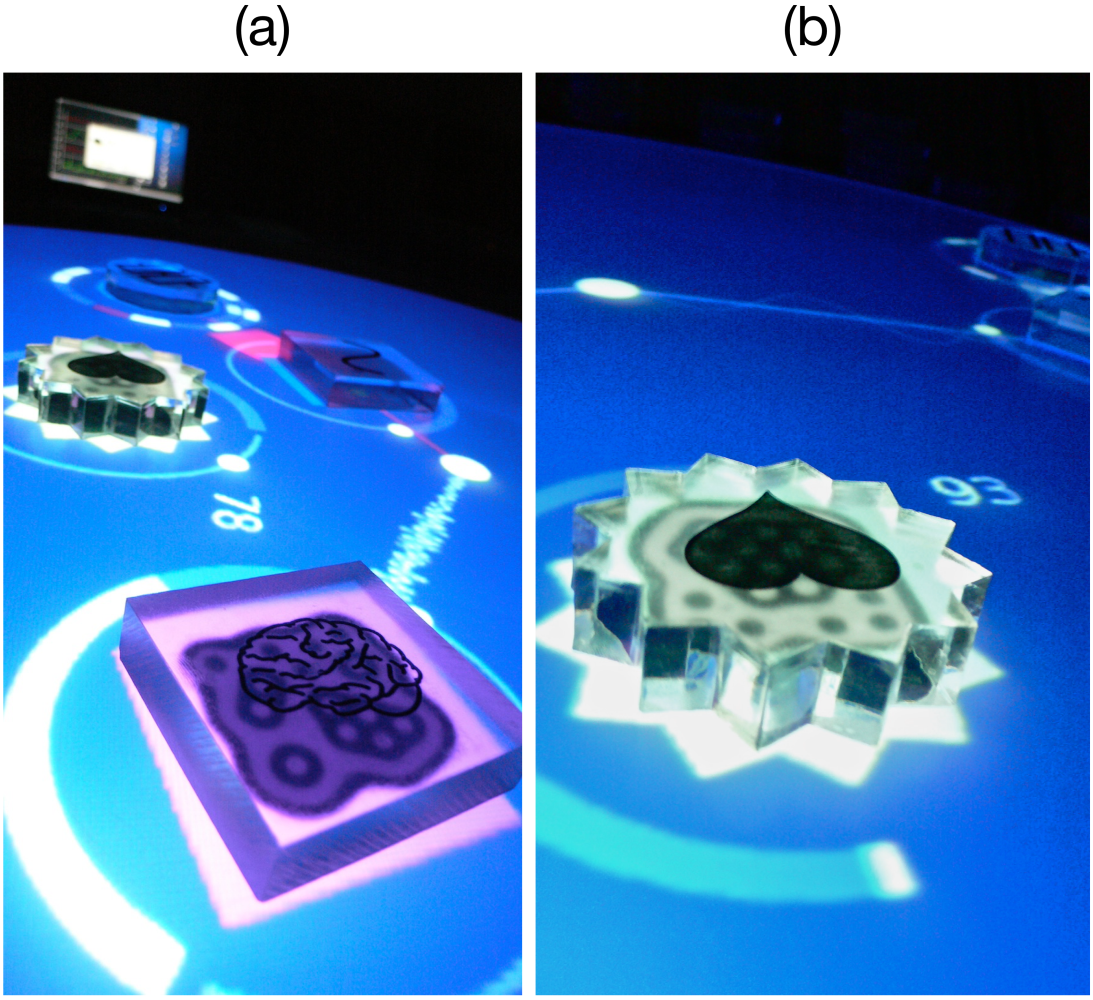


Figure 3.4.: The two *physiopucks* added to the Reactable framework. When placed on the tabletop the brain labeled physiopuck (a) produces an audification of EEG activity based on a white-noise signal. *Users* and *Emitters* can change the amplitude of the audification by moving the graphic slider that appears on the right side of the puck. The heart labeled physiopuck (b) drives the BPM of the sound composition according to the heart rate of the Emitter. The BPM values are displayed in the upper right corner of the physiopuck.

b-Reactable, BPM values were displayed in the upper right corner of the heart-labeled physiopuck. The sound engine was developed using Pure Data (Pd), a visual programming language for computer music [Puckette et al., 1996]. Pd was chosen due to its openness and suitability for performing such tasks, and for its flexibility when defining the mappings. This software also favors a robust integration with the Reactable framework, whose sound engine was also built with Pd.

3.2.5. EEG and ECG signal processing and sonic display

Following we provide a complete description of the DSP for EEG and ECG data, together with the correspondent sonification and BPM control mappings. Figure 3.5 shows a diagram of the main building blocks of the EEG sonification and ECG based BPM control. We processed the signal coming from the Enobio by first applying a DC block filter to contrast the signal drift and then performing a frequency magnitude analysis. Each block was multiplied by a Hann window function of the same size. An FFT with a size of 256 samples is then performed, leading to a spectral resolution of $0.97Hz$ per frequency bin.

The computed magnitude spectrum for each block is then used to shape the spectrum of a white noise signal. Each frequency bin of the EEG is used to weight the first 128 frequency bins of a 256 bins white noise FFT. Working at $44.1kHz$ for audio synthesis, we have a covered frequency range going from $0Hz$ to $11025Hz$, with each audio frequency bin covering about $86Hz$. The spectral magnitudes have been equalized by mean of weighting the chosen curve in order to emphasize the weaker higher frequencies.

This straightforward audification approach was used to map dominant EEG activity (mainly alpha band, $8 - 12Hz$) to human audible frequencies. In this manner the alpha activity, which is known to be associated with activation/relaxation [Wheeler et al., 1993], dominated the audification. This allows listeners to hear periodic components as frequencies. As demonstrated by Pauletto and Hunt [2005], through audification users are able to detect attributes such as repetitive elements, regular oscillations, discontinuities, and signal power to a degree comparable with using visual inspection of spectrograms. The EEG sonification is finally streamed over a TCP- IP/LAN connection to a server running the Reactable software, thus allowing to map it to the physiopucks, the objects that allow direct manipulation of physiological signal audification through hand gestures, and their combination with other Reactable objects (i.e. filters and controllers).

The overall EEG DSP and audification process implied an inherent latency of about 1 second. This could represent a problem in case of discrete control of a sound process

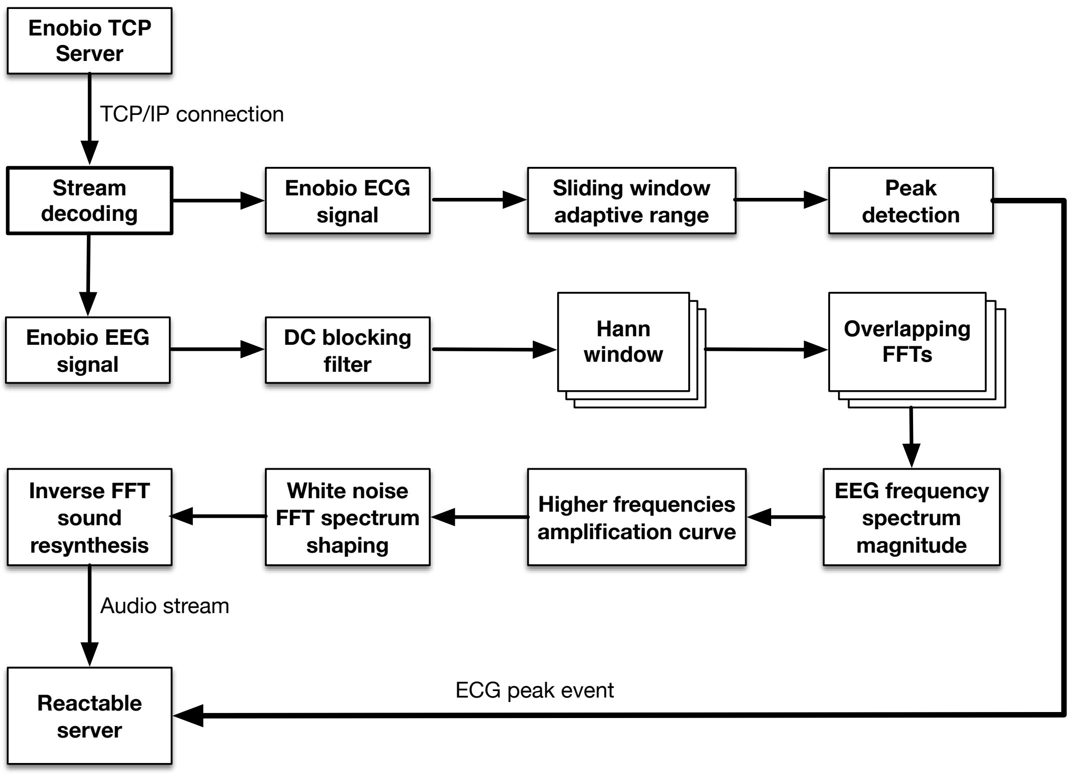


Figure 3.5.: Signal processing for EEG audification and ECG tempo control.

(i.e. triggers), which would require a maximum latency of 20 msec. However, higher latencies can be considered as tolerable for a continuous control, as the one applied in our sonification system [Wessel and Wright, 2002, Lago and Kon, 2004], also given the low frequency and low-variability of alpha rhythms.

ECG measures were used as a control mechanism to adjust the BPM of the system to the average heart rate of the user. We applied an adaptive rescaling to the ECG signal in order to smooth changes without losing heart rate peak resolution. A two-seconds sliding window (500 samples) is used to detect minimum and maximum values, and signal is normalized depending on that sliding window range. Afterwards, peaks in the ECG are detected by applying a simple threshold function. A heartbeat is detected if the normalized signal is above the 40% of the normalized range. A new heartbeat is then detected only if this signal falls below 30% of the normalized range.

3.3. Methods, experimental setup and procedure

To assess the effect of multimodal control on user motivation, we designed a task-oriented experiment involving two participants working together (i.e. a dyad). We chose this configuration as it allowed us to study how participants perceive their physiology-driven operations during music performance, and how a partner perceives implicit interaction involved in the same task. The experiment took around 45 minutes (equal for all experimental groups), and was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later revisions Association et al. [2001]. A total of 56 participants were distributed in three groups: the Physio group, where the multimodal system was fully functional; the Sham group, where *physiopucks* were driven by pre-recorded physiological signals; and the Control group, where no *physiopucks* were present, thus participants were using the conventional Reactable. The Physio group contained 11 dyads, with age mean of 28 ($SD = 3$), 10 females. The Sham group contained 11 dyads, with age mean of 27.1 ($SD = 3.5$), 9 females. The Control group contained 6 dyads, with age mean of 27.7 ($SD = 4.1$), 6 females. It is important to clarify that the reason for the Control group having half the dyads compared to the other groups is because its participants were all of the same type (*User)*. Following we provide a detailed description of each experimental group.

3.3.1. The Physio group

The Physio group involved a pair of participants with two distinct roles: a *User*, who operated the Reactable in a conventional manner (i.e controlling pucks with her hands) and an *Emitter*, who also manipulated the interface through standard pucks and gestures, but with the addition of providing physiological signals for the *physiopucks* (i.e. EEG audification and BPM controlled through heart rate).

In this way, changes in *Emitter's* arousal state drove the audification produced by the EEG *physiopuck* and the tempo was defined by the ECG *physiopuck*. This means that *Emitters* were continuously -and implicitly- controlling. They could also control *physiopucks* by trying to change their arousal state at will, and through hand gestures (i.e. reducing the volume of the EEG audification). *Physiopucks* were accessible to both *Emitters* and *Users*, therefore any of them could put them or take them out of the tabletop, or combine them with other Reactable pucks (i.e. filters, controllers).

3.3.2. The Sham group

Dyads in the Sham group were also integrated by *Users* and *Emitters*. However, *physiopucks* were not connected to the *Emitters'* physiological states, but driven by pre-recorded EEG and ECG signals. A placebo effect in this group was induced by making *Emitters* wear the physiological sensors, and by telling both members of the dyads that physiopucks were connected to *Emitter's* EEG and ECG activity.

With the aim of keeping the pre-recorded data as close as possible to the physiological signals of a *Physio-Emitter*, we collected and reused the EEG and ECG activities from a participant working on the same task as the ones applied in the experiment, but during a pilot session. All *Emitters* within the Sham Group used the same physiological recording.

3.3.3. The Control group

Participants in the Control Group also worked in dyads, but in this case both participants were *Users*. Therefore, dyads operated the Reactable only through hand gestures with no *physiopucks* or physiology involved. These *User-User* dyads worked with a tempo controller and an additional sound generator that replaced the two *physiopucks*. This allowed *Users* to perform similar operations to the ones allowed by the *physiopucks*. EEG headsets were also placed on both *Users*, explaining them that they were used for measurement purposes. In any case physiological data was connected to the sonification system; headsets were used to create a similar setup to the other two experimental groups

Type	Subtype	Connection
Generator	Oscillator	1 audio out N control in
Generator	Sampler	1 audio out N control in
Generator	EEG Physiopuck for resynthesis	1 audio out N control in
Audio Filter	Delay	1 audio out N control in
Audio Filter	Granular Filter	1 audio out N control in
Controller	Sin wave low frequency oscillator	N control out
Controller	Sequencer	N control out
Global	Metronome	N control in
Global	ECG Physiopuck for BPM control	N control out

Table 3.1.: Description of the Pucks and Physiopucks available during the experiments, according to Reactable taxonomy, subtype, and connections allowed.

(i.e. wearing and seeing physiological equipment). The Control Group performed the same tasks and followed the same procedure as the Physio and Sham groups.

3.3.4. Experimental procedure

The three experimental groups followed the same procedure, with different dyad composition as described in the previous section. Before starting the experiment, participants were asked to sign a consent form and fill out a pre-test questionnaire (see Section 3.3.5 for more details). Then the physiological sensors were placed on the *Emitter's* scalp and on the wrist of the non-dominant hand respectively. In order to reduce movement artifacts during ECG acquisition, we asked *Emitters* to keep their non-dominant hand in a resting position or on top of the tabletop interface. In the Control Group, EEG headsets were placed on both participants. A testing period followed for around 10 minutes, where electrode impedance and data acquisition was checked. Afterwards, all dyads went through a 5-minute explanation session, with the aim of introducing them to the *b-Reactable* and the *physiopucks* (or the standard Reactable in the case of the Control group). This session included seven Reactable pucks plus the two *physiopucks*, and were the same set of objects available for carrying on the tasks (see Table 3.2). After the explanation session, dyads had 5 minutes to freely explore the interface through gestures and physiology with the same set of objects. In the Sham group, once the exploration was finished, *Sham-Emitters* were secretly disconnected from the interface (while keeping the electrodes) to carry on the tasks with pre-recorded physiological data.

The experiment included two tasks, each one consisting on the replication of a 15-second music excerpt that was created with the same set of pucks that were available to the

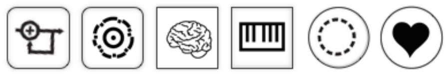
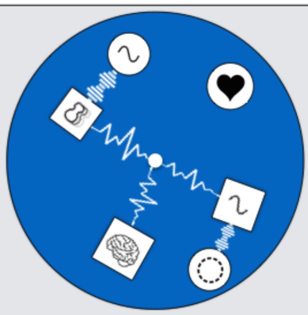
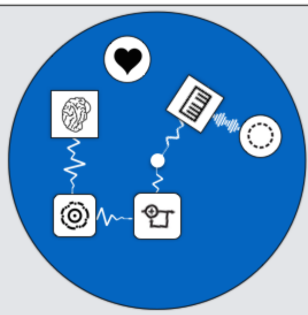
	Task 1	Task 2
Description	A fast pace composition that includes a synthesizer (violin) with a modulator for pitch variation, and an oscillator controlled by a step-sequencer. EEG audification generates a high frequency white noise.	A low pace composition that includes a synthesizer (piano) triggering the same note through a step sequencer, and low frequency EEG audification transformed by a granular filter and delay.
Pucks involved		
Task Solved	 High Physiological Activity	 Low Physiological Activity

Table 3.2.: Description of tasks involved in the experiment, including pucks involved and a representation of the solved tasks in the *b-Reactable*.

participants during the test. The tasks also required specific EEG and ECG states, namely high or low level of physiological activation, since the two music excerpts differed in musical pace and audification frequency (see Table 3.2)

First, dyads listened to the music reference. Then dyads had up to 5 minutes to mimic the reference music excerpt. Matching sounds required a concrete combination of pucks and *physiopucks*, and differed in the activity states required from *Emitters* (see Table 3.2). The participants could ask the experiment leader to replay the sound reference at any time. The task finished either when the dyads declared to have matched the reference sound, or once the 5 minutes period ended up. After the task participants were asked to fill in the post-test questionnaire (see next Section for more details). Finally, sensors were removed; participants were debriefed and given a small reward (chocolate bars).

3.3.5. Measures

Demographic information from participants (including music knowledge, and familiarity with the Reactable) was gathered through a pre-test questionnaire based on a 5-point interval scale. Statements for general music knowledge included "I can play music" and "I can compose music", whereas electronic music knowledge included "I can play an electronic instrument" and "I understand how an electronic music instrument works".

After each task, we used a post-test questionnaire that included a 9-point Bidimensional Self-Assessment Manikin pictorial scale (SAM) for assessment of valence and arousal [Bradley and Lang, 1994], and eleven motivational aspects based on CSCW literature ([Jones and Issroff, 2005] and references therein). A *Curiosity* measure was added from the attitude scale of Eagly and Chaiken [1998]. Each of motivation measures is based on one or several statements to be rated on 5 or 10-point interval scale. The ratings for each measure were collected through computer-based questionnaires, and the mean was taken if several questions were corresponding to a specific motivation aspect. Below are listed motivation measures with an example statement in the questionnaire.

- **Curiosity (M1)**: Task perceived as unusual, example of statement: "Performing with Reactable was an unusual experience". 5-point interval-scale, 3 questions.
- **Difficulty (M2)**: Rates the difficulty level of the task, example statement: "The tasks were too difficult to be accomplished". 5-point interval-scale, 3 questions.
- **Confidence (M3)**: Self-efficacy on achieving the tasks, example statement: "I've accomplished the tasks with efficacy". 5-point interval-scale, 4 questions.
- **Distribution of Control (M4)**: Balance of control among subjects, example statement: "I feel I was leading most of the work in every task". 5-point interval-scale, 5 questions.
- **Social Affinity (M5)**: Willingness to work together and collaborate, example statement: "I have a relation of friendship with my partner". 5-point interval-scale, 2 questions.
- **Interface feedback (M6)**: Measures how multimodal feedback (visual, sound) aids the collaboration between pairs, e.g. "The visual interface of Reactable helped me to understand how to create sound compositions". 5-point interval-scale, 2 questions.
- **Motivation Time (M7)**: How does motivation change over time? Does the subject lose interest as time pass by? Example statement: "The first tasks were more compelling and interesting than the latter". 5-point interval-scale, inverted,

2 questions.

- **Satisfaction (M8)**: Measures enjoyment and positive attitudes towards the experience, example statement: “Playing with my partner was a positive experience”. 5-point interval-scale, 3 questions.
- **Verbal Communication (M9)**: Measures the importance of verbal communication for solving tasks. Example statement: “Please rate Verbal communication according to the importance you think it had during the tasks”. 10-point interval-scale, 1 question.
- **Visual Communication (M10)**: measures the importance of the visual feedback provided by the system as a mean of communication with partners. Example statement: “Please rate Visual Feedback according to the importance you think it had during the tasks”. 10-point interval-scale, 1 question.
- **Gestural Communication (M11)**: measures the importance of physical manipulation of tangible objects when communicating with partners during tasks. Example statement: “Please rate Body Gestures of your partner according to the importance you think it had during the tasks”, 10-point interval-scale, 1 question.

3.4. Results

The IBM SPSS v20 software suite has was used for statistical analyses. The alpha significance level was fixed at 0.05 for all statistical tests, and a Greenhouse-Geisser correction was used to compensate for unequal variances (Greenhouse and Geisser, 1959). For multivariate analysis, Wilks’ Lambda was used as the multivariate criterion. All variables were normally distributed according to the Kolmogorov-Smirnov test. The reported results were not correlated with participants’ age, gender or music knowledge indexes, so this analysis is not included in the sections below. For the correlation analyses, an adjusted Pearson’s correlation coefficient rho was used to compensate for the small number of observations [Howell, 2013]. Since the measures were averaged over several questions, we considered them as continuous variables.

In our analysis we first ran a 1-way ANOVA, comparing all ratings from the four types of participants of the Physio and Sham groups with the ratings from the Control group (i.e. standard Reactable users). Second, leaving out the Control group data, we ran a 2-way MANOVA comparing participant’s ratings using two between-subjects factors: “group” (Physio vs. Sham) and participant’s “role-in-dyad” (*Emitter* vs. *User*). Third, to study in depth possible similarity between *Users* and *Emitters* in dyads, we applied a correlation analysis. Finally, we compared the correlations between participants' emotional state (SAM scale) and reported motivation aspects.

3.4.1. Comparing the four b-Reactable participant types to the Control group

To address our first hypothesis, we compared all four participant types: Sham-users, Sham-emitters, Physio-users and Physio-emitters with regular Reactable users in the Control group using 1-way ANOVA and 2-sided Dunnett post-hoc test. From 13 measures, only 3 showed significant changes in the reported motivation aspects.

First, the **Confidence** ratings showed significant differences among participant types with $F(4, 51) = 3.18, p < 0.021, h_{p2} = 0.2$ (see Figure 3.6a). From the four participant types, two were significantly higher than the Control group ($M = 2.6, SE = 0.2$), both coming from the Physio group: Physio-users at $p < 0.028 (M = 3.4, SE = 0.3)$, and Physio-emitters at $p < 0.037 (M = 3.4, SE = 0.2)$.

Second, the **Satisfaction** ratings also showed significant differences among participant types with $F(4, 51) = 3.57, p < 0.012, h_{p2} = 0.22$ (see Figure 3.6b). From the four participant types, only *Users*' ratings were significantly higher than ratings in the Control group ($M = 3.9, SE = 0.2$): Physio-users at $p < 0.038 (M = 4.4, SE = 0.2)$ and Sham-users at $p < 0.006 (M = 4.6, SE = 0.2)$.

Third, the **Gestural Communication** ratings also showed significant difference among participant types with $F(4, 51) = 2.89, p < 0.031, h_{p2} = 0.19$ (see Figure 3.6 c). Only Physio-user's ratings ($M = 7.3, SE = 0.6$) were significantly higher than in the Control group ($M = 4.8, SE = 0.7), p < 0.03$. The second highest rating was from Physio-emitters ($M = 6.7, SE = 0.7$), followed by Sham-users ($M = 6.1, SE = 0.6$) and Sham-emitters ($M = 4.9, SE = 0.6$). It should be noted that no significant differences between *b-Reactable* and Control groups could be seen for the **Visual and Verbal Communication** ratings.

3.4.2. Comparing Physio and Sham groups

To address our second hypothesis, all 11 measures of motivation, together with the two SAM scale ratings of valence and arousal were submitted to a multivariate analysis with two between-subjects factors: experimental "group" (Physio vs. Sham) and participant's "role-in-dyad" (*User* vs. *Emitter*).

The multivariate effect of the "group" factor reached significance at $p < 0.044$ with

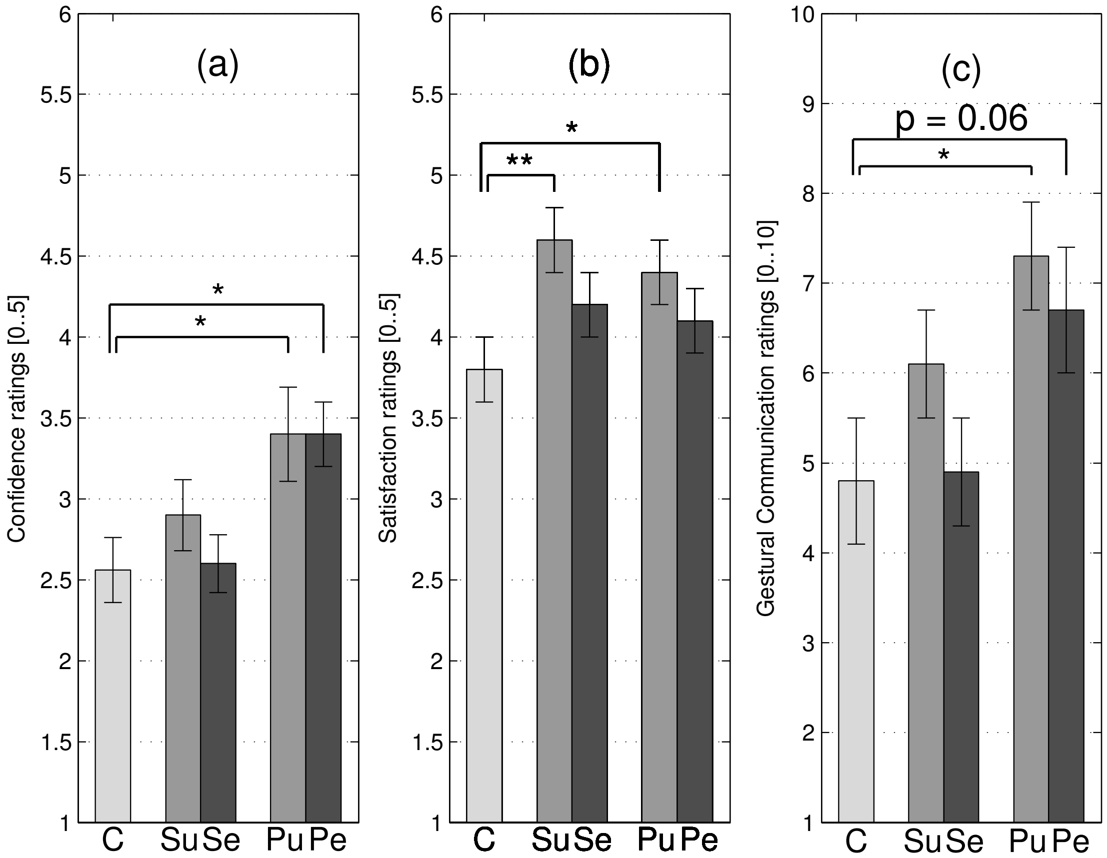


Figure 3.6.: Confidence (a), Satisfaction (b) and Gestural Communication ratings (c) across five participant types: the users from the Control group (C), Sham-users (S-u), Sham-emitters (S-e), Physio-users (P-u) and Physio-emitters (P-e). Error bars represent SEM values. *** - significance at$p < 0.005$ level, ** - significance at $p < 0.01$ level, * - significance at $p < 0.05$ level.

$F(13, 28) = 2.1, \Lambda = 0.51, h_{p2} = 0.5$. In a post-hoc analysis, two measures showed significant differences within this factor, similarly to the analysis in the previous section. First, the Physio group showed significantly higher **Confidence** levels than the Sham group, $F(1, 40) = 7.7, p < 0.008, h_{p2} = 0.16$. The corresponding means were $M = 3.4(SE = 0.2)$ for dyads in the Physio group and $M = 2.8(SE = 0.2)$ for the Sham group (see FIGURE 3.6a). Second, dyads in the Physio group also paid more attention to **Gestural Communication** than in the Sham group, $F(1, 40) = 5.62, p < 0.023, h_{p2} = 0.12$. The corresponding means were $M = 7(SE = 0.4)$ for dyads in the Physio group vs. $M = 5.5(SE = 0.4)$ for the Sham group.

The "role-in-dyad" factor reached significance only for the **Satisfaction** measure. In this case, *Users* declared higher level of **Satisfaction** than *Emitters*, $F(1, 40) = 4.28, p < 0.009, h_{p2} = 0.157$, in both experimental Groups. The corresponding means were $M = 4.5, SE = 0.1$ vs. $M = 4.1, SE = 0.1$ (see also Figure 3.6b).

Two significant interaction effects between the "group" and the "role-in-dyad" factors also emerged, and these response patterns are illustrated in Figure 3.7. The **Motivation Time** ratings showed significant differences, with Physio-users and Sham-Emitters declaring significantly more motivation than their partners in the corresponding dyad, $F(1, 40) = 6.61, p < 0.014, h_{p2} = 0.142$ (see Figure 3.7a). Similar interaction trend could be observed for the **Visual Communication** ratings with $F(1, 40) = 3.97, p = 0.053, h_{p2} = 0.09$. Here, same as for the Motivation ratings, Physio-users and Sham-emitters were paying more attention to visual feedback than their partners in the dyad (see Figure 3.7b).

3.4.3. Correlations between User-Emitter dyad

To complement direct comparisons in sections 3.4.1 and 3.4.2 and to study in depth the collaborative aspects of the experiment, we applied a correlation analysis to evaluate the consistency of the *User-Emitter dyads'* responses to each measure in the questionnaire (see TABLE 3.3). When all the questions were combined together, only the Physio group showed a significant level of consistency between participants' responses in *User-Emitter* dyad. When correlations were analyzed measure by measure, a more detailed picture emerged.

First, the **Curiosity** measure showed a significant positive correlation within the user-emitter dyad, but only for the Sham group. Second, the **Confidence** measure showed a significant positive correlation within the *User-Emitter* dyad, but only for the Physio group. This is coherent with the results from the previous analysis (Physio-dyads show-

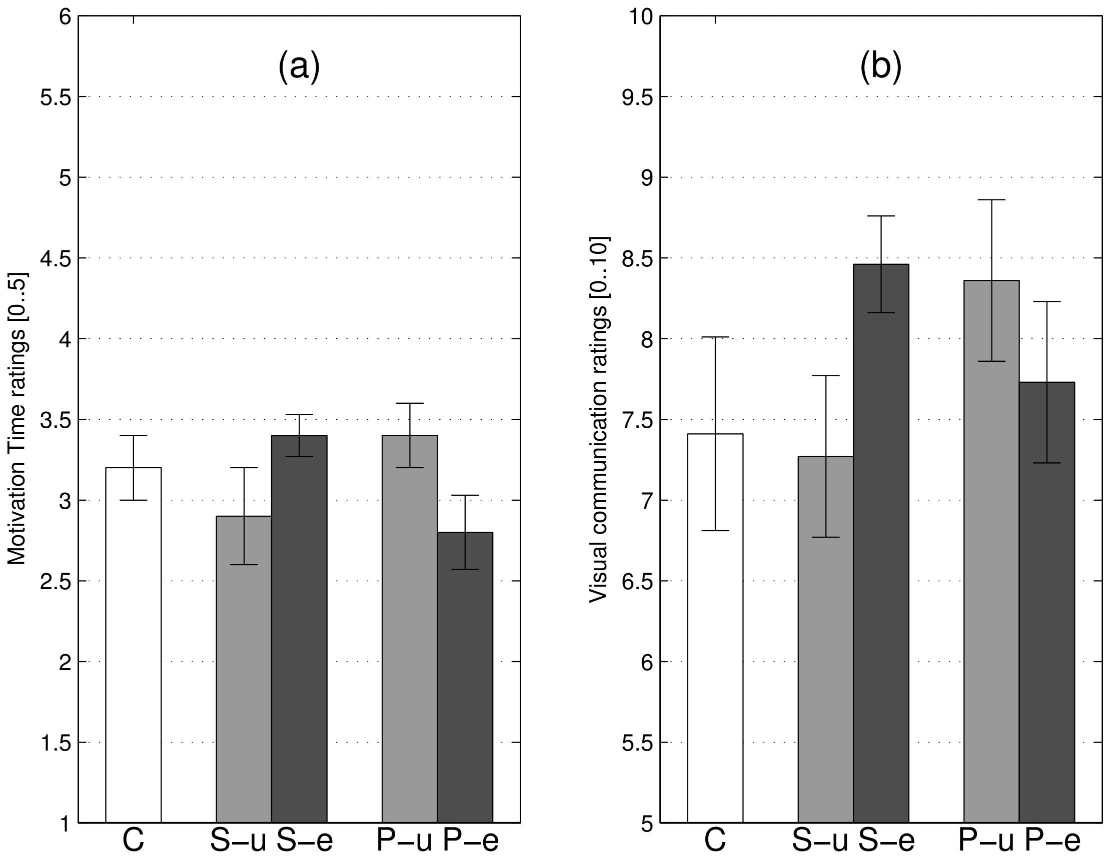


Figure 3.7.: The illustration of significant interaction patterns between "group" and "role-in-dyad" factors in the Physio and Sham groups for Motivation (a) and Visual communication ratings (b) for the Control group (C), Sham-users (S-u), Sham-emitters (S-e), Physio-users (P-u) and Physio-emitters (P-e). Error bars represent SEM values.

Measures	Physio rho (rho adjusted)	Sham rho (rho adjusted)
All measures together	0.80***	0.71
Arousal	0.25 (-)	0.38 (0.21)
Valence	0.13 (-)	-0.09 (-)
Curiosity (M1)	-0.0 (-)	0.66* (0.61)
Difficulty (M2)	0.36 (0.17)	0.21 (-)
Confidence (M3)	0.92*** (0.91)	0.19 (-)
Distribution of Control (M4)	0.30 (-)	0.54 (0.46)
Social Affinity (M5)	0.44 (0.31)	0.44 (0.33)
Interface Feedback (M6)	0.3 (-)	-0.3 (-)
Motivation Time (M7)	0.59 (0.52)	0.38 (0.22)
Satisfaction (M8)	-0.29 (-)	-0.17 (-)
Verbal Communication (M9)	0.35 (0.15)	0.30 (-)
Visual Communication (M10)	0.67* (0.62)	0.04 (-)
Gestural Communication (M11)	0.24 (-)	-0.25 (-)

Table 3.3.: The Pearson correlation coefficients (and observations adjusted values) showing the consistency within user-emitter dyad ratings. Separate data is shown for Physio and Sham groups (* marks significance at 0.05, ** - at 0.01, *** - at 0.005 level).

ing greater **Confidence** than Sham-dyads) and with the results of the role-in-dyad analysis (*Emitters* and *Users* showing high **Confidence** ratings in the Physio group). Finally, the **Visual Communication** measure showed a significant positive correlation within the *User-Emitter* dyad, but again only for the Physio group.

3.4.4. Correlations between motivational characteristics and valence-arousal ratings

All the results reported in previous sections showed no effects for the participants' SAM ratings of valence or arousal. However, we also decided to check the correlation between these subjective ratings of emotional state and each of the eleven motivational characteristics. Table 3.4 summarizes only the correlations that showed to be both significant and strong/moderate, with the intention of illustrating several recurrent patterns that complement the main findings of sections 3.4.1 and 3.4.2.

Second, two correlations between measures showed an opposite sign for the Control and the Physio-Emitter group. One was the relation between **Difficulty** and **Social Affinity** (see Figure 3.8). While this correlation was positive for the Control group

Measure	Control	Sham-users	Sham-emitters	Physio-Users	Physio-emitters
Arousal (A)		M8, 0.8*** M9, 0.7**		M1, 0.8** M11, 0.6*	M8, 0.6*
Valence (V)		M11, 0.5*	M1, 0.6* M4, -0.5*		M5, 0.7** M7, 0.7** M8, 0.8***
Curiosity (M1)	M11, 0.5*		V, 0.6* M4, -0.5*	A, 0.8** M6, -0.6*	
Difficulty (M2)	M5, 0.7**			M3, -0.8**	M5, -0.7* M10, -0.6*
Confidence (M3)				M2, -0.8** M7, 0.6*	M4, 0.6* M7, 0.6*
Distr. of Control (M4)		M7, 0.6*	V, -0.5* M1, -0.5* M9, 0.8**	M6, 0.6* M11, -0.7*	M3, 0.6* M6, 0.8** M7, 0.6*
Social Affinity (M5)	M2, 0.7**	M6, 0.6*			V, 0.7** M2, -0.7* M8, 0.6*
Interface Feedback (M6)		M5, 0.6*		M1, -0.6* M4, 0.6* M9, 0.6*	M4, 0.8** M7, 0.6*
Motivation Time (M7)		M4, 0.6* M11, 0.7*		M3, 0.6*	V, 0.7** M3, 0.6* M4, 0.6* M6, 0.6* M10, 0.6*
Satisfaction (M8)	M9, 0.6*	A, 0.8*** M9, 0.7**		M11, 0.6*	A, 0.6* V, 0.8*** M5, 0.6*
Verbal Comm. (M9)	M8, 0.6* M10, -0.6*	A, 0.7** M8, 0.7**	M4, 0.8**	M6, 0.6*	M10, 0.7**
Visual Comm. (M10)	M9, -0.6*				M2, -0.6* M7, 0.6* M9, 0.7**
Gestural Comm. (M11)	M1, 0.5*	V, 0.5* M7, 0.7*		A, 0.6* M4, -0.7* M8, 0.6*	

Table 3.4.: Significant, strong and moderate correlations between measures of motivational and emotional experience. The adjusted Pearson's rho is provided for each of five user types: Control group (C), Sham-users (S-u), Sham-emitters (S-e), Physio-users (P-u) and Physio-emitters (P-e); * marks 2-tailed significance at 0.05, ** - at 0.01, *** - at 0.005 level.

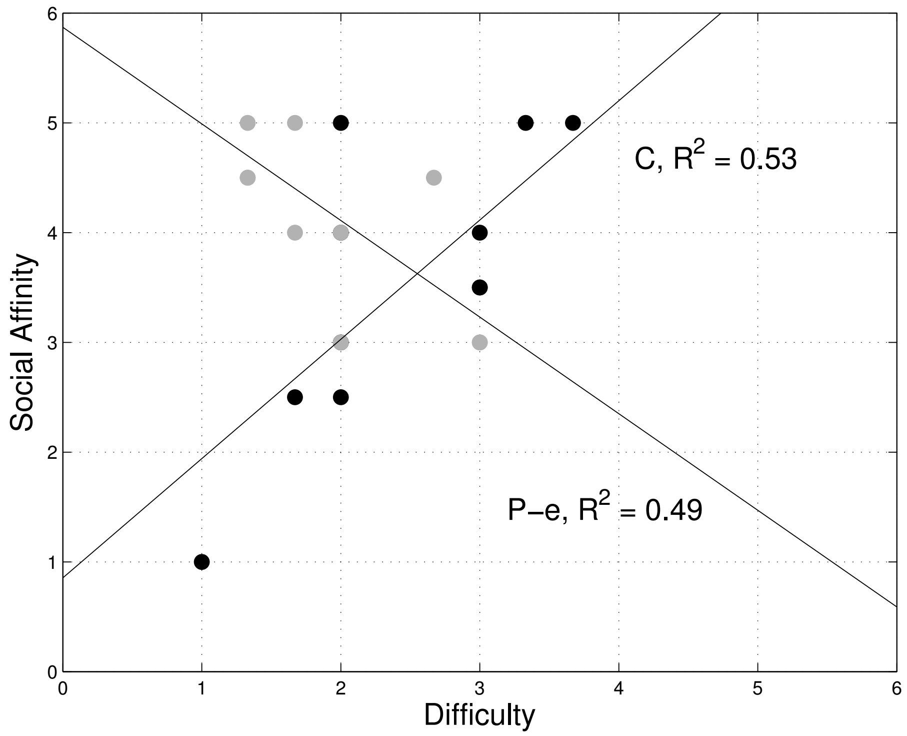


Figure 3.8.: The correlations between Social Affinity and task Difficulty measures for User in the Control group (C, black dots) and Physio-emitters (P-e, gray dots).

($adjrho = 0.7, p < 0.01$), Physio-emitters reported higher **Social Affinity** when **Difficulty** was lower ($adjrho = -0.7, p < 0.05$). The second relation was between **Visual** and **Verbal Communication**. Here Physio-emitters had a positive correlation between the two measures ($adjrho = 0.7, p < 0.01$). On the contrary, the participants in the Control group rated **Verbal Communication** as important while judging the **Visual Communication** aspects as low ($adjrho = -0.6, p < 0.05$).

Third, *Emitters* and *Users* in the Physio Group had two similar patterns of positive correlation between measures. One of them was between **Distribution of Control** and **Interface Feedback**, and the other between **Confidence** and **Motivation Time**. Finally, several correlations again highlighted the interaction effects between the "group" and "role-in-dyad" factors for **Motivation Time** and **Visual Communication** ratings that were described in Section 3.4.2 on page 90 (Figure 3.7 on page 93). Both Physio-emitters and Sham-users had high positive correlations between **Motivation Time** and **Distribution of**

Control, and between **Arousal** and **Satisfaction**. In addition, Sham-users shared similar patterns with the Control group: positive correlations between **Satisfaction** and **Verbal Communication**, and between **Satisfaction** and emotional ratings.

3.5. Discussion

The results of this study offer a number of insights on the effects of multimodal (gesture and implicit) interaction in user experience, by measuring multidimensional aspects of participants' motivation during a musical task. We compared subjective ratings of three groups of participants - the Control, the Physio and the Sham groups – using the Reactable and *b-Reactable*, a DMI that introduces physiology-based implicit interaction according to EEG and ECG measures. The experiment showed that (i) the motivation brought by *b-Reactable* significantly differs from the one of the Reactable; (ii) the dyads using *b-Reactable* had a different motivational experience depending on whether the physiological feedback was real or fake; and (iii) Physio-emitters and Physio-users had different motivational patterns.

Our first hypothesis postulated that *User-Emitter* dyads using *b-Reactable* would have a stronger motivation than users of the standard Reactable when performing the same collaborative musical tasks. Eleven motivation dimensions together with valence and arousal pictorial scales were measured through a post-experimental questionnaire. Indeed, several of these measures showed significantly higher ratings for *User-Emitter* dyads compared to the *Users* dyads in the Control group. Specifically, Physio-users had significantly higher ratings for **Confidence**, **Satisfaction** and **Gestural Communication** than conventional tabletop users. In addition, Physio-emitters had higher ratings for **Confidence,** and Sham-users for **Satisfaction** as compared to the Control group. Importantly, no ratings have been significantly lower than the ones given by Reactable *Users* in the Control group, showing that *b-Reactable* and the use of *physiopucks* do not affect participants' motivation negatively. It should also be noted that the observed effects are most likely caused by the presence of implicit interaction through *physiopucks*, and not due to the unusual experience of using EEG headsets, as participants in the Control group also wore headsets. We should also stress the fact that implicit interaction paradigms like the ones presented in this Chapter can also lead to self-adaptation (i.e. neuro and biofeedback) with participants trying to alter their EEG and ECG activity to match a given sound or tempo. However, this is more likely to happen after a number of training sessions, and this requires further investigation that goes beyond the current study.

The observed differences between *b-Reactable* and the standard Reactable could stem from *Emitters*, who would have a higher level of motivation compared to *Users* due to a new, implicit communication channel available to them. Indeed, taking a closer look at the differences between the four participant types and the Control group that performed the Reactable, we can see that motivation levels of *Users* and *Emitters* differed from each other. However, not fully in line with our predictions in hypothesis 1, Physio-emitters gave significantly higher ratings only for one measure (**Confidence**), while the Physio-users's ratings were significantly higher for **Confidence**, **Satisfaction** and **Gestural Communication**. In addition, Sham-users also had a significantly higher **Satisfaction** scores than the Control group. In other words, *Users'* motivation level in the *b-Reactable* groups also differed from the Control group. One possible explanation for Sham and Physio-users reporting the highest levels of **Satisfaction** compared to *Emitters* in the both groups, is because *Users* could only collaborate via hand gestures. *Users'* operation with *physiopucks* did not depend on the quality of physiological feedback, real or sham one, which could be only (or mostly) noticed by *Emitters*. Furthermore, *Users* in both groups had a less complex and a more conventional control of the interface. On the other hand, *Emitters* might not have reached high levels of satisfaction given the expressive limitations of the physiological control and the quality of the sonic designs applied in *b-Reactable*. Whereas *Users* had a quite varied repertoire of sound generators to play with, *Emitters* could have felt tied to white noise audification generated by the EEG-based *physiopuck*. Perhaps different sonic design strategies would result in better satisfaction ratings from *Emitters*. As *Emitters* might not have been able to achieve high expressiveness with the provided audification, we have to explore whether their satisfaction change in the case of applying other sonic designs for implicit interaction.

When comparing the differences between the Physio and Control groups, a few issues should also be stressed. First, Physio-users showed higher ratings of **Gestural Communication**, as opposed to the **Verbal Communication** preference showed by the dyads within the Control group. Second, **Social Affinity** between Physio-emitters and their partners was highly correlated with reported **positive Valence**. Interestingly, the **Social Affinity** ratings were higher when *Emitters* reported lower difficulty of the task, while an opposite correlation could be seen in the Control group, in which higher scores on **Difficulty** level matched the stronger affinity with the partner.

Our second hypothesis was that the Physio condition would provide higher levels of motivation aspects than the Sham one, due to dyads perceiving the difference in the implicit feedback quality (i.e. real-time vs. pre-recorded). Indeed, the multivariate comparison with all measures combined together showed a significant difference between the two experimental groups. Specifically, two motivation aspects were significantly higher in ratings for the Physio than for the Sham group, namely **Confidence** and **Gestural Communication**. This result suggests that physiology-driven implicit interaction could be a relevant technique to enhance quality of collaboration and non-verbal communication during multi-user music performance, through sonic or multimodal feedback.

In addition, a significant interaction was observed where Sham-users and Physio-emitters showed lower **Motivation Time** ratings than Physio-users and Sham-emitters respectively. The observed interaction may be explained by two factors: (i) task complexity and (ii) the use of sham recordings. The two tasks differed in complexity, with the second one being more difficult to solve than the first one (as it involved more pucks and deeper configurations). The low **Motivation Time** measures in Sham-users could be explained by the challenge that could imply solving tasks with an increasing difficulty with a partner (Sham-emitter), whose physiology-driven control parameters were not responding according to the Sham-emitter physiological state (i.e. working with pre-recorded EEG and ECG measures). This did not happen in the Physio Group, as Physio-users were collaborating with a partner whose physiology is driving the *physiopucks*. In the case of Physio-emitters, solving the second task could have been a significant challenge (especially without previous training sessions), and thus diminishing their motivation towards the end of the experiment. In sum, this interaction may suggest that it was easier for Sham-Emitters and Physio-users to solve the tasks, and that was reflected in higher **Motivation Time** ratings.

The direct between-groups comparisons were complemented by the results from within dyads correlations that showed whether *Emitter* and *User* shared similar motivational ratings. The dyads' ratings from the Physio group showed high and significant correlations for the **Confidence** and **Visual Communication**. This again stresses the fact that Physio dyads effectively collaborated and relied on gestural or visual communication resources rather than on speech. This effect does not mean that there is a trade-off between communication channels (i.e. more gestural communication, less speech) but a participant preference for one or another. Conversely, Sham dyads' ratings were only correlated for **Curiosity** measures. Finally, the ratings of **Distribution of Control** and **Interface Feedback**, **Confidence** and **Motivation Time** were positively and significantly correlated for both *Users* and *Emitters* in the Physio but not in the Sham group. Although we did not address this directly, these results suggest that there might have been some transfer of subjective experience, like confidence or curiosity, between the partners working together at the task.

A specific effect could be also observed where Sham-users and Physio-emitters had high positive correlations between the ratings of **Motivation Time** and **Distribution of Control**, and between the **Arousal** and **Satisfaction** ratings. Again, one possible explanation for this result is that Sham-users and Physio-emitters had a clearer task compared to their partners in the dyad. In the Sham group, Sham-emitters perceived and tried to solve the problem with the erroneous feedback, which could explain their high ratings for **Visual Communication**. Sham-users in this case were in the control of the situation and reported high levels of **Satisfaction**. On the contrary, Physio-emitters were the ones who had the most satisfying experience, as shown by the highest number of motivation measures correlated with **positive Valence**. Taken together, it is clear that the Physio and Sham groups differed significantly on several motivation dimensions.

Both variance and correlation analyses support our second hypothesis and suggest that participants in both groups were able to perceive the quality (real-time vs. sham feedback) of physiological activity displayed sonically and visually on the *b-Reactable*. Corroborating this result, the significant correlations of **Curiosity** ratings for Sham-dyads together with their low levels of **Confidence** can be explained by participants perceiving the sham feedback from the interface or, at least, having a certain inability to implicitly interact through the physiological channels. As a sign of coping with this situation, Sham dyads had a high correlation between the **Verbal Communication** and **Satisfaction** scores, and the scores of **Distribution of Control** and **Arousal**. In sum, the very fact of preferring gestural communication shows the potential of implicit indirect interaction for reinforcing a more intuitive, seamless and body-centered multi-user interaction.

It should be noted that that the fake nature of sham feedback could be perceived more easily via the audiovisual display of ECG activity. Although we did not assess directly the difference between the perception of EEG and ECG, participants might perceive these two signals differently. Unlike EEG feedback, ECG display was more direct and intuitive (a pulse blasting from the center of the tabletop interface) making it easier to perceive and to compare with the Emitters' own body states. Moreover, the ECG values were displayed in the upper right corner of the *physiopuck*.

Regarding the validity of our DMI (*b-Reactable*), the experiment showed that it could be used as a platform for collaborative music re-creation. Previous research on collaborative learning has highlighted that aspects like confidence, social affinity and distribution of control are of utter importance for increasing learning motivation [Jones and Issroff, 2005]. All of these measures have been sensitive to our between-subject manipulation, showing that *b-Reactable* can increase motivation as compared to standard tabletop systems based on hand gesture control. We also see that aspects like Motivation Time and preference for feedback type (i.e. visual, gestural, verbal) can be used as informative measures when studying participants' motivation during collaborative musical tasks. In our case, the proposed combination of implicit and explicit interaction fostered non-verbal communication and body-centric interactions as compared to gesture-only tabletop system.

When running the experiment we noticed that almost all participants were surprised and curious about the physiological sensors, even more than about the tabletop system itself. These reactions, together with the fact that the *Users'* task was not directly affected by the quality of the physiology-based feedback, could account for an increase in *User's* **Satisfaction**, keeping their ratings high even within the Sham group. Hence, **Satisfaction** ratings may have captured a sort of wow effect, representing the participant's impression of a "cool" and novel interface. However, it is important to note that while *Users* themselves were not connected to the system, the fact of being able to use their partners 'physiology through *physiopucks* still made the interface very exciting for them. The lower scores of the Control group, where both partners also had EEG headsets, further support this explanation. Since we are not aware of any other works that use physiology driven tangible objects in collaborative tasks (there are, however, other studies that combine physiology and tangible interaction, like Hermann et al. [2007], but without addressing music or multi-user experiences) many aspects of this unique *User-Emitter* situation are still to be tested concerning the *anthropomorphic* potential of the *physiopucks*' concept (e.g. "touching someone's heart"). Other uses of multitouch display may also bring interesting insights about collaborative scenarios based physiology-driven objects, e.g., transformation of parallel displays for teamwork, as the case of the CityWall public installation by Peltonen et al. [2008].

It should be noted that the effects found in our study might be of temporal nature, thus future experiments should address the impact of prolonged use of physiopucks. However, it should be noted that this kind of physiology-based interaction is likely to produce subjective experiences different from gesture-based control. Consider, for example, the BRAAMHS project, a novel musical instrument based on fNIRS (functional near-infrared spectroscopy) that adapts implicitly to users' changing cognitive state during musical improvisation Yuksel et al. [2015]. As one of the users commented, "I couldn't tell if I was influencing them [patterns] but for some reason it didn't feel random, I don't know why".

3.6. Conclusions & next steps

This chapter has presented the *b-Reactable,* a first system prototype supporting both implicit and gesture interaction for sound generation and control. Through this prototype we have studied how physiology-driven implicit sonic interactions can affect user motivation in a musical task, compared to the use of a gesture-only tabletop system (the Reactable). We have compared multiple dimensions of motivation between three experimental groups -Control, Physio and Sham - using the Reactable and *b-Reactable* with either real and fake physiology feedback. The experiment has shown that the motivation brought by *b-Reactable* is stronger than the one of the tangible interface just based on gestural inputs in terms of **Confidence**. On the other hand, the introduction of a fake (sham) physiological feedback significantly changes **Confidence** and **Communication** of participants. Importantly, Physio-emitters have shown different experiences than their partners, e.g, in terms of positive emotions **(Valence)** and **Social Affinity**. These results strongly support the potential of physiology-based interfaces and implicit interaction for improving single and multi user HCI, within or beyond the musical domain. Further developments in this regard could therefore explore how implicit interaction could widen multiuser communication, foster user entrainment by perceiving brain and body signals via *physiopucks* after training, or increase interpersonal synchronization in computer supported cooperative work (CSCW).

The results of this first study also allow us to formalize 3 specific aspects of sonic interaction design applied to implicit PhyComp that have to be further investigated:

- Although the sonic interactions presented in our first study had a significant effect on participants and were properly perceived by Physio-emitters, we cannot claim at this stage that they perform better than other sonic displays strategies (i.e. parameter mapping, musical mappings) for perceiving implicit physiological states. On the other hand, the fact of providing a fix, pre-defined physiology-to-sound mapping that does not consider user perception in its design could potentially affect user experience. It is highly likely that an *ad-hoc* mapping will not be perceived in the same manner by every user. Therefore, the aspect of **perceptualization,** understood as the process of associating a given display strategy (in this case sound) to the psychophysiological state that acts as the input for its rendering, has to be further explored. In line with Jovanov's studies on perceptualization of physiological data Jovanov et al. [1999], it defines how well a sonic design represents a given implicit physiological state, aiding user perception.
- Whereas the SID strategies implemented for this study were simple and straight-

forward (i.e. audification of EEG and BPM control through ECG) the fact of using two biosignals simultaneously made difficult to identify the specific impact that each of these had in the users during the experiment. Moreover, we cannot claim that such simple physiology-to-sound mappings are better than more complex mappings (e.g. parameter or musical mappings) for perceptualizing implicit PhyComp. Therefore, the role of **mapping complexity** in physiology-based sonic interaction, understood as *the number of physiological streams and sound parameters used in a given SID strategy*, has o be further explored. This will require an experimental design and SID strategy that allow to use several physiological features for informing adaptive systems (like in this study) but also to evaluate the effect of of each of them in isolation.

- Although the multimodality of the *b-Reactable* accounts for rich musical operations (as it is based on an extensively used DMI) the physiology-driven sonic interactions embedded in the *physiopucks* were rather simple in terms of expressiveness (white noise and BPM control of the overall music composition). In this study we did not directly assesses how meaningful these concrete SIDs were in the context of music performance. Therefore, the aspect of **meaningfulness** in the NIME context has to be explored further, understood as *the potential of physiology-driven implicit interaction for being perceived as an expressive component of the DMI, through which the player can produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to.*

These three aspects will guide our next steps on the exploration of implicit, physiology-driven sonic interaction. Concretely, in the next chapter we tackle *perceptualization* and *mapping complexity*. We will thus implement more sophisticated algorithms and signal processing techniques to measure and classify high-level EEG features (i.e. relaxation and affective states). We also proceed to extend our sound engine to support multi-parameter sonic display, musical mappings, and end-user configurations for providing personalized sonic interactions. This sound engine allows multiple mappings between physiological features and sound, ranging from direct audification to more complex paradigms, such as parameter-based sonifications and musical mappings (see Section 2.10 on page 46 for a summary on these techniques). In order to assess the *perceptualization* of these different SIDs, we temporally move away from the musical domain to apply our sonic designs in a perception-based context: neurofeedback. We design a set of experiments were participants are exposed to their own physiological activity (displayed through different SID strategies) in order to learn about their psychophysiological state (e.g. relaxation) and aiming at controlling this state by training.

To carry on evaluations through neurofeedback training imply to focus on EEG features. However, EEG estimations are more complex in this case (e.g looking at alpha/theta ratios) as well as the sonic interactions presented to users (multi parametric and musical mappings). Finally, in the next study we will also allow participants to personalize SIDs, to produce sound mappings tailored to their own perception and subjective preference.

4. Perceptualization and Mapping Complexity

This Chapter presents a set of studies meant to address two of the SID aspects that emerged from our previous experiment with *b-Reactable*: *perceptualization* and *mapping complexity* in physiology-based implicit sonic interaction. The goals of this chapter are (i) to determine what types of sonic designs perform best in representing a given implicit physiological state (e.g. relaxation) according to end-user perception (*perceptualization*) and (ii) to determine whether *mapping complexity* and personalization by end-users play a role in the perceptualization process. We address these issues separately by mean of two experiments based on neurofeedback training. The first one assesses the *perceptualization* quality of the most used sonic designs for displaying EEG activity (as suggested by our literature review). The evaluation is based on end-user perception (both subjective and through physiological measures) of own relaxation states estimated from EEG alpha activity. The second study leverages on the findings of the first one to implement a sound engine capable of (i) generating designs of different *complexity* (both in terms of physiological streams and sound parameters) and (ii) able to be personalized by end-users. Both *mapping complexity* and personalization are tested in alpha/theta neurofeedback training, collecting subjective and objective (EEG) measures of relaxation. Results from the first experiment suggest that parameter mapping sonification and musical mappings are good candidates for *perceptualizing* implicit physiological states, whereas the second experiment provides empirical evidence about the positive effect of *mapping complexity* and end-user personalization in the *perceptualization* of sonic designs for physiology-based implicit interaction. The studies also demonstrate that personalization becomes less instrumental when multiple physiological features are displayed through sound.

4.1. Introduction and motivation

Our previous study on the *b-Reactable* revealed the positive impact of physiology-based implicit interaction in user experience when expressed through sound. However, it also highlighted specific design issues that have to be further investigated. Under the light of our previous experiment, on the other hand, we still cannot claim that the implemented sonic design (i.e. audification) performs better than other sonic strategies (i.e. parameter mapping, musical mappings) for perceiving implicit physiological states. On the other hand, the fact of providing a pre-defined and ad-hoc physiology-to-sound mapping could potentially affect user experience, as it does not consider user perception in the design process.

One way to address these issues is to assess the perceptual link between a given sonic display strategy and the psychophysiological state that acts as its input. In line with Jovanov's studies [1999], we define this aspect as **perceptualization,** which basically describes how well a sonic design represents a given implicit physiological state, aiding user perception.

The SID strategies implemented in the *b-Reactable* were simple and straightforward (i.e. audification of EEG and BPM control through ECG) and we did not compare its performance with more sophisticated approaches (e.g. parameter or musical mappings). Therefore, the role of **mapping complexity** has to be further explored. This aspect specifically refers to the number of physiological streams and sound parameters used in a given SID strategy[1].

Perceptualization and *mapping complexity* of SID applied to implicit PhyComp are directly related to what we have already discussed in 3.6 on page 102: the selection of EEG-to-sound mappings is always a difficult task involving perceptual and aesthetic trade-offs, and it heavily depends on the application domain and goal. On the one hand, a simple EEG-to-sound mapping (as the one applied in the *b-Reactable*) accounts for a direct perception of changes in implicit physiological states, producing an almost reversible signal. However, such sonic designs tend to be unnatural, dull and not well suited for hearing out multiple, simultaneous physiological events. In short, they might convey good *perceptualization* quality*,* but they lack *complexity* (and most likely expressiveness). On the other hand, more indirect and complex mappings account for more naturalistic, rich and perceptually pleasant sounds. This is the case of parameter

[1]In Chapter 3 we also identified *meaningfulness* as a third aspect to be further explored in the design of implicit sonic interaction for PhyComp. However, as it requires to deploy SID in an expressive and creative context (e.g. music performance) it will be addressed in a separated study (see Chapter 6).

mapping sonification (see Section 2.4.1) or musical interfaces as the ones presented in Section 2.5, which were in fact designed with the aim of aiding music expression. But by applying more complex (and arbitrary) mappings, these sonic designs tend to hide or mask the nature of the physiological features that feed the sonic operations (therefore compromising the *perceptualization* of implicit states). In this context, the evaluation of *perceptualization* of sonic designs from an end-users perspective (i.e. directly assessing user perception) becomes a crucial factor for validating SID for implicit PhyComp, and to further explore its application in a given HCI context.

In this chapter we therefore tackle *perceptualization* and *mapping complexity* by running two studies. The first one addresses *perceptualization* by presenting sonic designs that vary in *complexity* (ranging from simple, direct strategies to more complex sound designs) and comparing them according to end-user perception in neurofeedback sessions, where sonic designs are used to display participant relaxation states estimated from EEG alpha activity. Following the literature review presented in Chapter 2 (see Section 2.6.3 on page 61) we choose the sonic designs most used for displaying real time EEG data: parameter mapping sonification (of different complexities) and musical mapping. Twelve participants take part on this experiment, where both subjective and EEG measures of user perception are collected to evaluate the performance of each sonic design in terms of *perceptualization* (relaxation perception and representation) and neurofeedback effect (relaxation state after training).

The second study leverages on the findings of experiment 1 to implement a sonic engine capable of generating designs of different *mapping complexity,* both in terms of physiological streams (i.e. number of EEG features being displayed) and sound parameters. To further explore *perceptualization,* experiment 2 also introduces personalization of physiology-to-sound mappings by end-users. This approach is tested through a more complex neurofeedback paradigm (alpha/theta training) with 31 participants, involving longer training sessions than experiment 1.

4.2. Experiment 1: assessing perceptualization of sonic designs

As discussed in the literature review (Chapter 2), the *perceptualization* of sonic designs for implicit PhyComp has not been widely explored. It is still not clear what type of sonic strategies perform best in terms of representing a given implicit physiological state, aiding its perception by the end-user (i.e. the listener). Therefore, this experiment has been designed for determining the *perceptualization* quality of diverse sonic designs for representing implicit physiological states (concretely relaxation) by assessing subjective and objective perception. Following the insights gathered in our literature review (Section 2.6) we selected three different sonic designs for the experiment: a simple and direct sound synthesis method (similar to the one implemented in the *b-Reactable*), parameter mapping sonification (the most used in the field of EEG sonic display) and musical mappings (the most explored within NIME and HCI). User perception was evaluated during alpha neurofeedback training, one of the most common neurofeedback paradigms which targets an increase in the EEG alpha band ($8-13Hz$) to reduce anxiety levels and increase relaxation states [Angelakis et al., 2007]. Levels of alpha activity in such training are often translated into sound-based feedback, either continuous or threshold-based, as it predominantly originates from the occipital lobe during wakeful relaxation with closed eyes [Hardt and Kamiya, 1978]. All sound designs were tested in both neurofeedback and offline (sham) conditions, the last one being based on the pre-recorded physiological data of a volunteer performing the same neurofeedback training as the participants in the experiment.

Three main hypotheses drove the study:

- **H1**: Simple and direct sonic designs will better *represent* relaxation states estimated through EEG, as they directly translate EEG data into sound.
- **H2**: Sonic designs with more complex mappings, such as parameter and musical mappings, will reach greater *relaxation effect*.
- **H3**: Relaxation will be *induced* in participants as result of the neurofeedback training, being reflected in both subjective and EEG measures.

4.2.1. Material & methods

System design

In terms of system design, specific technical requirements need to be satisfied. Since our research deals with real-time sonic display, it is of the upmost importance that the processing time for the entire signal chain -from EEG acquisition to auditory display- be minimized. Additionally, the EEG data processing need to be done at a constant rate in order to properly represent participant's relaxation states and promote neurofeedback training. Figure 4.1 shows the overall system architecture for experiment 1. The system is mainly composed of two modules: one devoted to EEG signal acquisition and processing, and a second module dedicated to the sonic display of the selected EEG features.

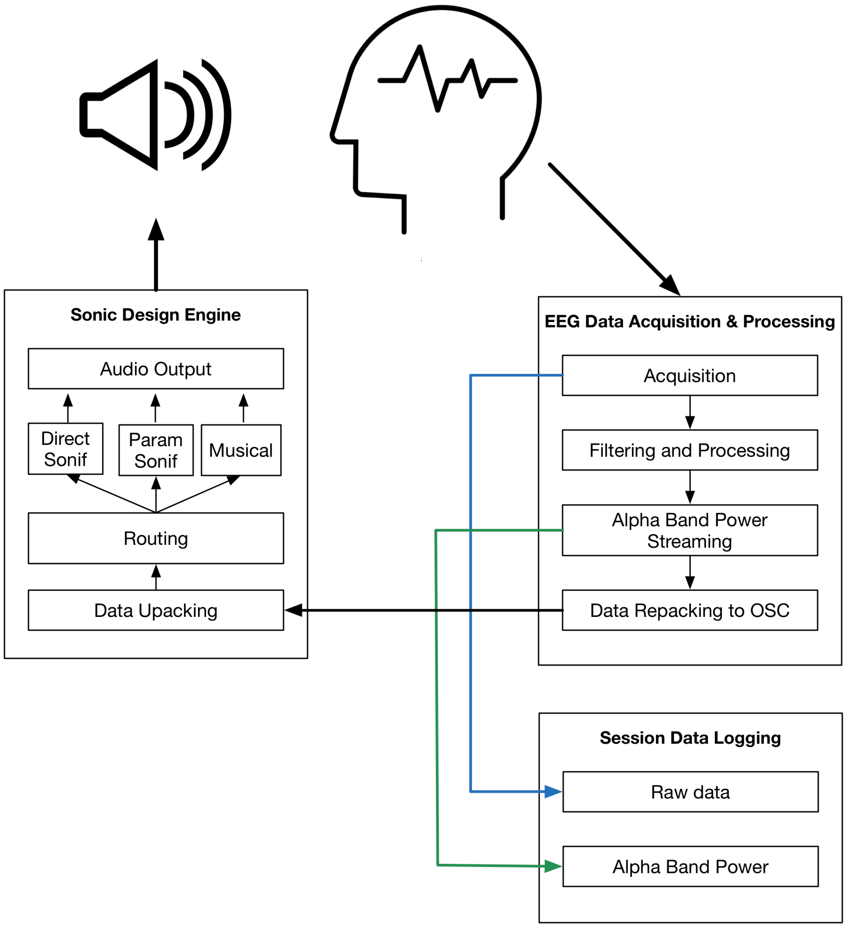


Figure 4.1.: Graphical overview of the system architecture for experiment 1.

Each module runs in a different computer (Apple MacBook Pro laptop, late 2013) and communicates through a middleware, as described below.

EEG signal acquisition and processing

Figure 4.2 presents a diagram illustrating each step of EEG signal acquisition and processing. Participants' EEG activity is acquired using an Emotiv EPOC[2], a 14-channel wireless and non-invasive neuroheadset that accounts for a fast and ergonomic placement. The EPOC is one of the leaders in low-cost consumer EEG sensing devices, and although previous studies have shown it unreliable for specific BCI applications (since

[2]`https://emotiv.com/epoc.php` (accessed on October, 2015).

its signal-to-noise levels are lower than in professional devices)[Duvinage et al., 2013], it has been successfully used for monitoring [Rodríguez et al., 2013], which is the process underlying neurofeedback. The EPOC's electrodes require a saline solution to improve conductivity, and their placement is related to the 10-20 international system [Homan et al., 1987]. Following the literature on alpha activity estimation [Kropotov, 2010, Angelakis et al., 2007], we acquire raw EEG data from 6 channels placed at O1, O2, P7, P8, T7, T8 of the 10-20 international system. Such placement accounts for a robust signal acquisition, reducing artifacts caused by facial muscle contractions.

Raw EEG data is wirelessly streamed at 128 samples per second to a signal processing module implemented in OpenVibe[3]. The six above mentioned channels are selected and epochs of 15ms are created for further processing. The incoming signal is later filtered using 4th order Butterworth filters to isolate the alpha band $(7 - 13Hz)$ and the full EEG bandwidth $(0.1 - 40Hz)$. The values of both signals are squared and spatial filters are applied to compute a single value from the six incoming input channels. At a final stage, relative alpha power is estimated by dividing the estimated alpha power by the full EEG band power.

Communication from OpenVibe to the sonic design engine is established using a Virtual Reality Peripheral Network (VRPN) server/client structure that outputs OpenSound-Control (OSC) [Wright, 2005] messages in real time. This software was designed as a middleware that reads from the VRPN server within OpenVibe and receives two data streams simultaneously [Vamvakousis and Ramirez, 2014]. OSC packages are then sent to the sonic design engine.

Sonic design engine

Like in the case of the *b-Reactable* prototype (see Chapter 3) the engine for displaying EEG activity through sound was implemented in Pure Data (Pd) [Puckette et al., 1996], a real-time graphical programming environment for audio synthesis and processing. The engine routes the incoming alpha values (received as OSC messages from the communication middleware) so that they can be displayed according to the desired sonic design.

Once the EEG data is received in Pd, it is then sent simultaneously to three processing blocks, corresponding to the sonic designs (direct sonification, parameter mapping sonification, and musical mapping) that will be discussed in detail in the following section. It is important to note that the engine and the implemented sonic designs work

[3] `http://openvibe.inria.fr` (accessed on October, 2015).

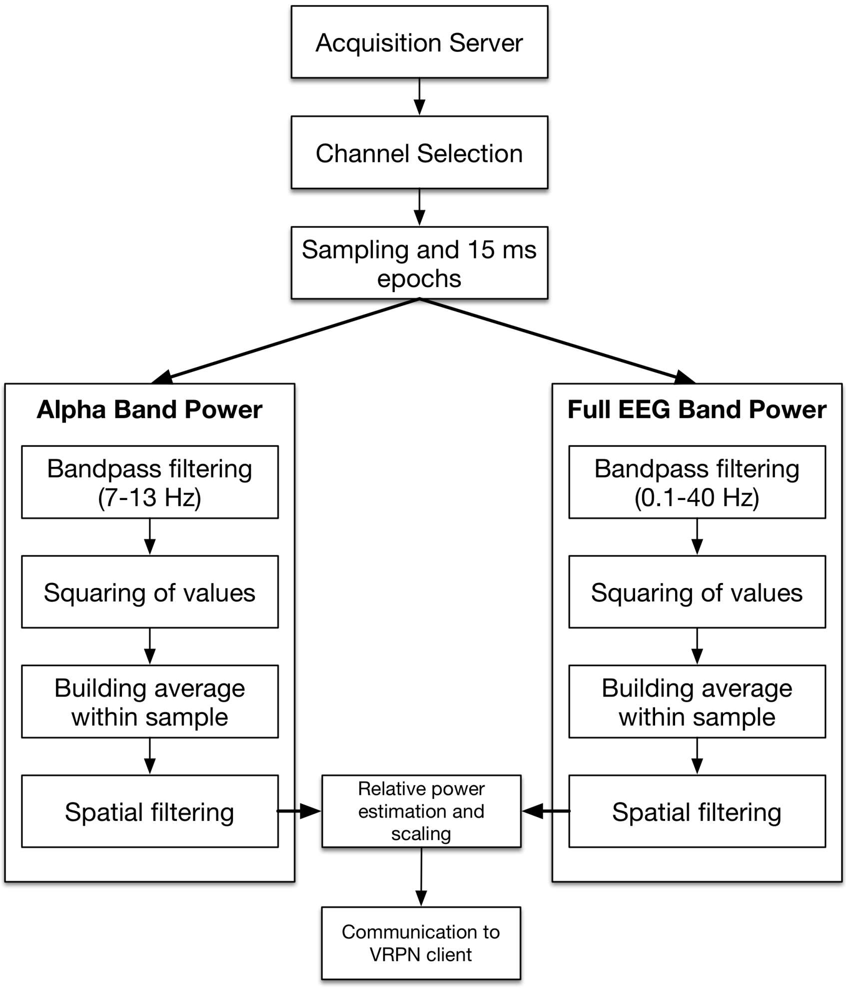


Figure 4.2.: Flowchart illustrating EEG signal acquisition and processing.

with both real-time EEG data (online mode), and prerecorded data (offline mode). This functionality allows testing sonic designs without the need of a real-time stream coming from an EEG headset.

Sonic designs and mapping criteria

For this experiment we implemented 3 sonic designs that represent widely used techniques in the field of EEG auditory display. However, as pointed out by our review and by other recent assessments such as Williams et al. [2014], a great variety of sound and musical features have been used to perceive and/or induce states, such as emotions or relaxation. The mapping approaches employed by such systems vary quite significantly, with little agreement on which sound features are essential, desirable, dispensable, etc. Thus, deducing an ubiquitous sound feature-set for representing and inducing relaxation is not a straight forward task. In order to overcome this limitation, we defined a mapping criteria based on previous studies on relaxation neurofeedback training, and on the ontology for affective algorithmic composition (AAC) systems proposed by Williams et al. [2014], where pitch, rhythm and timbre are proposed as the most common sound features found in AAC. By following this approach, we established five mapping guidelines for all our sonic designs:

- *Positive reinforcement*: the main objective of neurofeedback is to reach a greater awareness and, eventually, a voluntary control of physiological processes. The success of neurofeedback therapy is thus related to the effectiveness of this learning and self-regulation process. In this regard, *positive reinforcement* has shown to be an important factor for successful self-regulation, both at conscious and unconscious levels [Siniatchkin et al., 2000]. Following this premise, our sonic feedback will be designed with the goal of reinforcing the target state (i.e. relaxation) in a positive manner: higher relaxation estimations will yield more relaxing sounds.
- *Pitch variation*: as suggested by previous psychoacoustic and biofeedback studies [Batty et al., 2006, Budzynski and Stoyva, 1969, Hevner, 1937] pitch center (high, low) has shown to be correlated with a range of emotional descriptors, including arousal. Leveraging on this research, our physiology-to-sound mappings will reflect an inverse relationship between the measured relaxation state and pitch frequency. That is, when relaxation increases the pitch becomes lower, and when relaxation decreases the pitch becomes higher.
- *Harmonicity*: harmonicity is a component that suggest an inter-relationship between pitch and timbre. Although pitch content can influence the harmonicity and

noisiness of a musical timbre, the majority of the literature identifies harmonicity as a contributory timbral component rather than a pitch derived one (see Williams et al. [2014] for a review). Previous studies in emotional and psychophysiological response to music have shown that the level of relaxation increases with an increase in harmonicity of tones [van der Zwaag et al., 2011, Juslin, 1997]. However, there is also evidence of complex and dissonant harmony being correlated with low relaxation (i.e. excitement and arousal) [Williams et al., 2014]. Therefore, our sonic designs will seek harmonicity (when possible) but avoiding complex harmony structures.

- *Tempo*: previous investigations have shown a correlation between low tempo and low arousal [Hevner, 1937, Thompson and Robitaille, 1992, Collins, 1989]. Therefore our sonic designs will follow the same rule.
- *Loudness*: although previous research has established a relationship between loudness and affective states such as anger and happiness [Gabrielsson and Juslin, 1996] there is no agreement on the role of loudness in perceived and/or induce relaxation through neurofeedback. Therefore loudness will be a parameter defined by the participant prior the beginning of the experiment.

Following this mapping criteria, we have implemented 3 sonic designs[4]:

1. **Direct sonification**: this technique involves the direct translation of data into sound, and represents our starting point for designing a simple auditory display strategy. For implementing it, we use FM modulation [Chowning, 1973] as a synthesis method, as it is fairly simple from a signal processing standpoint, but it can result in harmonically-rich waveforms. FM synthesis works on the principle that one signal (the *carrier*) is modulated by a second signal (the *modulator*). The amount of alteration in the resulting signal is dependent on the amplitude of the modulator. For the sake of synthesizing a sound that was not unpleasant to listen to, we choose a simple FM technique with one carrier and one modulator. Since only one stream of values is provided by the EEG acquisition module (relative alpha power), it affects both the carrier frequency and the modulator, but in different ways. For the carrier, rather than immediately scaling the input values to an audible range, we instead subtract them from a fixed initial value of $450Hz$. In this way, a higher input value yields to a lower carrier frequency. This is done following the above mentioned mapping criteria, which establishes that higher EEG input values (associated with higher relaxation) should convey low pitch sound.

[4] An excerpt of each sonic design can be found at `http://www.dtic.upf.edu/~smealla/PhD_Material/Sounds.html`

For the modulator, the mean of the input signal is calculated in real-time and used as the frequency for the modulator. The *depth* or amount of amplitude given to the modulator is preset and left fixed throughout our tests and experiments, providing enough of the modulator signal to create harmonic richness, without promoting an uncomfortable listening experience. In fact, rapidly changing the *depth* amount by tying it to the input value tends to generate uncomfortable sounds. The resulting sound is a dynamic rapidly-changing modulated sine wave, which has harmonic richness, but does not lead to an uncomfortable listening experience.

2. **Parameter mapping sonification**: this strategy involves features of the input signal being mapped to some features of the resultant output sound [Hinterberger and Baier, 2005]. In our case, the data is first sent to a feature extractor module and selected features are then mapped to different sound parameters. The sound in this case is an oscillator put through a voltage controlled filter (VCF). When data is streamed, the feature extractor detects each time a new local *maxima* or *minima* occurs. These values affect separate yet similar output sounds, which differ mainly in the way that the center frequency of the VCF is determined. The sounds, when triggered, take on the following parameters as determined by the data:
    - *Sustain*: the distance in time between each *maxima* or *minima* is used to calculate the sustain of the resulting sounds' amplitude envelope. The sonification is designed so that the output will have a generally fast envelope, but the length of the sustain is somewhat variable within a predetermined range.
    - *Center frequency*: by keeping track of how many *maxima* or *minima* occurs each second, a crude estimation of the input signal's frequency is calculated. This value is scaled, and the frequency of the *maxima* is used to modulate the center frequency of one sounds VCF, while the *minima* frequency modulates the center frequency of the second VCF.
3. **Musical mapping**: to achieve a more inherently musical-sounding output, we map the incoming stream of data to frequency values and trigger bursts of sounds set to these notes. To favor more musical outputs, input data is scaled to integers to trigger recognized MIDI notes. We restrict the MIDI output to values associated with certain notes. In this case, the 7 diatonic notes of a major scale (minus the octave) are used. The input stream is scaled to select one of these notes at a time, and the input data are also used to send groups of MIDI messages corresponding to the diatonic chords within the same major scale. By doing so, the data create a melodic line in addition to the underlying chords. Since these all

belonged to the same scale, a relatively consonant, if not melodically memorable, composition is created in real time. In terms of rhythm, separate approaches are taken for the melodic line and the chords. The notes from the melody are played in euclidian rhythms [Toussaint et al., 2005], which have onsets that are as evenly divided as possible given the input parameters. By adjusting these parameters, the tempo of the output can be adjusted, as can the number of beats per measure and how many notes can happen in each measure as a maximum. Meanwhile, the chords' output is set to be triggered every 1000, 1500, 2500 or 3000 milliseconds, based on a random selection. At the MIDI output, the Ableton Live[5] *Operator* is used, a versatile synthesizer that combines classic analog sounds and frequency modulation synthesis. Separate patches are chosen for the melody and the chords, and modifications are made from the stock patches to create longer reverb times and decay times, promoting a pleasant sound. In fact, two identical sounds are used for the chords in Ableton, and their output is alternated within Pd, so that each new triggered chord would not prematurely cut off the decay of the previous chord. With these elements combined, a seemingly random, although actually rather complex, musical output is created. As in the case of the previous sonic designs, higher input values produced lower notes within the scale.

All sonic designs were handled by a MacBook Pro laptop (late 2013), and resulting sounds were displayed using an M-Audio Fast Track Pro sound card, and a pair of Genelec 8010A studio monitors.

Participants

The tested sample consisted of 12 students (5 females), from Universitat Pompeu Fabra, Barcelona. Their ages ranged from 23 to 36 years, with the mean being 28.1 ($SD = 1.4$). The participants did not have prior knowledge on the experiment, although several were familiar with EEG and the Emotiv EPOC in particular. None had participated in experiments involving EEG sonic display before. The study was conducted in accordance with the Declaration of Helsinki.

Experimental procedure and measures

Before the experiment, each participant signed a consent form, provided demographic data (age, gender, place of birth and academic background) and was exposed to a 5-

[5] `https://www.ableton.com` (accessed on October, 2015)

Block	**Neurofeedback Session**					
	S1	S2	S3	S4	S5	S6
B1	Direct_{RT}	Direct_{S}	Parameter_{RT}	Parameter_{S}	Musical_{RT}	Musical_{S}
B2	Direct_{RT}	Direct_{S}	Parameter_{RT}	Parameter_{S}	Musical_{RT}	Musical_{S}

Table 4.1.: Procedure for Experiment 1. Each experiment consisted of two blocks (B1 and B2) of six neurofeedback sessions (S1-S6), one per type of sonic display in real time (RT) and sham (S) conditions. All sessions were randomized and separated by a 1-minute break. Blocks were separated by a 2-minute break.

minute relaxation induction, using the sound of sea waves, after the EEG headset was mounted.

The experiment consisted of two blocks separated by a two-minute break (see Table 4.1). In each block participants were exposed to six neurofeedback session of 2 minutes length. These six conditions included the three types sonic designs, triggered either by participants EEG (real time) or by pre-recorded EEG (sham) acquired from a volunteer that followed the same experimental procedure. All sessions were randomized and were followed by a one-minute break where the participant was asked to answer three questions.

Three subjectives measures (5-point Likert scale) were used to assessed (i) participant's own *relaxation state* (question: "how relaxed do you feel now?") , (ii) the quality of the sonic design for representing the relaxation state, called *sound tranquility* (question: "how relaxing was the sound you just heard?"), and (iii) the *perceived congruency* between own relaxation state and the relaxation represented by the sonic display (question: "how well did the sound reflect your state of relaxation?").

Data analysis

All data satisfied the normality criterion as verified using the Kolmogorov-Smirnov test. To test our hypotheses, we applied 3-way ANOVAs with the factors of (i) *sonic design* (direct, parameter mapping or musical mapping), (ii) *type of feedback* (real or sham) and (iii) *presentation block* (1st and 2nd) for 3 subjective measures and relative alpha power values. Greenhouse-Geisser correction was used to correct for unequal variances. Alpha level was fixed at 0.05 for all statistical tests.

4.2.2. Results and discussion

From the three experimental factors only *sonic design* showed significant differences both for *sound tranquility* ($F_{(1.9,20.5)} = 12.42, p < 0.001, \eta_P^2 = 0.53$) and for the *reported relaxation* ($F_{(1.5,16.3)} = 8.53; p < 0.005, \eta_P^2 = 0.44$). We used Bonferroni-corrected pairwise comparisons to see these effects in detail (see Figure 4.3). For *sound tranquility*, the direct sonification was perceived significantly less relaxing than the two other sonic designs. Moreover, parameter mapping sonification was perceived as the most relaxing. A similar pattern can be observed for *reported relaxation*, with parameter mapping sonification leading to the highest reported measures. No significant differences were observed for *perceived congruency* ($F_{(1.7,19)} = 3.27, p = 0.07, \hat{\eta}_P^2 = 0.23$). However, as shown in Figure 4.3 (panel C), a similar trend for parameter mapping sonification was found, being perceived as the most congruent between participants relaxation and sound representation. Relative alpha power data did not show significant differences across all three factors. Finally, the factors of *type of feedback* and *presentation block* did not reach any significance.

Going back to our initial hypotheses, **H1** ("simple and direct sonic designs will *represent* relaxation states estimated through EEG better") is discarded, as direct sonification performed significantly worst than parameter mapping sonification and musical mapping in conveying relaxation (*sound tranquility*). The same trend was detected when participants were asked about the *perceived congruency* between the sound and their own relaxation state after the experiment, although the lack of significant effects in this aspect demands a deeper study, probably including longer exposure to sonic designs. This preliminary results nonetheless show promising *perceptualization* quality for parametric mapping sonification and musical mappings for representing participants' relaxation state, leading us to consider parameter mapping sonification as a good candidate for further develop sonic interactions for implicit PhyComp.

Regarding **H2** ("sonic designs with more complex mappings, such as parameter and musical mappings, will reach greater *relaxation effect*"), the study showed that parameter mapping sonification also achieved significantly better relaxation effect than the direct sonic strategy (see Figure 4.3, panel B), but it is still comparable with the performance of musical mapping. It is important to note that this trend was only observed in the subjective data. Also related with this issue, **H3** ("relaxation will be *induced* in participants as result of the neurofeedback training") has to be rejected. Although participants declared a significantly higher relaxation effect when exposed to parameter mapping sonification, there is no evidence of induced neurofeedback effect by any of the sound designs when

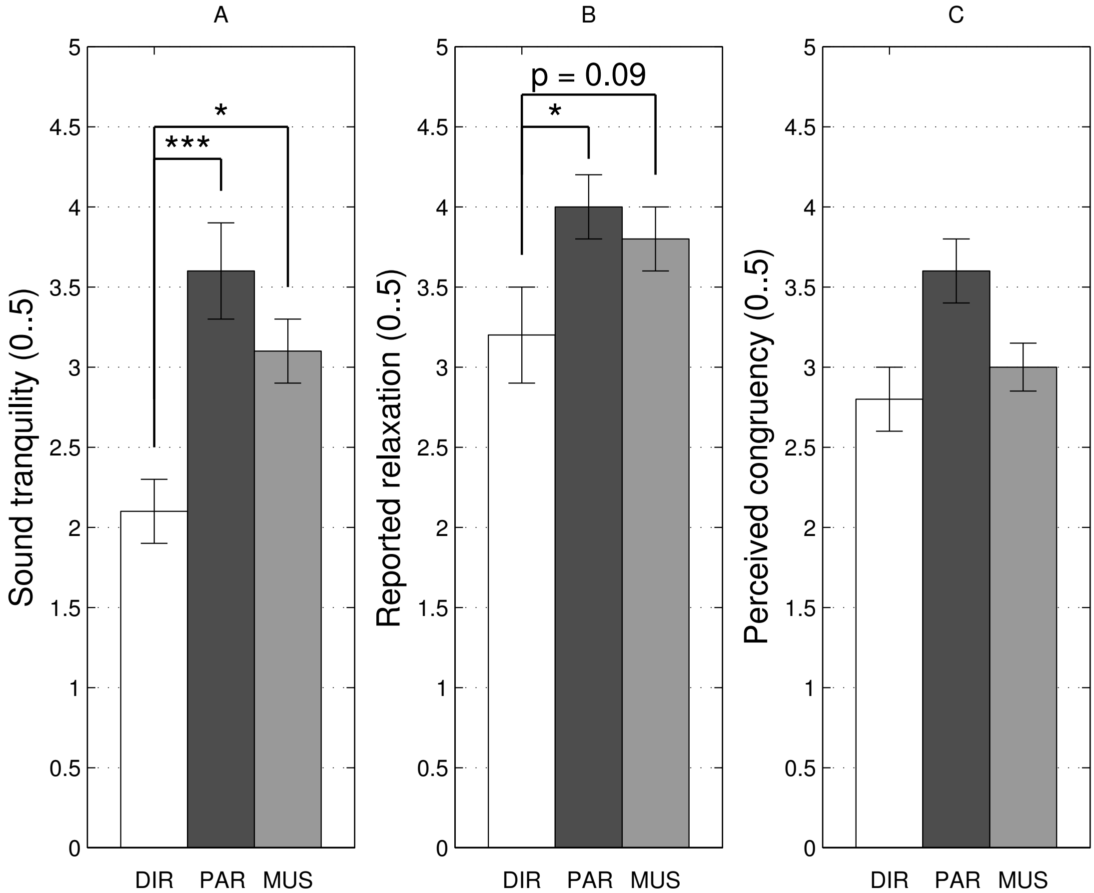


Figure 4.3.: Ratings of *sound tranquility* (A), *reported relaxation* (B), and *perceived congruency* (C) for the three sonic designs: audification (AUD), parametric mapping (PAR) and musical mapping(MUS). Significant differences from Bonferroni-corrected pairwise comparisons (A) and t-tests to 0 (B) are marked at $p < 0:05$ (*), and at $p < 0:005$ (***) levels. Error bars represent standard error values.

looking at the EEG recordings.

At this point it is important to make a distinction between *perceived* and *induced* relaxation to fully understand the contributions of this experiment. These results offer empirical evidence on the advantages of parameter mapping sonification for the *perceptualization* of a given implicit psychophysiological state (relaxation). Although participants have declared to *felt* more relaxed after being exposed to parameter mapping sonification (see Figure 4.3, panel B) this result by itself does not provide enough evidence for claiming that relaxation has been *induced* to participants as result of the neurofeedback training, mainly because no differences emerged when compared real and sham feedback. In fact, Zenter and colleagues [2008] carried out experiments to precisely examine the differences in *felt* and *perceived* emotions when conveyed through sound (particularly music). Their conclusion highlights that emotions were less frequently *felt* in response to music than they were *perceived* as expressive properties of the music (p. 502). This distinction has been also well documented by Västfjäll [2002], Gabrielsson and Lindström [2001], Vuoskoski and Eerola [2011]. By examining the results of this first study, on the one hand, we cannot claim that any of the tested sonic designs perform better than the rest for inducing relaxation through neurofeedback training, but on the other hand, we have found evidence about parameter mapping sonification performing better in terms of *perceptualization* (the representation of relaxation through sound, aiding user perception) with potential for inducing relaxation, as shown by the subjective data.

We leveraged on these findings for designing a second experiment that will compensate the weaknesses of experiment 1. In order to further explore relaxation induction and neurofeedback effect, experiment 2 will introduce longer training sessions (i.e. longer exposure to the stimuli), a more complex EEG estimation of relaxation states (i.e. alpha/theta ratio) to determine whether a more sophisticated physiological estimation plays a role in perception and induction of implicit PhyComp. We will also involve end-user perception in the sonic design loop, allowing participants to personalize physiology-to-sound mappings before the experiment. Finally, it should be noted that the similarity observed between parametric and musical sonic designs is somehow expected, as both were more complex in nature than the direct sonification strategy. We consider the aesthetic dimension of these sonic designs as a main contributor to the difference between parametric and musical mapping.

4.3. Experiment 2: mapping complexity and personalization

The results of experiment 1 showed the advantages of parameter and musical mappings for perceiving implicit physiological states through sound. However, these positive effects have been only demonstrated through subjective perception, without participants' EEG recordings showing any statistical difference, or noticing a significant neurofeedback effect. This second experiment leverages on these finding, focusing on the sound designs that showed better perceptualization performance (parameter and musical mapping) applying a more complex relaxation estimation (alpha/theta ratio), and strengthening neurofeedback training with longer sessions.

Moreover, in this experiment we also further explore a second aspect of SID applied to implicit PhyComp, namely *mapping complexity* (the number of physiological streams and sound parameters used in a given SID strategy). As pointed out in the review presented in Section 2.6, most of the available EEG sonic designs work with a constant number of EEG features, and also apply fixed, predefined EEG-to-sound mappings that cannot be changed by end users. De Campo and colleagues [2007] conducted one of the few studies addressing the personalization of EEG sonification by end users. Tests were done with medical specialists performing evaluation of EEG data containing epileptic events and seizures, and showed the advantages of sound parameters personalization.

The goal of this second experiment is thus to determine whether *mapping complexity* and end-user personalization play a role in the *perceptualization* of implicit physiological states through sound. In order to do so, we improve our sonic engine to account for a flexible and customizable multi-parametric EEG signal transformation into sound. We test this approach in an empirical study based on the well-established alpha/theta (*a/t*) neurofeedback training paradigm (see [Gruzelier, 2009] for a review). During this training, users try to relax with their eyes closed. The most important moment is the so-called *theta/alpha crossover* (*t/a*), when alpha activity slowly subsides accompanying sleep onset and theta activity becomes more dominant [Egner et al., 2002]. The increase of *t/a* power ratio with eyes closed is a well known accompaniment of states of deep relaxation such as stage 1 of sleep and meditation [Gruzelier, 2009]. The change on the relaxation estimation and training based on *a/t* neurofeedback (instead of solely use alpha relative power) responds to the need of further study the lack of relaxation induction found in experiment 1. On the other hand, multidimensional physiological features are a prerequisite for creating sonic designs with greater *mapping complexity*. Finally, as the *a/t* neurofeedback training is typically based on auditory feedback due to its closed-eyes condition, it is a good candidate to validate the personalization and

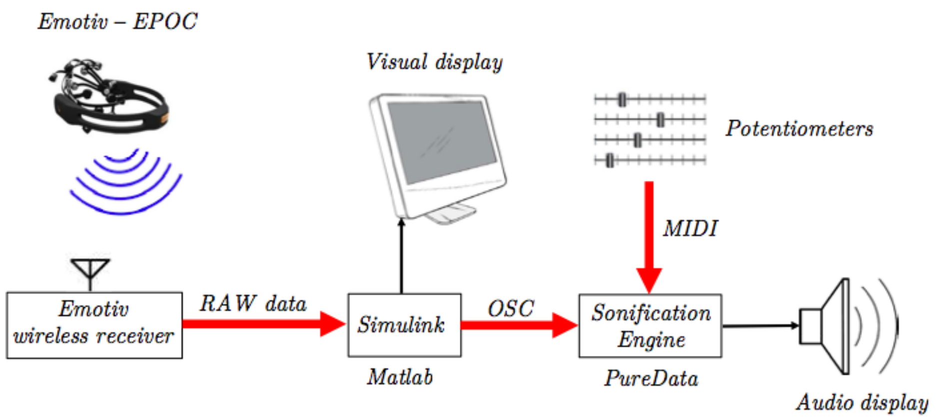


Figure 4.4.: Components and data streams of the EEG sonic display system: EEG signal acquisition, signal extraction and processing, and sonic engine.

complexity of our sonic designs.

The driving hypotheses for experiment 2 are:

- **H1**: *a/t* neurofeedback training will have more impact for participants using *personalized* sonic designs than for the ones using pre-defined, fixed sound mappings.
- **H2**: *a/t* neurofeedback training will have more impact for participants using more *complex* sonic designs (based on multiple EEG features) compared to the ones using only one (*t/a ratio*).

4.3.1. Material & methods

System design

The system developed for experiment 2 is composed of three main blocks: (i) EEG signal acquisition, (ii) signal extraction and processing, and (iii) the sonic engine. This modular approach allow us to create personalized and versatile EEG-to-sound mappings. Figure 4.4 illustrates each component of the system architecture, and in the following sections we describe them in depth.

EEG signal acquisition

As in experiment 1, EEG activity is acquired using the Emotiv EPOC. The data is sampled at $128Hz$ and low- and high-pass filtered internally at $85Hz$ and $0.16Hz$ respectively. This semi-raw data is accessed through the EPOC SDK and then sent to

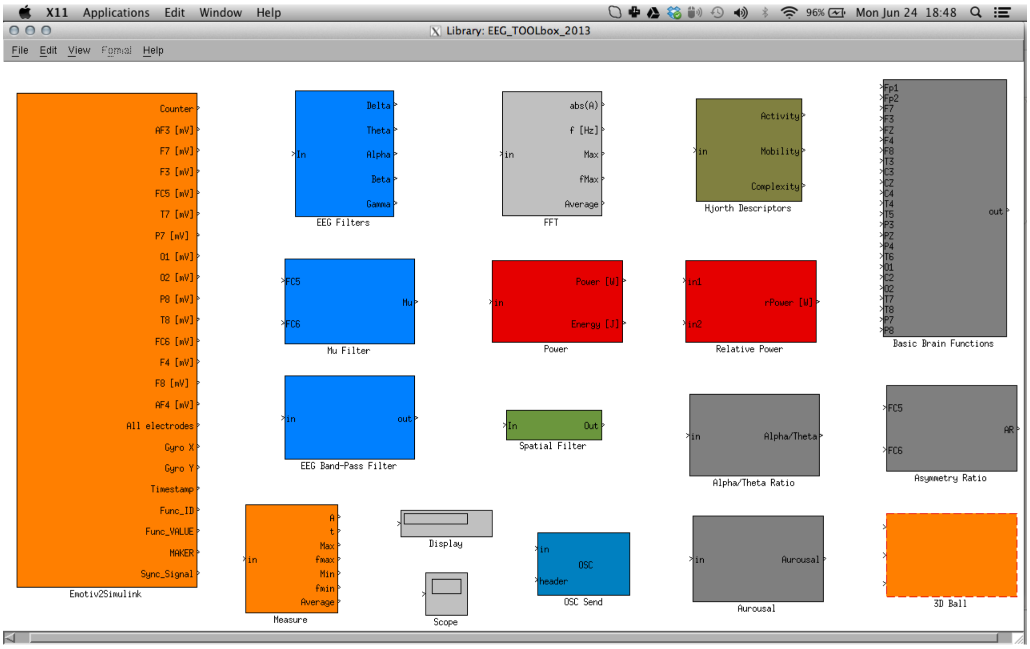


Figure 4.5.: The *EEG-TOOLbox* integrated into the Matlab/Simulink library.

Matlab/Simulink[6], as described below.

Signal extractions and processing

Differently from experiment 1 (where an ad hoc signal processing was applied to specifically estimate EEG alpha relative power), for this study we created a custom-made toolbox for Matlab/Simulink (denominated *EEG-TOOLbox*) to process the EPOC EEG data, both online and offline, and to feed the sonic engine. Although other tools have been designed for real time processing of EEG in Simulink (see Arslan et al. [2005, 2006], Filatriau et al. [2006]), they have been conceived as processing systems rather than a modular toolbox. This was a main motivation for creating the *EEG-TOOLbox*, composed of different processing blocks (see Figure 4.5) which allow extracting EEG features in a modular way, adapting to different neurofeedback or monitoring scenarios, beyond the scope of this particular study. Each block present a menu for adjusting processing parameters and for configuring its internal properties, as described below.

Emotiv2Simulink: through this block it is possible to access to the EPOC raw EEG data stream for further processing within Simulink. The block is based on a Mex S-function and on the drivers provided by the manufacturer. The Mex S-function outputs a vector with the data, and a demultiplexer is used to separate 21 data types (EEG channels, the

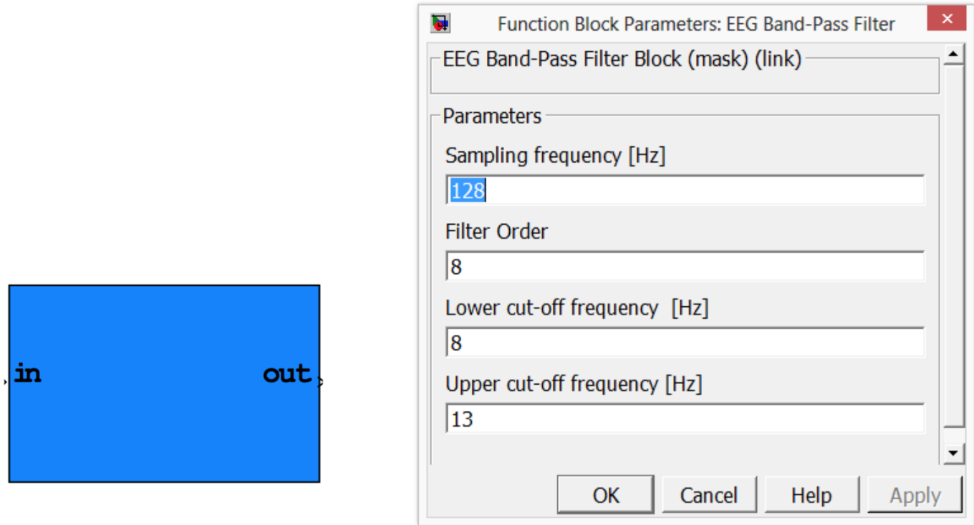
[6] http://www.mathworks.com/products/simulink/ (accessed on October, 2015).

Figure 4.6.: The EEG-TOOLbox. Emotiv2Simulink block.

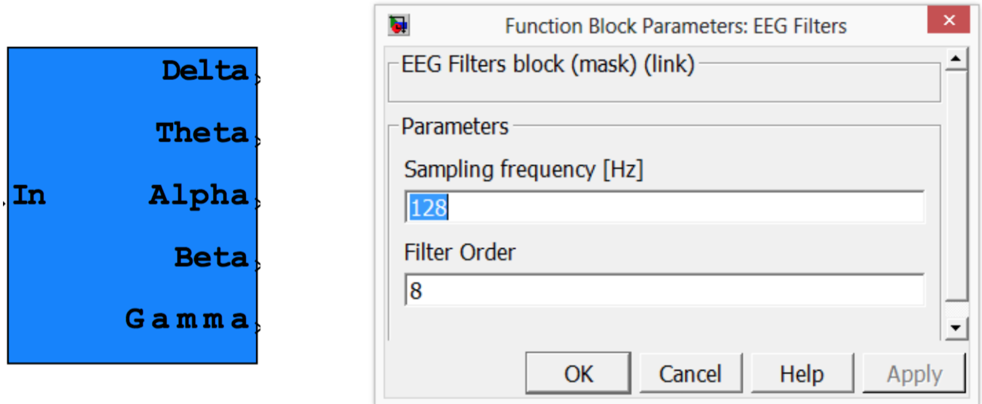
EPOC gyroscope, and synchronization signals) as depicted in Figure 4.6.

BP-filters: this block is used for applying a band-pass filter to the input signal. It is based on a Butterworth IIR discrete band-pass (BP) filter that allows direct configuration of sampling frequency, filter order, lower and upper cut-off frequencies (see Figure 4.7a). The *EEG-TOOLbox* also offers a filtering block with predefined EEG frequency ranges: delta $(0.5-3.5Hz)$, Theta $(4-7.5Hz)$, alpha $(8-13Hz)$, beta $(14-30Hz)$, and gamma $(30-63Hz)$. This block only requires to select a given sampling frequency and filter order (see Figure 4.7b).

Envelope: a block that squares the input signal and then applies a FIR low-pass filter and down-sampling to estimate its envelope (See Figure 4.8). The signal is squared to demodulate the input signal, using the input as its own carrier wave. This means that half of the energy of the signal is shifted up to higher frequencies, and the average moves down to DC. Next step is to reduce the signal resolution in order to reduce the sampling frequency. As the signal may contain frequencies which can cause aliasing, a FIR low-pass filter is applied before performing the downsampling. A minimum phase low-pass



(a)



(b)

Figure 4.7.: The EEG-TOOLbox. Band-pass filter blocks.

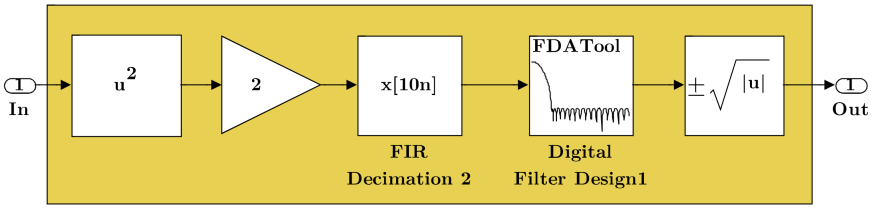


Figure 4.8.: The *EEG-TOOLbox*. Processing for envelope estimation.

filter is applied to remove the energy of high frequencies. The result of this process is the envelope of the signal. In order to maintain the correct scale, the signal is amplified by a factor of two; since we are keeping only the lower half of the energy of the signal, this gain allows to make it match with the original energy. Finally, the square root of the signal is calculated to reverse the distortion resulting from squaring the signal.

MaxMin: this block is based on a configurable window of size L, and it estimates the maximum, minimum and median value of a signal in the temporal and frequency domain.

(Relative)Power: this block calculates the power of the input signal applying the Parseval formula

$$P_x(k) = \frac{1}{N} \sum_{n=0}^{N-1} \mid x(n) \mid^2 \tag{4.1}$$

based on a sliding window that performs the cumulative sum and squaring for each instant of time. The Relative Power block works the same way, calculating the power of *in1* and dividing it by *in2*, where *in1* is a signal filtered on a specific frequency band, and *in2* is the full frequency spectrum of the same signal. Being the result a relative value, the output can only go out from 0 to 1.

Hjorth-Descriptors: this block calculates the Hjorth parameters for activity, mobility, and complexity [Liang et al., 2013, Hjorth, 1975]. Hjorth activity is estimated as

$$A_{(x)} = std(x)^2 = Var(x) \tag{4.2}$$

the Hjorth mobility parameter is calculated as

$$M(x) = \left(\frac{\dot{A}(x)}{A(x)} \right)^2 \tag{4.3}$$

the Hjorth complexity parameter is calculated as

$$C(x) = \left(\frac{\dot{M}(x)}{M(x)} \right)^2 \tag{4.4}$$

OSCSend: it allows direct communication via UDP with virtually any modern real-time sound synthesis environment through the Open Sound Control (OSC) protocol.

The sonic engine

We updated the sonic engine presented in experiment 1, following a modular approach for allowing: (i) to choose among different EEG features (as processed by the above mentioned Simulink toolbox), (ii) to decide on the number of EEG features to be display sonically (*complexity*), and (iii) to flexible define EEG-to-sound mappings (*sound modules*). Although for this experiment the system is tested with 3 EEG features feeding 5 sound modules in parallel, there are no restrictions on the number of features and modules to be displayed, other than hardware limitations (e.g. processing power). Multiple EEG streams can also feed a single sound module. Importantly, the engine also allows further personalization by end users via sliders.

Below we provide a description of the different sound modules that compose the engine at this stage, following the same mapping criteria for relaxation neurofeedback training used in experiment 1 (Section 4.2.1). For flexibility and easiness, all inputs of the sonic engine modules are normalized between 0 and 100.

- **Pedal module:** a fixed pitch tonic sound is initialized on loading (D2, MIDI note n° 38). A tremolo effect is applied and phase shifted to each of the even harmonics, giving a slowly moving chorus-like timbre to the drone. Inputs for this module are able to control tremolo, panning and gain (see Figure 4.9).

- **Melody module:** the melody is constructed using a major scale stemming from five semitones (one fourth) bellow the central tonic to sixteen semitones above it (major third). The slope of the incoming signal controls note triggering speed and pitch, whereas the volume is controlled by a linear function of the input signal (see Figure 4.10).

- **Rewarding module:** this module triggers short duration nature sounds (birds, owls, and crickets) when the input signal goes over a predefined threshold value. Therefore, every time the input signal is over the threshold set, the user is rewarded with a more varying yet still relaxing soundscape. A dynamic amplitude panning allows use of spatial audio (see Figure 4.11).

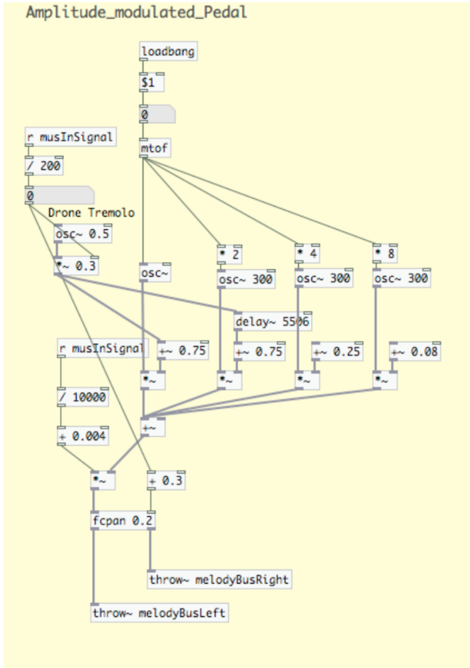


Figure 4.9.: Sonic Engine. Pd patch of the Pedal module.

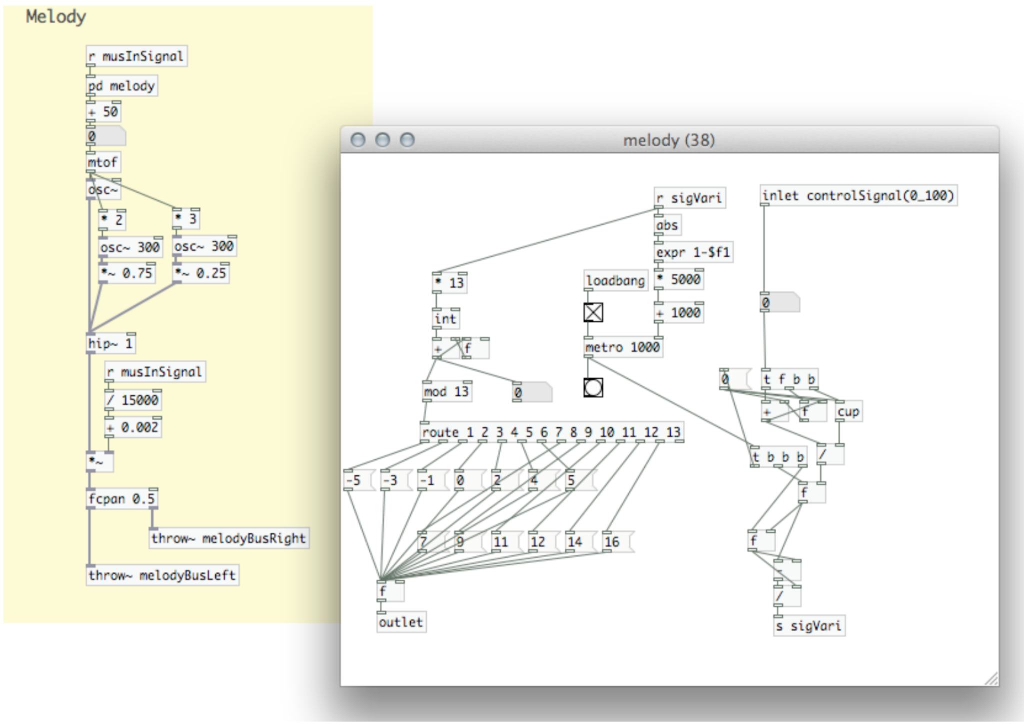


Figure 4.10.: Sonic Engine. Pd patch of the Melody module.

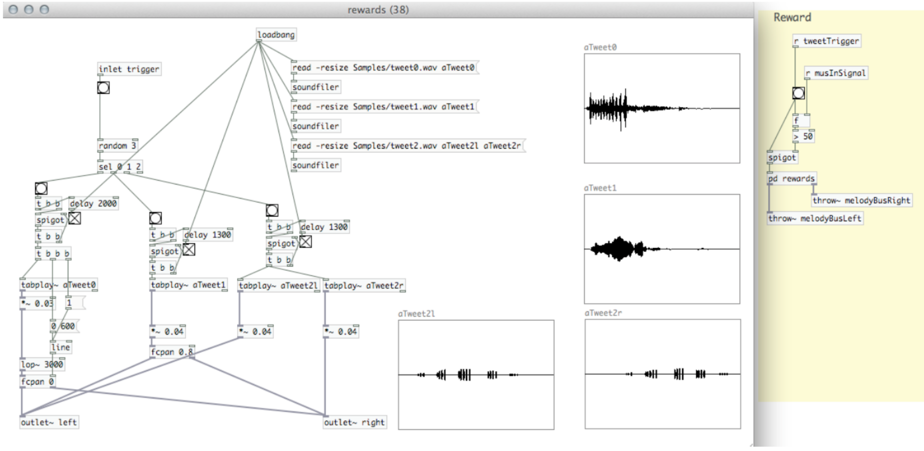


Figure 4.11.: Sonic Engine. Pd patch of the Rewarding module.

- **Wind model module:** based on a procedural audio, this module uses a series of white noise generators to create a wind soundscape. The use of procedural audio (in contrast to a sample-based approach) gives a complete control over sound parameters (e.g. wind speed). Therefore, the perceptual quality of the windy scene dynamically changes according to mapped EEG features (see Figure 4.12).

- **Rain model module:** based on a sample-based approach, this module preserves soundscape fidelity using dynamic cross fading between different rain excerpts. It also allows a comparison with the procedural audio approach of the wind module. Input signal modulates the amount of excerpts in the resulting composite *rain density*, which varies from a light rain to a small thunderstorm (see Figure 4.13).

- **Graphic User Interface and personalization:** a graphic user interface (GUI) is used to simplify the patching between different sound modules and EEG features. It also contains time and frequency signal monitoring scopes for both input (EEG) and output (sound) data. The GUI also displays tools for users' personalization via sliders using a MIDI controller, allowing them to adjust their mapping *on the fly* while listening to the sonic results in real-time (see Figure 4.14).

Users are able to personalize (i.e. adjust) the mappings as follows. One EEG feature is associated to one sound element (e.g. alpha relative power is mapped to rain density of the rain model module). The user is then allowed to change the mapping in a continuous scale from -1 to +1. If the potentiometer is dragged to -1, the mapping is inverted, i.e. a greater value of the input data will yield a smaller value for the sound parameter (more alpha would lead to less rain). If it is positioned at 1, the mapping is positive (e.g. more

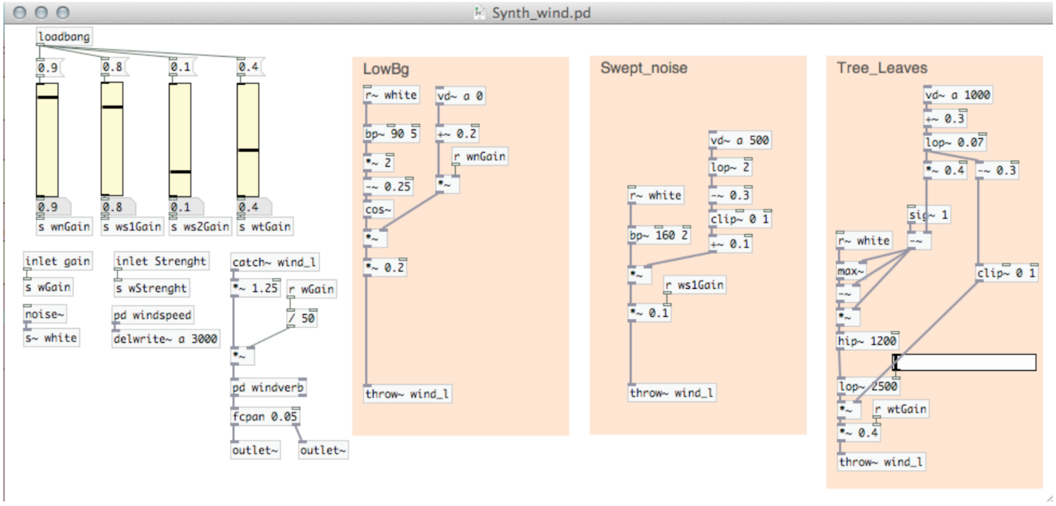


Figure 4.12.: Sonic Engine. Pd patch of the Wind module.

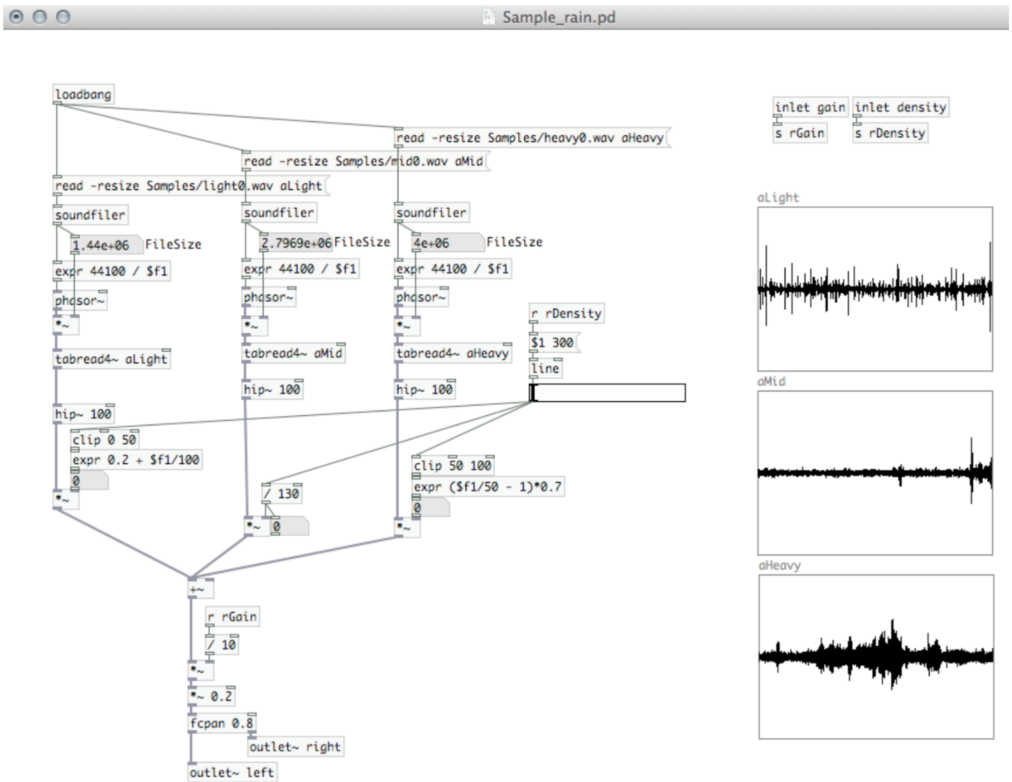


Figure 4.13.: Sonic Engine. Pd patch of the Rain module

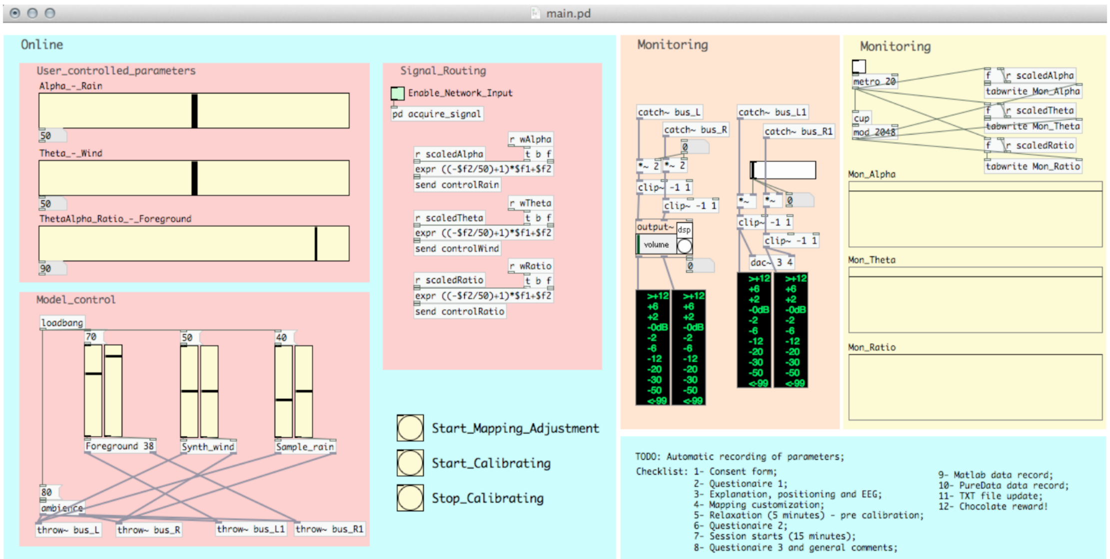


Figure 4.14.: Sonic Engine. Graphic User Interface

alpha yields more rain). If the potentiometer is place in the middle (zero), the sound is fixed, thus not influenced by the input data (e.g. a fixed amount of rain equivalent to 50% of its maximum value, independently of the alpha power).

As can be seen, this new version of the sonic engine[7] presents considerable technical and design advantages compared to the version used in experiment 1. In this new engine, modules can be combined and expanded easily using Pd, parameter mapping sonifications has been enriched with more complex models using procedural audio and samples, and end-user personalization is easily allowed through a GUI and physical sliders.

System settings used in the experiment

Figure 4.15 shows the system configuration for the experiment. The *EEG-TOOLbox* is configured to extract three main EEG features, according to the a/t neurofeedback protocol [Gruzelier, 2009]: alpha relative power, theta relative power, and t/a power ratio. In accordance with previous neurofeedback studies [Egner et al., 2002], alpha relative power ($7-13Hz$ range) is calculated from the occipital area where this type of activity occurs during close-eye conditions [Kropotov, 2010]. The second feature, theta relative power ($4-7.5Hz$ range) is calculated from the activity of all 14 channels, as

[7] The sound engine for Pd is available for downloading at `http://www.dtic.upf.edu/~smealla/phd_material.html`

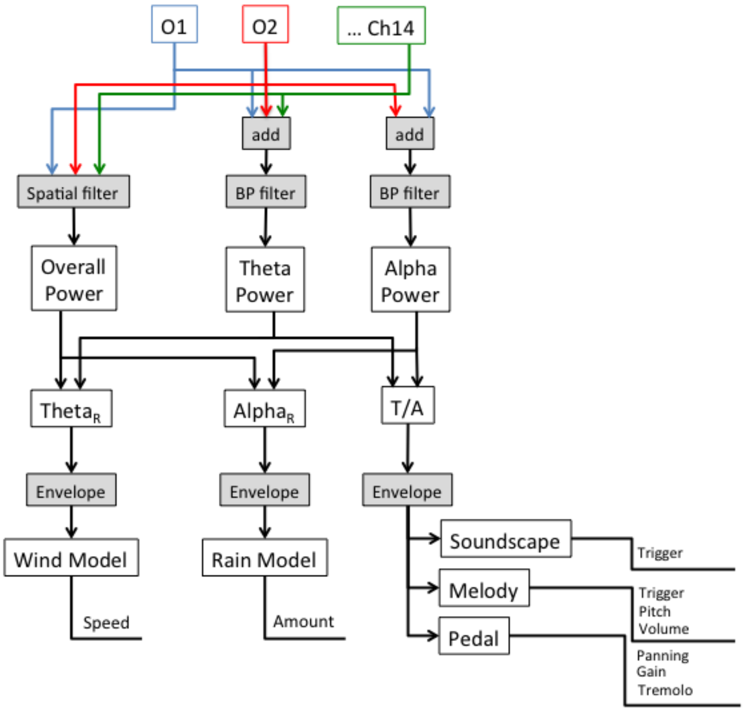


Figure 4.15.: Processing stages of EEG features and corresponding sound mappings for experiment 2.

cortical theta rhythms are small and diffuse when picked up by scalp electrodes, arising almost entirely from the cerebral cortex Kropotov [2010]. These two relative powers are obtained by dividing the band power by the overall signal power. In this manner, the output signal is kept within a range of 0 to 1. Finally, the third feature, *t/a* power ratio, is estimated as the main measure for the *a/t* neurofeedback procedure. A spatial filtering is applied to give more weight to the occipital channels in the calculation of the envelope of the alpha band and *t/a* ratio. For the calculation of the envelope of theta, equal weight for all channels is used. The envelope is estimated for all three EEG features using the envelope block based on FIR-based filter (order 35).

The sound engine has been configured as follows[8], according to the mapping criteria defined in 4.2.1 (see also Table 4.2 on the next page):

- *Pedal module*: tremolo, panning and gain are linearly mapped to *t/a* ratio.
- *Melody module*: note triggering speed and pitch are functions of the *t/a* ratio where its positive change yield descending pitches while negative ones leads to ascending

[8]A video of the sound engine configured for the experiment is available at `http://www.dtic.upf.edu/~smealla/PhD_Material/videos.html`

Group	Foreground	Reward	Wind	Rain
Fix-Single (F-S)	t/a ratio (Fixed)	t/a ratio	Constant	Constant
Personalized-Single (P-S)	t/a ratio (Personalized)	t/a ratio	Constant	Constant
Fix-Multiple (F-M)	t/a ratio (Fixed)	t/a ratio	theta (Fixed)	alpha (Fixed)
Personalized-Multiple (P-M)	t/a ratio (Personalized)	t/a ratio	theta (Personalized)	alpha (Personalized)

Table 4.2.: Configuration of the sonic engine for each experimental group, defined by type of mapping (Personalized or Fixed) and number of EEG features (Single or Multiple). EEG features include t/a ratio, and relative power of alpha and theta. "Personalized" stands for adjusted mappings driven by EEG, "Fixed" stands for fixed mapping driven by EEG, and "Constant" stands for sound feature held constant.

pitches. Higher rate of change yields bigger jumps in melody (more semitones between two consecutive notes). The note volume is a linear function of t/a ratio.

- *Rewarding module*: threshold is set at 50% of the calibrated maximum value of *t/a* ratio recorded prior to the main training session.
- *Wind model module*: theta relative power controls wind speed of the modeled sound object - the higher is the input, the faster the wind will *blow*.
- *Rain model*: alpha relative power is used to control rain density.

Figure 4.16 shows the experiment setup. The study was conducted in a room isolated from external noise. Participants were seated on a swivel chair equally distant to four loudspeakers (Roland active loudspeakers, Model MA15-D, and a M-Audio sound card, FastTrack Pro), as previous research has shown that spatial rendering increases affective impact of sound [Västfjäll, 2003]. The sonic engine and EEG processing were handled by two different computers. The lights were dimmed down during the personalization and pre-relaxation sessions and turned off during the *a/t* neurofeedback session. A 42 inches screen (not visible to the for participants) was used to monitor the EEG signal quality and levels.

Participants and experimental procedure

Our two hypotheses were tested using a between-subjects design. Thirty one participants, mean age 27.81 ($SD = 5.18$), 15 females, took part in the experiment. The study was conducted in accordance with the Declaration of Helsinki. Participants were equally and randomly distributed among four experimental groups (see Table 4.2):

- **F-S group**: *fixed* EEG-to-sound mappings, *single* EEG feature displayed ($M = 26, SD = 2$).

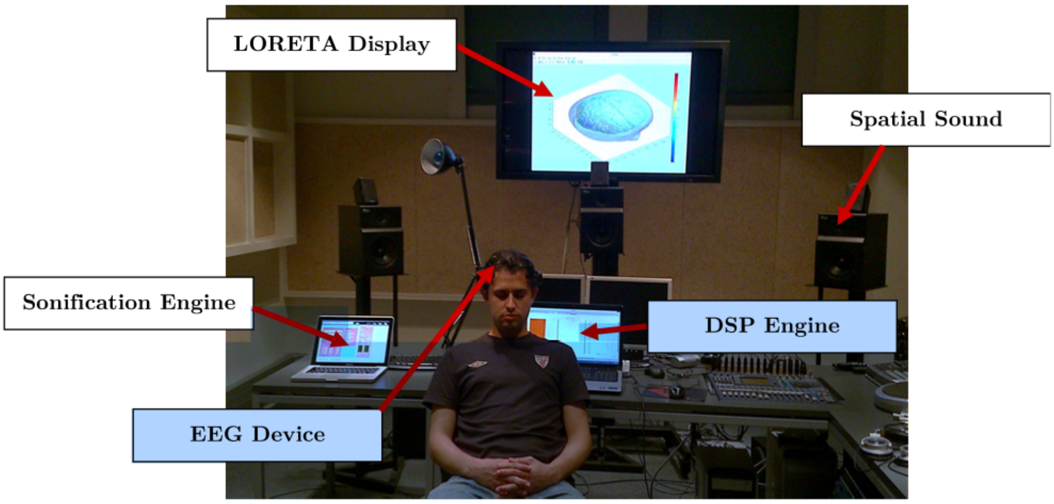


Figure 4.16.: Photo of the technical setup for experiment 2. The participant seated on a swivel chair surrounded by a 4-channel sound system (spatial sound). The DSP and sound design engines were executed from two different laptop computers. A 42 inches screen was used to monitor the EEG signal quality and levels.

- **P-S group**: *personalized* mappings, *single* EEG feature displayed ($M = 27.6, SD = 5$).
- **F-M group**: *fixed* mappings, *multiple* EEG features displayed ($M = 30, SD = 8$).
- **P-M group**: *personalized* mappings, *multiple* EEG features displayed ($M = 27.5, SD = 4$).

All groups listened to soundscapes of a comparable sound richness, but varying in the number of EEG features feeding the system, and in the sonic design applied. Table 4.2 shows the configuration of the sonic engine for each group. Each experiment lasted around 50 minutes, according to the following protocol:

- *Information and consent form*: the participant receives an explanation on each stage of the experiment, and the relation between brain activity and the created soundscape. The participant signs a consent form.
- *Initial emotional state self-assessment*: subjective measures of emotional valence and arousal are collected in paper through a 9-point Self-Assessment Manikin (SAM) scale [Bradley and Lang, 1994].
- *Sensor placement and baseline state recording*: the participant sits in a swivel chair,

and the Emotiv EPOC is mounted in her/his scalp. The baseline EEG activity is recorded.

- *Relaxation induction*: we request the participant to close her/his eyes and listen to a 5-minute sound of sea waves. EEG activity is recorded, and the thresholds for the neurofeedback session is calculated by looking at the maxima and minima values.
- *Pre-test emotional state self-assessment*: the participant is requested to fill in a SAM scale.
- *Mapping adjustment*: EEG features start to be sonically displayed in real time. The participant is then asked to personalize the mappings to reach the "most relaxing sound possible", using a MIDI interface with sliders placed in front of the chair (as described in Section 4.3.1). There is no time limit to personalize the mappings.
- *Neurofeedback training session*: the participant is asked to close her/his eyes and to turn the chair facing away from the experimenter. The participant is then asked to relax and to listen to the soundscape (different for each group) for 15 minutes. The participant is instructed to raise her/his hand if feeling uncomfortable or falling asleep.
- *Post-test emotional state self-assessment:* the participant is requested to fill-in a SAM scales again.
- *Headset removal and debriefing*: the EEG headset is removed. The participant is debriefed, thanked and receives candies as a small reward.

Data Analysis

For analysis purposes, the raw EEG data with sampling at $128Hz$ was first filtered $(0.5 - 30Hz)$. Using visual inspection and thresholding (over 3σ) data regions with artifacts were marked for removal in subsequent analyses. Closely following the design of the neurofeedback training protocol (see Section 4.3.1), we analyzed data from O1 and O2 electrodes. Signals were BP filtered to obtain alpha $(7-13Hz)$ and theta $(4-7.5Hz)$ components. The 15-minute data recordings were split into 10-second epochs and for each of them relative alpha, relative theta and t/a ratio where calculated. Next, we averaged the means from individual epochs for five 3-minute periods (18 epochs each) excluding the epochs marked as containing artifacts. This was done for O1 and O2 channels separately. For analysis we used the values averaged across both channels and the means from the first and the last 3-minute period of the experimental procedure.

In other words, we compared maximal changes caused by 15-minute training in relative alpha, relative theta and t/a ratio.

Both subjective (SAM ratings of pre and post assessment of emotional state) and objective measures (EEG features) of two relaxation periods were subjected to a 3-way MANOVA. Therefore, the within-subjects factor was *relaxation period* (1-3 min vs. 12-15 min period), whereas the between-subjects factors were *Number of sonified EEG features* (Single vs. Multiple) and *Feedback personalization* (Personalized vs. Fixed mapping). Alpha level was fixed at 0.05 for all statistical tests. Greenhouse-Geisser correction was used to correct for unequal variances. For multivariate analysis Wilks' Λ was used as the multivariate criterion. All variables were normally distributed according to the Kolmogorov-Smirnov test.

4.3.2. Results and Discussion

In accordance with the relaxing nature of the experimental procedure, the overall MANOVA effect of *relaxation period* was significant with $F(5, 23) = 6.89, p < 0.001, \Lambda = 0.4, \hat{\eta}_P^2 = 0.6$. Split by measures, this effect reached significance for subjective arousal ratings, $F(1, 27) = 26.06, p < 0.001, \hat{\eta}_P^2 = 0.49$, relative alpha power, $F(1, 27) = 5.81, p < 0.05, \hat{\eta}_P^2 = 0.18$, relative theta power, $F(1, 27) = 10.4, p < 0.005, \hat{\eta}_P^2 = 0.28$, and *t/a* ratio, $F(1, 27) = 5.45, p < 0.05, \hat{\eta}_P^2 = 0.17$. This shows that participants in all four groups reached greater relaxation as compared with the initial 3-minute period of the experiment. More importantly, the interaction between the *relaxation period* and between-group factors of *number of sonified EEG features* and *feedback personalization* also showed significance, as described below.

Confirming **H1** ("*a/t* neurofeedback training will have more impact for participants using personalized sonic designs than for the ones using pre-defined, fixed sound mappings") the overall MANOVA effect of *relaxation period* $\times$ *feedback personalization* was significant with $F(5, 23) = 2.58, p < 0.05, \Lambda = 0.64, \hat{\eta}_P^2 = 0.36$. This effect reached significance both for *t/a* ratio at $F(1, 27) = 13.14, p < 0.001, \hat{\eta}_P^2 = 0.33$; and for relative alpha power, $F(1, 27) = 10.08, p < 0.005, \hat{\eta}_P^2 = 0.27$. Changes in relative theta power did not reach significance. For the groups with personalized feedback, *t/a* ratios increased from 0.73 ($SE = 0.2$) to 1.42 ($SE = 0.2$), while for the groups with fixed mappings such change was not found; the period means were 0.93 ($SE = 0.2$) for the first and 0.79 ($SE = 0.2$) for the fifth one (see Figure 4.17 right panel). A similar pattern occurred for relative alpha levels, with the means for personalized groups dropping from 0.29 ($SE = 0.04$) to 0.15 ($SE = 0.03$), and steady means for groups with fixed mappings of 0.2 ($SE = 0.04$).

This effect can be also seen in Figure 4.18 (right panel) where differences between the last and the first period are plotted.

In line with **H2** ("*a/t* neurofeedback training will have more impact for participants using more complex mappings"), the overall MANOVA effect of *Relaxation period* × *Number of sonified EEG features* was significant with $F(5,23) = 5.09, p < 0.005, \Lambda = 0.48, \hat{\eta}_P^2 = 0.53$. This effect reached significance only for *t/a* ratio at $F(1,27) = 8.94, p < 0.01, \hat{\eta}_P^2 = 0.25$. Here, the higher number of sonified features in the feedback resulted in a greater *t/a* ratio increase from the mean of $0.83\,(SE = 0.2)$ at the initial experiment stage to $1.45\,(SE = 0.2)$ at the final 3-min period. In comparison, groups with single-feature based feedback showed no improvement, going from the mean of $0.84\,(SE = 0.2)$ to $0.76\,(SE = 0.2)$, see also Figure 4.17 (right panel).

A triple interaction between *Relaxation period* × *Feedback personalization* × *Number of sonified features* was also observed. It reached significance only for relative alpha, $F(1,27) = 5.13, p < 0.05, \hat{\eta}_P^2 = 0.16$. This effect can be better seen in Figure 4.18 (right panel), where differences between two periods are plotted for each of participants' group. While not significant when comparing the period differences within each group, this interaction is reflected in bigger difference between F-S and P-S groups as compared to the difference between F-M and P-M groups.

4.4. General discussion

As shown in the previous section, the results from our between-group analysis confirmed our initial hypotheses, both for subjective and objective data. The significant differences between initial and final relaxation periods for P-S and P-M groups using personalized feedback support our first hypothesis that personalized mappings are more instrumental than the fixed ones for displaying implicit physiological states through sound (in this case, relaxation). This is supported by a significant increase in *t/a* ratio and decrease in alpha relative power, observed after 15-min neurofeedback session. This trend did not occur in F-S and F-M groups, where training was done with fixed sonic mappings.

Our second hypothesis, stating that more complex mappings (based on multiple EEG features) will be more efficient than those relying on a single EEG feature, was also confirmed. A significant increase of *t/a* ratio and a significantly lower reported subjective arousal after the neurofeedback session was observed only for P-M and F-M groups with multiple feature feedback. As expected, the training effect was smaller for P-S and F-S groups undergoing neurofeedback with a single feature mapping.

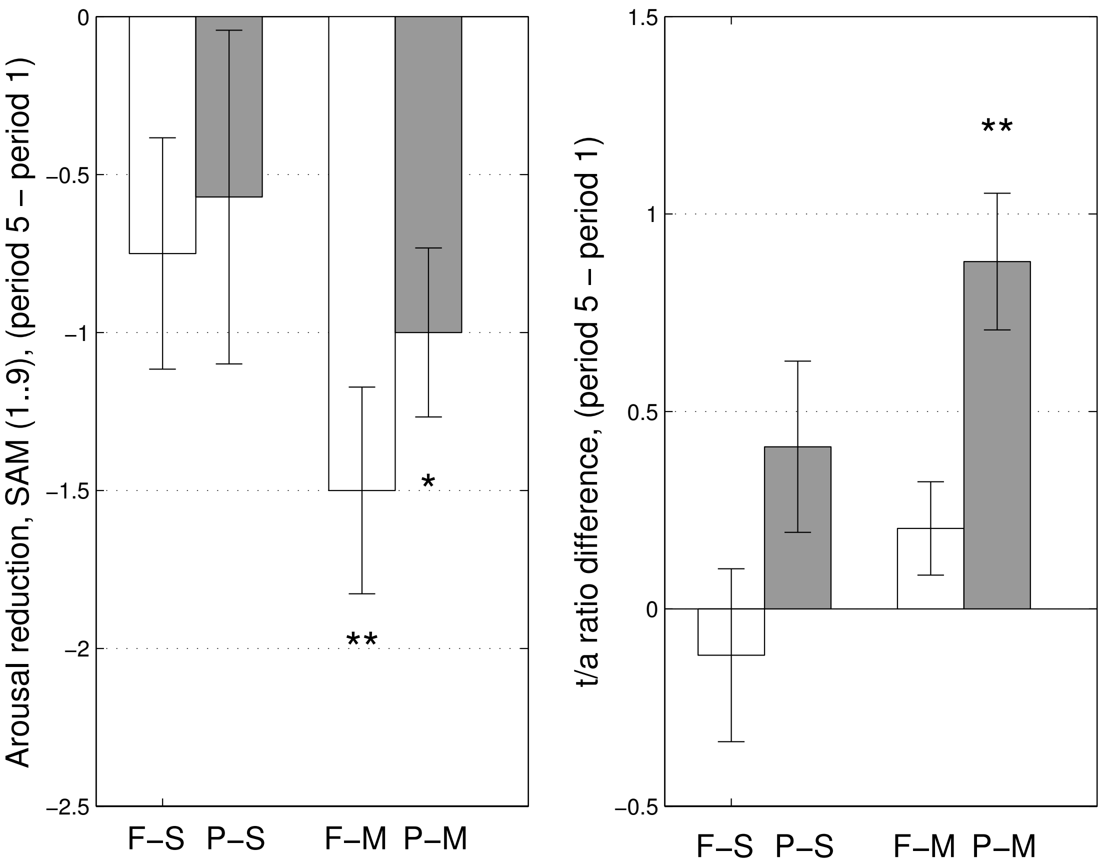


Figure 4.17.: The means of the difference between the first and the final 3-min period of the 15-minute session. Subjective arousal (left panel) and t/a ratios (right panel) are shown for four experimental groups: F-S (Fixed mapping/ Single feature), P-S (Personalized mapping/ Single feature), F-M (Fixed mapping/ Multiple features), P-M (Personalized mapping/ Multiple features). Error bars represent standard error values. Bonferroni-corrected significant difference from 0 at $p < 0.05$ (*), and at $p < 0.01$ (**) levels.

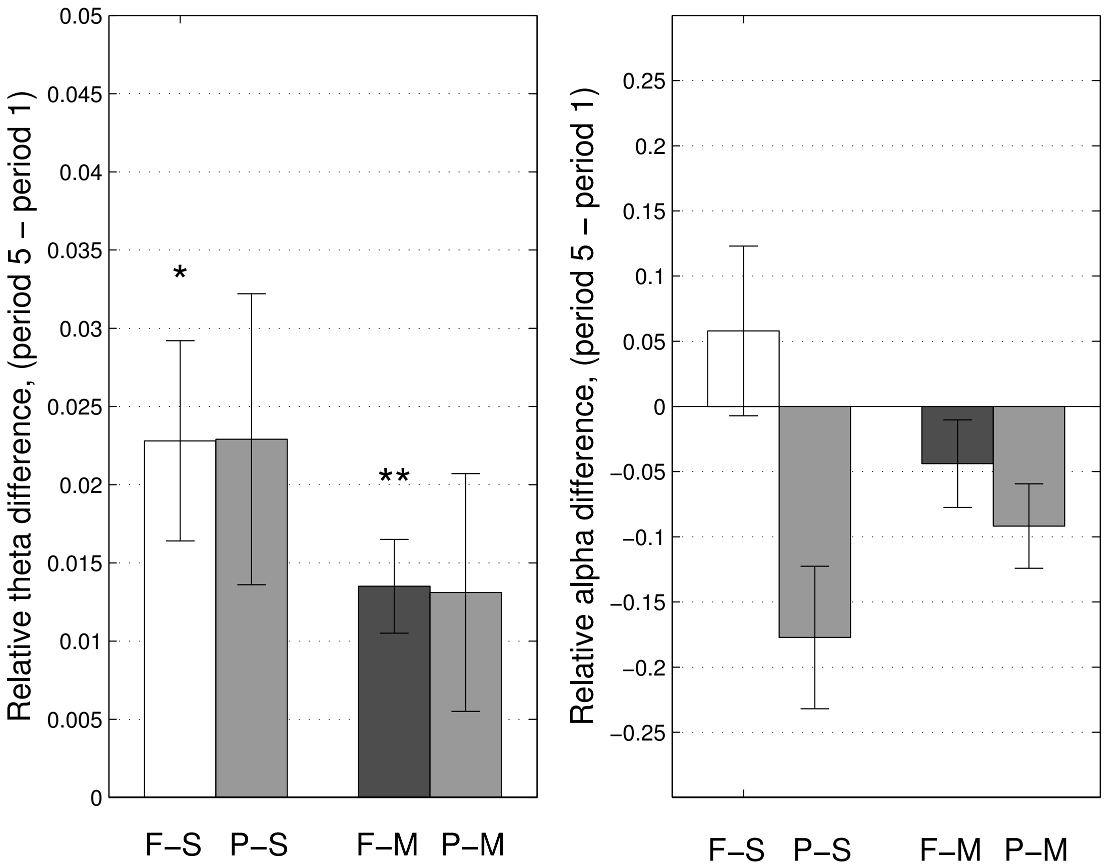


Figure 4.18.: The means of the difference between the first and the final 3-minute period of the 15-minute session. Relative theta power (left panel) and relative alpha power (right panel) are shown for the four experimental groups: F-S (Fixed mapping/ Single feature), P-S (Personalized mapping/ Single feature), F-M (Fixed mapping/ Multiple features), P-M (Personalized mapping/ Multiple features). Error bars represent standard error values. Bonferroni-corrected significant difference from 0 at $p < 0.05$ (*), and at $p < 0.01$ (**) levels.

Finally, changes in relative alpha power showed an interaction between *mapping complexity* and personalization. Here the differences between the P-S and F-S (groups that were using fixed and personalized sound mappings of a single EEG feature) are bigger than between P-M and F-M (groups whose training was based on multiple EEG features). In other words, personalization became less instrumental when multiple features are displayed. In this case, a decrease in user attentional resources could explain this effect. As the number of sound processes to attend to becomes larger, it is more likely that *change deafness* will reduce the efficiency of sonic designs, as stated by Shinn-Cunningham [2008] on auditory spatial attention. Certainly, *mapping complexity* plays a central role in the design of sonic interactions for implicit PhyComp, and has to be further explored.

Significant between-group differences in both subjective and physiological data show that the observed results are not obtained merely due to relaxing nature of presented sounds (wind, water, etc.). Thus, the effect caused by our *a/t* neurofeedback setup differed from simple relaxation. The observed results, that varied significantly between four test groups, support the functionality of the conducted neurofeedback training, and we can argue that both *mapping complexity* (the number of EEG features to be displayed) and end-user personalization play an important role in the effectiveness of sonic designs for physiology-based implicit interaction. More training sessions, larger number of participants, and more robust EEG equipment could be used to further study brain dynamics during *a/t* training (e.g. changes between training sessions, learning curves within each session, individual differences etc.).

Differently from our first study, experiment 2 did not include a sham group with fake neurofeedback, or a control group undergoing 15-minute relaxation to compare with our experimental conditions. As shown by the pre and post-training differences, the participants within the F-S group, with fixed sound mappings of a single EEG feature, can be seen as a strong baseline for the three other groups. Nonetheless, future works could complement this study by adding a control group to quantify the depth of relaxation induction with and without neurofeedback.

Also several observations can be made when comparing our results with previous works on EEG sonification and personalization. In our case, the transparency of the personalization controls was an important factor, as already suggested by previous research like De Campo et al. [2007]. The positive effect of personalization in our studies demonstrates that the task of sound mapping adjustment was clear and understandable for end-users. However, the development and evaluation of intuitive interfaces for personalization should be further explored. For instance, a natural step would be to integrate the sonic engine presented in this Chapter into the *b-Reactable* to further study whether *mapping complexity* and personalization of sonic designs for physiology-based implicit interaction also result in positive effects for music performance (e.g. enhanced expressiveness).

Another key issue tackled by our two experiments is the validation of sonic designs in the context of implicit PhyComp. As pointed out by our literature review (see 2.6 on page 56), while diagnostic applications are expected to have a rigorous assessment, most of sonic designs applied to PhyComp lack a systematic validation, making difficult to determine the efficiency of a particular sonic design for an specific goal (in this case for conveying relaxation estimated through EEG). In this context, we contributed to the methodologies for sonic design validation through a combination of objective and subjective measures using a well-known perception-based scenario (i.e. neurofeedback). In fact, some of the discrepancies between subjective ratings and physiological indices observed in our study should encourage the use of multimodal measures for correcting and further interpreting results. Our studies also provide grounding to further explore other aspects of SID applied to implicit PhyComp, such as expressiveness in HCI contexts (see Chapter 6).

Finally, it is important to note a number of advantages offered by the presented sonic design system. Firstly, the sonic models created for this experiment (wind and rain, with procedural audio and samples respectively) can be easily modified and further used for future research, promoting new EEG-to-sound mappings, and realistic dynamic sounds. Secondly, the use of Pd makes it easy for other practitioners and researchers to modify the existing sound modules and to add new ones. In this manner, the sonic engine could easily be expanded to present sounds that are directly related to other psychophysiological states estimated through different modalities such as breathing or heart rate [Tajadura-Jiménez et al., 2008]. Other future work could test different neurofeedback protocols, integration of other EEG devices, and deeper studies on user preference for sound mappings and their effectiveness for other domains such as data mining, diagnosis, entertainment, or music.

4.5. Conclusion & next steps

This Chapter presented a set of experiments that explored two main aspects of sonic interaction design applied to implicit PhyComp, as depicted by Chapter 3: *perceptualization* (how well a sonic design represents a given implicit physiological state, aiding user perception), and *mapping complexity* (the number of physiological streams and sound parameters used). The goal of these studies was to determine how these aspects affect the performance of sonic designs for implicit PhyComp. Whereas experiment 1 provided empirical insights on the *perceptualization* quality of parameter mapping sonification and musical mappings for representing implicit physiological states (specifically relaxation) and for aiding user perception, the results of experiment 2 provided evidence about the role that both *mapping complexity* and end-user personalization play in the *perceptualization* of sonic designs for implicit PhyComp. The studies also offered insights on the interaction of these issues, showing that personalization becomes less instrumental when multiple physiological features are displayed through sound (probably due to a decrease on the user attentional resources).

In the next Chapters we will then move forward towards *meaningfulness,* one of the three aspects that, together with *perceptualization* and *mapping complexity,* arose from our first study on sonic interaction design (see Chapter 3). In order to do so, we will use the findings of this Chapter to create a new version of the *b-Reactable* that will include a parameter mapping approach, and will allow user personalization directly through the tangible user interface. Through this approach we will aim at studying whether physiology-based, implicit sonic interaction can *meaningfully* support music performance. For undergoing this task, we first propose a methodological framework for NIME design, that focuses on the exploration of *expressiveness* and on the role of the mapping component in the NIME creation chain. This framework will make possible to directly assess system properties (such as mapping and synthesis) and performance aspects (such as musicality and expressiveness) of the updated version of the *b-Reactable.*

5. A Methodological Framework for Designing and Evaluating NIME

One of the main SID aspects to be explored in this dissertation is the *meaningful* integration of implicit PhyComp in a relevant HCI context, namely music performance. As this task requires systematic and specific evaluation methods, in this Chapter we present a framework for designing and evaluating new interfaces for musical expression (NIME), with special interest on *expressiveness,* a relevant aspect to determine the *meaningful* contribution of implicit PhyComp in the NIME field. The objectives of this Chapter are therefore (i) to analyze the most relevant NIME design and evaluation frameworks available in the literature, (ii) to identify how they tackle different stakeholders, and (iii) to propose and test a framework focused on *expressiveness*, *mapping,* and participants' previous musical knowledge. This framework is deployed in a one-trimester NIME master course where groups of participants (students) prototype DMIs within a restrictive setup, consisting of smart-phones controllers and the Pure Data (Pd) programming language, and perform with them in front of the rest of the class, which in turn evaluates the performances as *listeners*, in an iterative process. The insights gathered during the study suggest that students with different backgrounds were able to effectively engage in the NIME design processes; that the assessment tools proved to be consistent for the evaluation of *systems* and *performances* aspects of NIME; and that the outcome of the evaluation translated into a traceable progress in the students' DMIs.

5.1. Introduction and motivation

In the previous Chapter we have explored two of the three SID aspects that arose from our first experiment with the *b-Reactable*: *perceptualization* and *mapping complexity* of sonic designs for implicit PhyComp. Now we move towards the investigation of the third aspect: how physiology-based implicit interaction can *meaningfully* support musical expression. In order to achieve this goal, we propose to update the *b-Reactable* according to the guidelines of *perceptualization* and *mapping* discussed in the previous Chapter, and to test it in a music performance context. However, as our objective is to identify concrete contributions of implicit PhyComp to NIME, a thorough evaluation on key design aspects (both at system and music performance levels) is required to determine its potential for being perceived as an expressive component of a DMI, through which the player can produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to.

In this Chapter we therefore propose a framework for NIME design and evaluation, that is also meant to inform teaching and iterative design processes. In order to test the performance of this framework in a real world scenario, we deploy it in a one-trimester NIME master course focused on the exploration of expressiveness and on the crucial importance of the mapping component in the NIME creation chain. In this context, participants (i.e. students) prototype DMIs in groups, using a quite restrictive setup consisting only of smart-phones controllers and the Pure Data (Pd) programming language. This is done following a complete hands-on and self-reflective approach, in which the students (i) design a DMI (with the aforementioned important and predefined constraints), (ii) perform with the instrument in front of the rest of the class, and (iii) evaluate these performances as *listeners*, in an iterative process.

This Chapter is structured as follows. We first present an overview of the existing NIME frameworks in the context of education, design and evaluation. We then introduce the course, describing its context and its peculiarities. We describe the evaluation methods developed for assessing the projects created by the students, detailing how we apply this evaluation to inform iterative design, and we analyze and discuss the obtained results. We conclude discussing relevant findings and challenges, the manners in which the proposed framework can be applied to the evaluation of implicit PhyComp (concretely, the updated version of the *b-Reactable*) and how it could inform other NIME practitioners, educators or designers.

5.2. Designing and evaluating NIME

Since its birth in 2001, the NIME conference gathers researchers and practitioners that initially attempted to answer the question of *how to better play computer music* by exploring connections with the better-established field of human-computer interaction (HCI). As the NIME field matured, it integrated knowledge and practices from different disciplines. Moreover, in parallel to this maturation process, there has been a naturally growing interest in finding methodological and design frameworks that help evaluating the quality or the suitability of NIME, guiding teaching and iterative design processes.

In 1999, Michel Waisvisz, artistic director of the Dutch center for research and development of new musical instruments -STEIM- from 1981 until his death in 2008, and one of the few undeniable NIME virtuosi, highlighted the apparent lack of progress and the *permanent reinvention of the wheel* that seemed to be going on in the realm of musical gestural controllers: “A growing number of researchers/composers/performers work with gestural controllers but to my astonishment I hardly see a consistent development of systematic thought on the interpretation of gesture into music, and the notion of musical feed-back into gesture.”[1] [Cadoz et al., 2000].

5.2.1. Learning NIME design

The way DMIs are currently created and evaluated is strongly related to how design and validation methods are taught in different NIME venues. Sixteen years after Waisvisz’s concerns, the design of DMIs no longer relies solely on the efforts of some romantic and isolated pioneers. While a course on controllers taught at Stanford’s CCRMA was already presented in the first NIME Workshop in 2001 [Verplank et al., 2001], in the last years numerous NIME design courses have sprung up at universities around the world. A special workshop devoted to NIME education took place at NIME 2011, with the aim of providing a structured forum for NIME educators to share their approaches, experiences and perspectives on teaching NIME curricula [Gurevich et al., 2011]. While we do not aim at identifying the main differences and peculiarities between existing NIME design courses, we proceed to summarize the common features they typically share:

- Courses tend to be taught at the beginning of graduate or senior undergraduate levels [Lyons and Fels, 2013];
- They tend to be very multidisciplinary, often bridging the gap between art and science education [Lehrman and Ryan, 2005], and thus agglutinating students from very different backgrounds and different levels of knowledge (e.g. fine arts, music, computer science, engineering, interaction or product design, etc.).
- While some courses are more closely defined and more knowledge oriented than others (i.e. the competences to be acquired during the course may include a given set of tools, technologies or procedures), they mostly tend to be project oriented and students learn what they need in order to develop their own projects, which

[1] Waisvisz, Michel. "Gestural Round Table". Available at `www.steim.org/steim/texts.php` (accessed on October, 2015).

are then finally presented in live performances or demo scenarios [Lyons and Fels, 2013].

- While very often these projects give great freedom to the students - typically only limited by the technical resources and the know-how available at each center - the lack of a shared technological knowledge among these students, makes technological topics prevalent (e.g. how to use different sensors; how to connect them to a micro-controller; how to synthesize/process sound in a programming environment such as Pd, Max/MSP, SuperCollider or Chuck, etc.) and more important than design aspects or more conceptual criteria.
- Finally, most courses tend to instruct on how to create new DMIs, most often eluding the question of how to improve them or make better ones, whatever this adjective may mean.

These aspects lead us to the next section, in which we present an overview on the existing frameworks for the design and the evaluation of DMIs.

5.2.2. NIME design and evaluation frameworks

Is Waisvisz's initial quote still valid? Are the designers of DMIs (who may also be performers, composers and/or researchers) still blindly working in a field which shows no consistent development of systematic thought? Inseparable from concepts so complex and elusive as music or taste, the NIME realm may indeed always remain an area impossible to reduce and systematize, or as Perry Cook put it in the first NIME workshop in 2001, "musical interface construction proceeds as more art than science, and possibly this is the only way it can be done" [2001]. Creating DMIs is indeed in many respects, very similar to creating music. It involves a great deal of different know-how and many technical issues, while at the same time, like in music, there are no *inviolable laws*. But even if we may agree in the fact that the NIME discipline will never become a science, this should not prohibit us from thinking about it and analyzing its outcomes, and in particular, it should not prevent us building on the successes and the failures of experienced practitioners. Not unlike much research in HCI culminates in lists of guidelines and/or principles for design (and/or evaluation of design) based on research or practical experience relating to how people learn and work, it comes as no surprise that the first tentative NIME design frameworks have been mostly proposed by experienced digital luthiers [Jorda, 2005].

In his aforementioned paper, Cook also delivers his first principles for designing computer music controllers. As pointed out by O'Modhrain [2011], Cook's paper (as well as most of the following frameworks) which includes statements such as "copying an instrument is dumb, leveraging expert technique is smart", sets the goals for desirable properties of successful DMIs, yet saying little about how to achieve these goals. In this regard, Jordà proposes a conceptual framework that could serve in evaluating the potential, the possibilities, and the diversity of new digital musical instruments, focusing on the *expressive possibilities* these instruments can offer to their performers. This framework discusses in depth several DMIs desirable properties or goals such as the instrument's *playability*, *learnability*, *musical efficiency*, *variability*, *reproducibility*, *explorability* or *diversity* (the ability of an instrument to support diversity in musical style and performance), and how each of these different properties can promote/support different performance needs and approaches, such as the ones desirable in a instrument for novices, or the ones required for developing virtuosity [2005].

Dobrian and Koppelman , when approaching expression in digital musical performance, also stress the importance of *virtuosic mastery*, and how this can be promoted through intuitive but complex gesture-sound mappings (and obvious long-term practice) [2006]. All the above mentioned authors elude however the delicate issues of how to clearly attain these design goals and how to objectively evaluate them. The task of evaluating DMIs is, in fact, strongly linked to that of designing them, and knowledge gained in any side of the equation should complement the other. It is also clear that traditional evaluation methodologies coming from the field of HCI tend to be unsuited to the even more subjective evaluation of DMIs [Bellotti et al., 2002]. And yet, directly inspired by HCI, Wanderley and Orio [2002] provide one of the first sets of guidelines to aid in the selection suitable tasks for evaluating DMI designs. Although these guidelines and tasks do not constitute in themselves methods for evaluation, they definitely bring observations that can constitute good evaluation starting points.

More recently O'´Modhrain [2011] presents an excellent and detailed overview of previously existing DMI evaluation frameworks that we encourage the reader to consult, and proposes the evaluation of DMIs from the diverse and complementary perspectives of all the *stakeholders* involved in the process. This list includes performers, audiences, composers, instrument builders, component manufacturers and customers, and assumes that each of these stakeholders may have different ideas of what *evaluation* may mean, and that DMI designs should be therefore tackle from these multiple perspectives. O´Modhrain's paper also provides a list of goals such as *enjoyment*, *playability*, *robustness* or *achievement of design specifications*, meant to be confronted from the diverse perspectives of each stakeholder.

Following this stakeholder-based approach, Barbosa et al. [2011] deepen in evaluation methodologies from the perspective of the audience, while Gurevich and Cavan Fyans [2011] focus on the relationship between performers and digital systems and on the spectatorś' perception of these interactions. Among other recent publications, Gelineck and Serafin [2012] insist on the importance of longitudinal studies carried along longer periods of time, in an attempt to study the development of virtuosity. Along similar lines, Marquez-Borbon et al. [2011] study the evolution of skill development interviewing and following a group of users for several months, while they also propose the conception and design of *experimental DMIs* for specific evaluation purposes (as opposed to artistic purposes). Kiefer [2012] also uses his own DMIs for proposing the combination of HCI inspired methodologies and grounded theory methods for assisting the design, use and evaluation of creativity support tools with a focus on multi-parametric DMIs. In essence, while the search for solid and grounded design and evaluation frameworks is one of the main trends in current NIME research, general and formal methods that go beyond specific use cases have probably not yet emerged.

5.3. Framework and deployment

We now propose a methodology for NIME design and a set of evaluating tools intended to inform the design process. In order to test and refine all aspects of this framework, we deployed these methods in a one-trimester graduate course called Real-time Interaction, coordinated by Prof. Sergi Jordà. Following we provide a description of the course, background of participants, content structure, and specific results.

5.3.1. Context

The course was compulsory for two one-year master programs at Universitat Pompeu Fabra, Barcelona: (i) Sound and Music Computing -SMC-, and (ii) Cognitive Systems and Interactive Media (CSIM). For this reason, the background and interests of both types of students tend to be quite different. SMC students have clear musical interests, most often playing one or several musical instruments, and tend to come also from more technical and engineering backgrounds, thus often having prior experience in computer programming. CSIM students, on the other hand, come from a mix of backgrounds (psychology, sociology, humanities, design, mathematics, architecture, etc.) and most often do not have any prior experience in music performance nor in computer programming. Finding a suitable balance to satisfy such diversity is not an easy task. For this pragmatic reason, previous deliveries of the course did not explicitly focus on NIME design, but rather in analyzing the characteristics and differences of real-time interaction in different contexts (e.g. NIME, video games, augmented reality, etc.) from a more conceptual point of view. From 2013 on, we decided to face the challenge. Would it be possible to conceive a more hands-on course that (i) from a technological perspective, would be challenging and yet feasible for all types of students (musicians vs. non-musicians, programmers vs. non-programmers), and that (ii) from a conceptual and theoretical perspective would also provide enough food for thought and useful learning for all participants?

Taking into consideration some of the properties that constitute the intrinsic and more relevant features of real-time interaction when compared to more conventional WIMP interaction, namely the multidimensionality, multi-modality and the continuity of the input space [Jordà, 2008] we decided to focus on the systematic exploration of two advanced NIME topics that are tightly related to the scope of this dissertation, assuming that the conceptual challenges they would provide would not be substantially minor for the SMC students than for the CSIM students. Also the later could benefit from some of the learnings, being able to subsequently extrapolate them to their specific areas of research. The two chosen topics for exploration were *expressiveness* and the crucial importance of the *mapping* component in the NIME creation chain.

In this context, the introduction of an evaluation process within the course had two main motivations. From a pedagogical perspective, there was a clear objective of making the students fully aware of the intrinsic difficulties of evaluating complex and creative interaction contexts. On the other side, it allowed us to research and validate our framework in a relevant, hands-on environment. Through this approach, we aimed at:

1. Experiment with evaluation methods in which participants would swap between different roles (i.e. *designers*, *performers* and *listeners*) and analyze how previous music knowledge would affect each of these roles (following O´Modhrain's ideas on stakeholders [2011]).
2. Investigate to what extent the proposed evaluation method could effectively inform *iterative design processes.*
3. Shed some additional light on the elusive concept of *expressiveness,* needed to systematically assess the meaningful contributions of physiology-based implicit interaction to the NIME field.

From a technical point of view, and unlike most NIME courses that tend to offer a free or at least wide enough approach to technology, we decided to apply a restrictive approach based on two technological tools, namely smart-phones as controllers and the

Pd programming language for audio synthesis and processing. This strategy responded to two main reasons: eliminate all accessory technical information that would probably add confusion to the least tech-savvy students, and to carry on our framework evaluation in a reasonably constrained and controlled scenario.

5.3.2. Students profile & structure of the course

Real Time Interaction takes place in the first trimester (October-December) of the academic year and is composed of 12 weekly 2-hour classes. Thirty five students took the course that year, with approximately half coming from each of the two above mentioned master programs. SMC students had some musical knowledge, playing one or several instruments and being familiar with Digital Audio Workstations and electronic music production. This information was obtained through a questionnaire as described in Section 5.3.3. Most also had some computer programming knowledge and several were even familiar with Pd or Max/MSP, although none had worked on real-time electronic music performance. With some exceptions, most of the CSIM on their side, did not have any prior musical experience. Although some were engineers acquainted with computer programming, none of them had ever worked with digital audio or data flow programming languages. Special efforts were needed in order to find a right balance between novelty and viability that would satisfy almost everyone, a goal that was almost achieved. No student found the course too trivial and only two complained about its difficulty. Table 5.1 shows the topics covered in the course, along with the recommended readings and the assignments required for each of the 12 sessions.

The topics covered, which did progressively deepen from the more general to the more specific, can be synthesized as follow: starting with the concept of *interaction* and the problems deriving from the evaluation of interactivity (session #1), the special characteristics of *real-time interaction* were highlighted (#2), then the particular case of musical interaction with *DMI* was studied (#4), for subsequently focusing on *timbre control* and *navigation* (as opposed to more traditional pitch-based control), trying to elucidate the meanings of *expressiveness* (#5), and investigating how more complex (especially non-linear - many-to-many) mappings [Hunt et al., 2003, Kiefer, 2012] could affect achieving this objective (#6). To encourage participation and discussion, students were asked to read several papers before each new topic was introduced, the full list of which can be also consulted in Table 5.1.

Additionally, all these concepts were put in practice by the students with progressively sophisticated implementations using Pd. These started from a simple *Theremin* with vi-

Session	Content	Readings for next session	Assignments for next session
0		[Rafaeli, 1988, Svanaes, 2013]	
1	Introduction and discussion on ‘interaction’ and the ‘evaluation of interactivity’	Start reading the Pd tutorial	Think about potential real-time applications
2	Real-time interaction (technical, perceptual and design issues)	Selected and abridged info on sound and digital audio (Hz, pitch, dB...)	Build basic Theremin with sine oscillator in Pd. Control pitch, amplitude and add a ‘nice’ and natural vibrato control
3	Pd hands-on exercises	Interactive music: [Chadabe, 1984]	Build monophonic synth with 2-3 continuous parameters. Don’t worry about the interface: just put sliders
4	Interactive music overview (historical, conceptual)		Find videos of ‘expressive’ performances (acoustic, electric, electronic...) with a focus on timbre control
5	Expressiveness. Timbre navigation videos: Tuvan singing, didgeridoo, wah-wah brass, electric guitar	Selected and abridged info on audio filters, subtractive and modular synthesis	Add filter and LFO to your synth. Download OSC app for your smartphone/tablet (IOS/Android)
6	MIDI, OSC, sensors and accelerometers. Connecting smartphones/tablets to Pd		Create minimalistic smartphone interface for your synth. No sliders; just continuous control from accelerometers, compass, 2D multi-touch... Focus on timbre; forget pitch.
7	1st performance and on-line evaluation questionnaire	Control: [Ryan, 1991, Pressing, 1990]. Mapping: [Arfib et al., 2002, Hunt et al., 2003]	Check the feedback from your colleagues and continue enhancing your synth
8	Mapping and non-linearity	[Jorda, 2005, Chapter 7]	Continue enhancing your synth
9	Pd hands-on: feedback, distortion, non-linear many-to-many mappings	[Tahiroğlu, 2011]	Focus on non-linearity and many-to-many mappings. Get ready for the 2nd performance
10	2nd performance and on-line evaluation questionnaire		Document and upload your final synth
11	Machine learning in HCI	[Fiebrink, 2011, Gillian et al., 2011]	
12	Evaluation methods in HCI and NIME.	[O’Modhrain, 2011]	

Table 5.1.: Structure and contents of the course [Jordà and Mealla, 2014]

brato control (#2-3) and went into several iterations of a *monophonic synthesizer* with increasing timbral control parameters (#3-4, #5-6), that later was controlled from a *smartphone/tablet* using the OSC protocol with increasingly complex mappings (#6-7, #7-8, #8-9). While during the first sessions students worked individually, after session 3 they created 11 working groups (of 2-4 students) that remained stable for the rest of the course[2]. Sessions 7 and 10 constituted the backbone of the evaluation method, since in these two sessions each working group performed a 2 to 3 minutes piece/improvisation that was evaluated by all the other students, as described in detail in the next section. Performances were video-recorded[3] and made available to the students for a more detailed evaluation (see Figure 5.1).

After the second performance (#10), session 11 was devoted to the use of machine learning techniques for NIME control mappings [8],[10]. Although one of the initial objectives when envisaging this course was to include a 3rd performance/iteration using these techniques, it turned out clear from the beginning that it would be impossible to grab as much content in a 12 weeks course, so this topic remained at the theoretical level and was presented to the students as a potential follow-up to their work. Finally, session 12 provided an overview of evaluation methods and issues in HCI in general and NIME in particular, which concluded with a discussion of the results of the evaluation.

5.3.3. Evaluation tools

The methods applied during the master course were designed to assess both the *System* and the *Performance* aspects of the developed projects. Through this approach, we were able to evaluate DMIs in different stages, and explore how this evaluation can inform iterative design. Learning aspects, however, were not assessed.

Twenty two students (7 females), mean age $25.3 (SD = 2.15)$ participated in the evaluation. Participants' demographic data (age and gender) and previous music knowledge (capability for playing music and electronic music) were measured at the beginning of the master course through an electronic questionnaire.

During each performance session, all participants (in the *listeners* role) completed a 5-point Likert scale questionnaire to rate both the *System* (the DMI itself) and the *Performance* (related to the use of the DMI and the quality of the musical output). The *System's* properties were measured according to 3 variables:

[2]The reports on the projects developed by each group are available at http://www.dtic.upf.edu/~smealla/phd_material.html

[3]The videos of the performances are available at http://www.dtic.upf.edu/~smealla/PhD_Material/videos.html

Figure 5.1.: Performances by working groups using the DMIs developed during the master course.

- **Mapping richness**. Statement: "I have found the control mapping rich and interesting".
- **Synthesis richness**. Statement: "I have found the sound synthesis rich and interesting".
- **Potential**. Statement: "The system shows great potential as a DMI".

Performance's aspects, on the other hand, were assessed through the following variables:

- **Musicality**. Statement: "I have found the performance musical".
- **Expressiveness**. Statement: "I have found the performance expressive".
- **Virtuosity**. Statement: "The performers were able to control de instrument as real virtuosi".

These variables, whose choice was influenced by Jorda [2005], O'Modhrain [2011], were previously debated in class to assure a consistent interpretation during the evaluation process. Each *listener* fulfilled the questionnaire after each performance (except their own). Together with the questionnaire, tags and comments about the projects were also collected.

5.3.4. Results

For analysis purposes, the sample was divided in two groups: High Music Knowledge (HMK, 15 participants) and Low Music Knowledge (LMK, 7 participants), and in two stages (1st and 2nd Performances). A Pearson Correlation analysis was applied to test the coherence and strength of the two Categories of variables (*System* and *Performance* properties). An analysis of Variance (ANOVA) was applied to find significant differences between 1st and 2nd Performances. For this multiple comparison analysis, Bonferroni correction of significances was applied and alpha was fixed at 0.05 for all statistical tests.

Correlation Analysis

The correlation analysis showed significances for both 1st and 2nd Performance stages. More specifically, *Musicality*, *Mapping Richness* and *Synthesis Richness* were positively correlated for in both 1st and 2nd Performances. On the other hand, *Potentiality*, *Expressiveness*, and *Virtuosity* also showed a significant correlation for both stages. Table 5.2 shows the direction and strength of significant correlations in both stages.

Between Stages Analysis

When analyzing differences between 1st and 2nd Performances without considering previous music knowledge (all *listeners* together) three variables (*Potentiality*, *Expressiveness* and *Virtuosity*) reached significance for 3 of the 11 projects (see Table 5.3). When

Variable	Mapping	Synthesis	Potential	Musicality	Expressiveness	Virtuosity
Mapping		$r = .450^*$		$r = .334^*$		
Synthesis	$r = .450^*$			$r = .502^*$		
Potential					$r = .505^*$	$r = .391^*$
Musicality	$r = .334^*$	$r = .502^*$				
Expressiveness			$r = .505^*$			$r = .591^*$
Virtuosity			$r = .391^*$		$r = .591^*$	

Table 5.2.: Pearson correlations for *System* and *Performance* variables. Only significances are shown (* $p < 0.01$) [Jordà and Mealla, 2014]

DMI	All	HMK	LMK
4	Expressiveness ($F_{(9.20)} = 1.43$)*	No significances	No sig.
5	Potential ($F_{(5.90)} = 1.44$)* Expressiveness ($F_{(6.30)} = 1.6$)*	Expressiveness ($F_{(9.20)} = 1.62$)*	No sig.
11	Expressiveness ($F_{(9.20)} = 1.51$)* Virtuoso ($F_{(4.60)} = 2.10$)*	Virtuoso ($F_{(7.18)} = 2.1$)*	No sig.

Table 5.3.: Analysis of Variance between 1st and 2nd Performances for the whole sample (All), and High and Low Music Knowledge (HMK/LMK). Only significant differences are shown. Alpha for Bonferroni-corrected significances set at $p < 0.05(*)$ [Jordà and Mealla, 2014]

analyzing between-stages differences according to previous music knowledge, a different picture emerged. For the HMK group, only two variables (*Expressiveness* and *Virtuosity*) reached significance for two projects (see Table 5.3). The LMK, on the other hand, did not show any significance.

5.4. Discussion

5.4.1. Analysis of the Results

The statistical analysis showed that the proposed categories (*Performance* and *System*) were coherent and consistent independently of *listeners'* previous music knowledge. In spite of these findings, two variables (*Musicality* and *Potential*) showed to be correlated with the opposite category, meaning that *Musicality* was significantly correlated with the variables within the *System* category, and *Potential* with those of the *Performance* category. This can be explained by a certain level of ambiguity in the operationalization of these variables in the questionnaire (i.e. the way in which the questions were formulated). This has occurred despite the musical background of *listeners*. The outcome of this first study, however, helped us to rearrange the variables according to their coherence, making them a good fit for our next study on implicit PhyComp with the *b-Reactable*. The ANOVA, on the other hand, showed that the evaluation of DMIs through the questionnaire is sensitive to participants' music knowledge. In this regard, certain musical background facilitates a proper understanding of the framework and the questionnaire. We also have to mention that the disparities between groups (LMK was half of the size of the HMK group) represents a limitation in this study, therefore we cannot fully describe the real impact of musical knowledge in the empirical use of the proposed framework. In sum, future studies using our framework should address similar educational settings, bigger group sizes, and music knowledge should be normalized. Moreover, the addition of qualitative tools such as interviews and open questionnaires could complement statistical validity (as will be seen in the next Chapter). This could bring a better understanding of how the proposed framework can contribute to improve the DMI design process, beyond the natural enhancement resulting from mere iteration.

5.4.2. Analysis of the Initial Objectives

As stated in the Section 5.3, the main objectives of this course/deployment covered design, pedagogical and research issues. From the designs and pedagogical perspectives, the focus was on the role of *Expressiveness* and *Mapping* in the DMI design process, and on the value of our evaluation tools to feedback meaningful information to the iterative design process. In this regard, the Case Study has shown that the students got actively engaged in a design process, with the evaluation informing the development of prototypes. Although only 3 out of 11 projects reached significant differences when comparing the 2 design stages, almost all DMIs showed improvements after iteration. A bigger and equilibrated sample will help to reflect the contributions of these evaluation tools in the design process.

It is also worthy to discuss to what extent the design guidelines imposed during the course either constrained or helped students to focus on core aspects of the DMI design chain. In this sense, the fact that all groups achieved operative DMIs shows that the proposed guidelines helped to leverage the students' background, fostering collaboration between students with different skills. Our research goals, on the other hand, aimed at studying how these methodologies can cover different stakeholders. Although we present methodologies mainly focused on *listeners* and their music knowledge (meaning that although all participants exerted the 3 roles, they only evaluated from what they heard from their colleagues performers, without testing the other DMIs themselves) the results show the relevance of *listeners'* perception for informing iterative design of DMIs. Future studies should broaden the scope of this study by also considering *designers* and *performers*. Finally, the internal analysis of each project was not covered by this study. Since we did not analyze the relation between the implementations and the feedback received, no conclusion can be taken on the interactions between *mapping* and *expressiveness* at this point. Data was collected in this regard, in the form of smartphones GUIs, Pd patches (which incorporated all the mappings), video recordings and written reports, so future work can be devoted to such analysis.

5.4.3. Guidelines for future work

A number of guidelines for future work can be envisioned in response to the faced challenges and problems. Firstly, the proposed evaluation tools have to be tested in different NIME design scenarios (e.g. workshops) beyond the specific master course presented in this Chapter. Regarding the grouping of participants by music knowledge, experimental groups should be leveraged for achieving better statistical validation, and for analyzing in depth the effect of musical background in the design, performance and evaluation process. In the same direction, the roles of *designer* and *performer* could be detached and analyzed separately, together with the influence of music knowledge for both stakeholders. In this regard, we envision an experiment where performers could select their favorite DMIs designed by other working groups and perform with them for later evaluation as *performers*. Concerning the design guidelines, the proposed methods should be tested with other design constraints, and future work should also deal with the analysis of the DMIs themselves, to go beyond *listeners'* perception.

Finally, it is important to mention that the study presented in this Chapter represents a first step in the creation of a NIME design and evaluation framework. Our current approach could be complemented and expanded with qualitative methods such as interviews and focus groups. In fact, we have been applying this framework in a master course for the last two years, extending the evaluation from the *listener* perspective to the standpoint of *designers* and *performers,* including interviews group discussions.

5.5. Conclusion & next steps

In this Chapter we have presented a framework for NIME design and evaluation meant to inform the iterative design processes. These methods have been applied in a Case Study focused on the exploration of *expressiveness* and *mapping* as crucial components in the NIME creation chain, and making use of a quite restrictive setup consisting only of smart-phones controllers and the Pd programming language. Working groups were formed, and a 2-step DMI design process was applied, including 2 performance stages. The evaluation tools assess both *System* and *Performance* aspects of each DMI, according to *listeners*' impressions during each performance stage. *Listeners*' previous music knowledge was also considered.

The learning and knowledge that we have gained through this iterative methodology is threefold:

1. All the students (some of whom had never performed music, neither programmed computers) were able to effectively engage in the NIME design processes, being able to develop working DMIs that fulfilled all the asked requirements;
2. The assessment tools proved to be a consistent method for the evaluation of systems and performances, in the context of our master course.
3. The fact of informing the design processes with the outcome of the evaluation, showed a traceable progress in the students' outcomes.

Although these findings were obtained in the specific context of a NIME course, we believe that several of these solutions and learnings could be extrapolated to more generic contexts, being other NIME or even HCI courses, and used to inform teachers, designers and practitioners in general. Our next Chapter will therefore leverage on this design and evaluation framework to determine to what extent implicit PhyComp is perceived as a meaningful component of a DMI, through which the player can produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to.

6. Enhancing NIME with Implicit Physiological Computing

This Chapter addresses the issue of *meaningfulness* that, together with *perceptualization* and *mapping complexity*, constitutes one of the main aspects of physiology-based implicit sonic interaction in the context of this dissertation. The goal of this Chapter is to systematically explore how implicit PhyComp contributes to the design of a digital musical instrument (DMI) in a *meaningful* way, implying that it will be perceived as an expressive component of the DMI, through which the player is able to produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to. In order to do so, we create a new version of the *b-Reactable*, that incorporates a number of features informed by the results of our previous experiments in *perceptualization* and *physiology-to-sound mapping*, namely a parameter mapping approach, end-user personalization of *physiopucks*, and a more complex implicit psychophysiological input (valence and arousal estimated through EEG).

We test this new incarnation of the *b-Reactable* in an expressive context (i.e. music performance) involving 15 participants with different levels of music experience (novice, knowledgeable and expert) who perform musical improvisation exercises under two conditions (Global and Local implicit interaction). Four different measures are collected for evaluating user experience and *meaningfulness*: affective data (valence/arousal) estimated through EEG, behavioral data (based on the use of the *physiopucks*), System and Performance aspects of the DMI (in accordance with the framework defined in Chapter 5) and open interviews on user experience.

The main results show that our affective estimations are valid for the context of music performance, and that participants use these implicit sonic interactions (both Global and Local) in a distinctive and meaningful manner. Subjective, behavioral and psychophysiological data show that Global and Local implicit interaction are perceived in significantly different ways according to participants' previous musical experience, with preference for the latter.

6.1. Introduction and motivation

In the previous Chapter we have proposed and tested a framework for the design and evaluation of new interfaces for musical expression (NIME), also meant to inform teaching and iterative design processes. By assessing System and Performance components of digital musical instruments (DMI), this framework offers methodological tools to systematically explore how physiology-based implicit interaction could contribute to NIME design in a *meaningful* way, implying that it will be perceived as an expressive component of the DMI, through which the player is able to produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to.

At this point, it is important to recall the outcomes of our first study with the *b-Reactable* (Chapter 3). It shows that physiology-driven sonic interactions embedded in *physiopucks* improve user experience in terms of motivation. However, its sonic designs were rather simple and limited in terms of *expressiveness* (white noise shaped by EEG alpha activity and BPM control of the overall music composition through ECG), and we did not directly assessed how *meaningful* these specific sonic designs were for performing music, according to players with different musical background.

Guided by the *perceptualization* and *mapping complexity* findings presented in Chapter 4, we are now in a better position to integrate new sonic strategies into the *b-Reactable* to further study the *meaningfulness* of physiology-driven, implicit interaction in a musical context. In order to do so, we update the *b-Reactable* with the following features:

- **Perceptualization**: the tangible objects used to deploy physiology-based sonic interactions (*physiopucks)* now work under a parameter mapping approach (the sonic design that showed better *perceptualization* in Chapter 4) based on transposing affective states (valence/arousal) into a sound output according to timbre characteristics. As end-user personalization has shown to improve the *perceptualization* of implicit sonic interactions, we add a feature called Gain, that allows players to interpolate between physiological and gesture input, thus fostering multimodality.
- **Mapping complexity**: we use a two-dimensional representation of user affective states (valence/arousal) estimated through EEG, as implicit physiological input. This is a more complex data stream compared to the one used in the first prototype (EEG alpha activity). Given that the manner in which sonic interactions are controlled constitutes a crucial component in the NIMEs creation chain (see Chapter

5), two implicit interaction strategies are applied for driving sound through PhyComp (i.e. Global and Local implicit interactions).

We test this new incarnation of the *b-Reactable* in an expressive context (i.e. music improvisation) involving novice (N), knowledgeable (K) and expert (E) musicians. Moreover, we also ran comparative studies with the standard Reactable, as we did in Chapter 3. Participants are asked to perform musical improvisation exercises under two conditions, defined by the above mentioned sonic implicit interactions (Global and Local). Four different measures were collected for evaluating user experience and *meaningfulness*:

- **M1**: affective data (self-reported and physiological measures of valence and arousal)
- **M2**: behavioral data, based on the use of the *physiopucks*
- **M3**: system and performance aspects of the DMI, as defined by the framework presented in Chapter 5.
- **M4**: open interviews on user experience

Our main hypotheses are:

- **H1**: Players will be able to *perceive* the provided physiology-based implicit sonic interactions based on valence and arousal.
- **H2**: Players will perceive these sonic interaction as a *meaningful music resource* during performance.
- **H3**: Players' perception and preference on sonic interaction will vary according to their *music knowledge*.

This chapter is structured as follows. We first provide a short summary on specific topics (feedback design in PhyComp, EEG and emotion estimation, expressiveness and HCI) that although are present all across this dissertation, will help the reader to put the study in context. Then we provide a detailed description of the new *b-Reactable* and its implementation, followed by the use case definition, methods and experimental protocol. We then present the results of the experiment and conclude discussing the main findings of this study and future work.

6.2. Context

6.2.1. Feedback design for physiology-based implicit interaction

As shown in our initial study with the *b-Reactable* (Chapter 3) and in the perceptualization experiments of Chapter 4, when an interactive system uses physiological activity (e.g. EEG) to implicitly and indirectly drive sonic interactions in real time, feedback design becomes crucial. As explained in Chapter 2, the feedback design process in PhyComp is traditionally defined by its *directionality.* In this sense, research efforts have mainly focused on two modes: positive and negative feedback control[1][Carver et al., 2000] and their combination, as they are not mutually exclusive. Negative feedback creates behavioral stability by reducing the discrepancy between the input state (e.g. physiological correlate of engagement) and a desired standard (a given level of engagement). Negative feedback is ideal for adaptive systems that are designed to keep the user in the *flow* [Csikszentmihalyi and Csikzentmihaly, 1991]. Think for instance on a first-person shooter; negative feedback would be desirable for parents that will like to keep the stress of their children low if they game becomes too challenging (i.e. more user stress, less difficulty).

Positive feedback, on the contrary, is designed to amplify the discrepancy between the input state and the desired state in an exponential fashion. This leads to performance instability, and it may therefore be used to adjust the desired target state upwards as the user becomes more involved with the task. If we go back to the fist-person shooter example, positive feedback would be preferred by a gamer willing to improve her game skills (i.e. more perceived difficulty will trigger more challenging obstacles).

Both types of feedback can be combined for toggling unstable episodes of skill acquisition (positive feedback) and stable moments for skill consolidation (negative feedback). Design decisions in this regard are essential, especially for goal-oriented HCI domains such as video games or training applications.

However, together with feedback *directionality*, the dimension of *feedback complexity* is also of utter importance, as demonstrated in Chapter 4. In the concrete case of SID for implicit PhyComp, this complexity is defined by the number of physiological inputs and sound parameters used to produce and display system adaptations. In this Chapter we explore this dimension because, whereas *directionality* is particularly important when the agenda of the system is based on training (e.g. biofeedback) or a goal oriented application (e.g. an adaptive video game), *feedback complexity* plays a major role when PhyComp coexists with other input methods (e.g. tangible interaction, as in the case of the *b-Reactable*), and in expressive HCI domains like NIME. As pointed out by Jordà and Mealla [2014], mapping strategies are crucial for expression and for domains where content exploration or creation are as relevant as task solving. In this context, for instance, a performer playing alone might prefer a physiology-to-sound mapping with a great *complexity* to drive several musical operations overtly and automatically, while fo-

[1]For a comprehensive re- view on this topic, please refer to Fairclough [2009]

cusing on specific musical aspects explicitly. For collaborative performance, on the other hand, a reduced *complexity* might be preferred for avoiding undesired sonic outcomes.

6.2.2. Music, EEG and emotion estimation

According to Scherer and colleagues [1984], emotions can be conceived as a process that consists of various components: cognitive appraisal, physiological activation, motor expression, behavior intentions, and subjective feeling. Emotional states can be therefore described as particular configurations of these components. For a long time, cognitive sciences have been studying the foundations of emotions. More recently, computational models have also been proposed and applied in several domains such as music [Schubert, 1999], body movement [Castellano et al., 2007], and films [Fredrickson and Branigan, 2005]. There are distinct approaches and techniques used to generate music with appropriate affective content. For instance, Livingstone and Brown [2007] established relations between music features and emotions, based on the findings of earlier studies by Emery Schubert [1999]. To this end, both emotions and a set of musical structural rules were represented in a two-dimensional emotion space with an octal form (see Figure 6.1) know as *the circumplex model of affect*, as proposed by Russell [1980]. Each emotional expression is placed at an approximate point in a two dimensional emotion space constituted by valence (positive and negative) and arousal (active and passive) levels.

A recent work by Cichocki BCI lab [Valenzi et al., 2014] compares state-of-the-art algorithms and electrode placement for emotion estimation. It confirms the well established findings that frontal EEG alpha asymmetry is linked to the withdrawal model and can be used to estimate emotional valence [Davidson et al., 1990]. Valenzi's classification also shows the known association between higher brain activity (arousal) and lower alpha band activation [Lindsley and Wicke, 1974]. Several affective estimation systems have been implemented following these considerations, both offline [Takahashi, 2004, Bos, 2006, Petrantonakis and Hadjileontiadis, 2010], and online [Ramirez and Vamvakousis, 2012, Lin et al., 2010].

6.2.3. Expressiveness in NIME and HCI

As mentioned in the introduction of this Chapter and as discussed in Chapter 5, expressiveness plays a major role in the perception of a given input method as a *meaningful* component of a DMI. In fact, previous studies on HCI have stressed on the importance of expressiveness for designing interactive systems [Dearden and Harrison, 1997], for

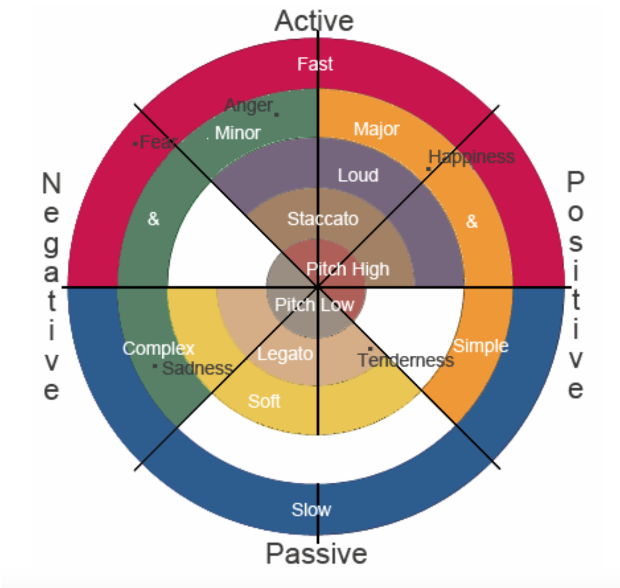


Figure 6.1.: The primary music-emotion structural rules graphed on the Two Dimensional Emotion Space [Livingstone et al., 2007]

real-time audiovisual performances [Hook et al., 2011], or to inform the design of control surfaces [Bodanzky, 2012]. However, the assessment of expressiveness in PhyComp systems remains almost unexplored. As mentioned in Chapter 2, most of the evaluation of PhyComp systems from an HCI perspective have focused on control and perceptual aspects in task oriented applications, as in the case of video games [Allison et al., 2007, Nacke et al., 2011, Nijholt et al., 2009]. This has brought valuable information about the integration of PhyComp into interactive systems, user preference and multimodal control. However, assessment of expressiveness in domain-specific contexts (e.g. NIME) is still pretty much uncovered, with the exception of biomedical studies on expressive EMG-based control of prosthetic devices [Shenoy et al., 2008].

As shown in Chapter 5, NIME provides a valuable corpus for understanding the role of expressiveness in interface design, HCI and computer music. These efforts have shaped NIME methodologies and design frameworks like the one presented in this dissertation, that aim to explore the potential and diversity of new digital musical instruments and sonic interactions, focusing on the expressive possibilities these interfaces can offer. By deploying PhyComp in a NIME context, we can then explore its *expressiveness* in depth, looking at System and Performance aspects, as proposed by our NIME design and evaluation framework (i.e. mapping and synthesis richness, potential, musicality, expressiveness, and virtuosity). In this manner, we can analyze how each of these properties support different user needs, such as the ones desirable in a instrument for novices, or the ones required for developing virtuosity. NIME thus appears as a perfect candidate for exploring the expressive possibilities of implicit PhyComp, and for studying it in a context that goes beyond goal-oriented scenarios. It is important to note that very few studies have been carried out in this direction, as most of the work on the intersection of NIME and PhyComp focuses on technical development, prototyping and interface design, and tend to elude the delicate issue of how to better attain *expressiveness* or how to empirically evaluate it.

6.3. Materials and methods

6.3.1. System design

In this section we describe the main aspects of the new version of the *b-Reactable*, including physiological signal acquisition and treatment, affective classifiers, sound display and implicit interaction strategies. From an interaction design perspective, three main goals drove our design process:

- *Robust signal acquisition with care on ergonomics*: the accurate measurement and processing of user's physiological signals are central for any PhyComp system. That can be achieved by combining professional physiological sensing hardware and previously validated processing algorithms. However, this type of equipment tends to be rather restrictive and invasive for music performers. In order to cope with these constraints, we use a Mitsar-EEG 201 amplifier[2], which complies with high monitoring standards, but still constitutes a portable solution for *out of the lab* EEG recordings.
- *Perceptible physiology-based sonic interaction*: following the outcomes of our *perceptualization* and *mapping complexity* studies (Chapter 4) we aim to design a DMI where implicit PhyComp interaction will be coherently perceived by participants, and will be meaningfully applied in a musical context. Since the link between music and emotions has been sufficiently demonstrated (see Section 6.2.2) we decide to estimate the real time affective response of participants (valence and arousal) from the EEG, and apply them to drive sound processors implicitly.

[2] `http://www.mitsar-medical.com/eeg-machine/eeg-amplifier-201/` (accessed on October, 2015)

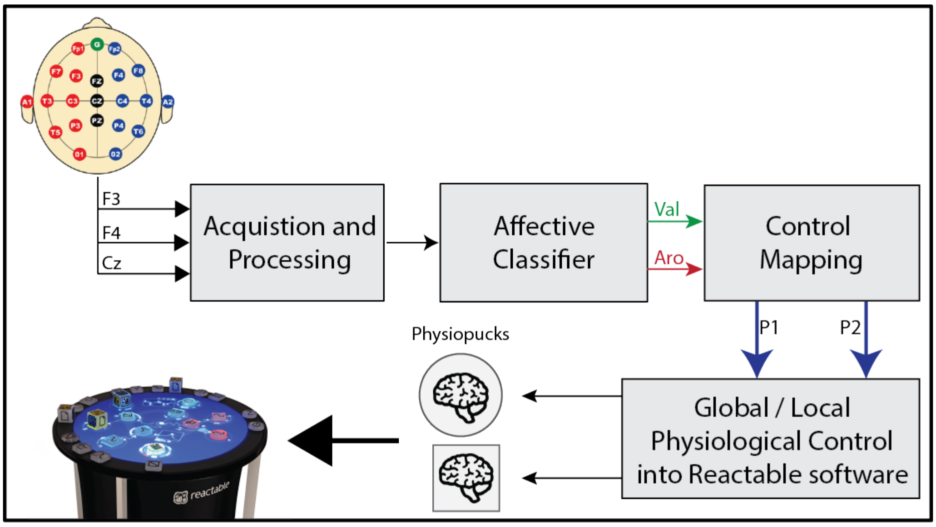


Figure 6.2.: The updated *b-Reactable*: system architecture diagram.

- *Multimodal control and personalization*: based on the affordances shown by the first *b-Reactable* for combining implicit physiological control with other input methods (i.e. gestures), in this study we go further, studying the preference among different implicit interaction modes, and comparing them with tangible interaction. As in Chapter 3, this is done by using the Reactable framework, which supports both tangible and multitouch inputs, and which has been widely used in music performance.

Figure 6.2 shows the main components of the system architecture.

6.3.2. Physiological signal acquisition and processing

Brain activity from participants is measured with a 21-channel EEG amplifier (Mitsar-EEG 201). The amplifier is interfaced to a desktop computer via USB at a sampling rate of 500Hz. A five-meters length secure wired connection allows the player to stand in front of the *b-Reactable* without the need of being stuck close to the amplifier. Sintered Ag\AgCl electrodes are placed in the player's scalp using conductive gel and an elastic cap (MCSCap) with plastic holders positioned according to the 10-20 International System [Homan et al., 1987]. We use F3 and F4 electrodes for recording, whereas Cz was used as reference. Unipolar recording is performed using the EEGStudio software[3]. The signal is band-pass filtered between 0.5 and 30Hz in real time, using a notch filter to

[3]http://www.mitsar-medical.com/eeg-software/erp-software/ (accessed October, 2015).

eliminate environmental noise. ICA decomposition of raw EEG is also applied for eye movement suppression. An FFT conversion with a size of 500 samples is applied to treat the signal in the frequency domain, and to calculate alpha and beta power spectrums. For reducing the effect of movement artifacts, we relied on alpha/beta ratios, adaptive rectification (max/min value updating) and a 5-second averaging window to smooth the signal.

6.3.3. Valence and arousal estimation

In line with the considerations mentioned in Section 6.2.2, we apply a two-dimensional affective model based on valence and arousal indexes (V/A), well-known within emotion research [Bradley and Lang, 1994]. V/A indexes are estimated as following: we use 1-second blocks transformed by FFT to the frequency domain. Relative power of alpha ($8-13Hz$) and beta ($13-30Hz$) bands are calculated for F3 and F4 channels. The use of F3 and F4 is meant to reduce montage time, and it is consistent with previous work on real time V/A estimation [Ramirez and Vamvakousis, 2012]. The valence estimator is calculated as:

$$EEG_{val} = F3_{alpha}/F3_{beta} - F4_{alpha}/F4_{beta} \tag{6.1}$$

The arousal estimator is calculated as:

$$EEG_{aro} = (F4_{beta} + F3_{beta})/(F4_{alpha} + F3_{alpha}) \tag{6.2}$$

The first 20 seconds of the recording are used to calculate initial max and min values for the estimator (mean $+/-1\sigma$). During the task, this 20-second buffer is updated with new values, and new max/min are recalculated every second. The values above max/min are rectified to extreme values. The estimated V/A values are then normalized to a range of 1..9. Finally, a 5-second moving average is applied to assure smooth changes of the data sent to the sonic engine. Previous affective estimations were developed and validated following the same model. Takahashi [2004] and Petrantonakis and Hadjileontiadis [2010] applied it for emotion recognition using multi-modal bio-potential (EEG, skin conductance, and face recognition), and Bos [2006] validated this approach using the IAPS and IADS databases, two libraries that contain emotion-annotated images and sounds respectively, which are widely used in emotion research [Lang et al., 1999, Lang and Bradley, 1999].

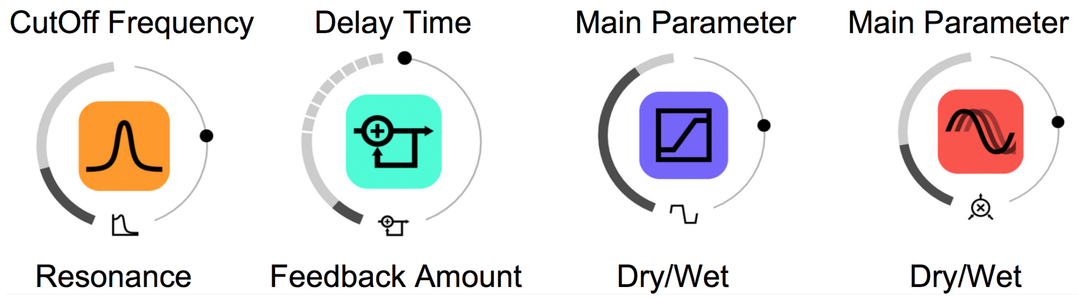


Figure 6.3.: Reactable sound processors capable to be controlled through physiology-based implicit interaction mappings. From left to right: low-pass filter, a delay, distortion and chorus. The upper and lower texts indicate their two control parameters.

6.3.4. Mapping design

Following previous research on the affective potential of musical parameters such as tempo, loudness, timing, timbre or vibrato [Juslin, 1997] and our own studies on *perceptualization* (see Chapter 4), we apply a parameter mapping strategy based on transposing affective states into a sound output according to timbre characteristics. The mapping is created ad hoc by applying two bi-linear interpolations that associated the four extreme coordinates of the two-dimensional V/A space, to the four extreme coordinates of a two-dimensional sound parameter space defined by $P1 - P2$, where $P1$ and $P2$ are directly proportional to V/A (see Figure 6.4). Mathematically, the desired interpolation point is calculated as:

$$f(x,y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)}(f(Q_{11})(x_2 - x)(y_2 - y) + f(Q_{21})(x - x_1)(y_2 - y) + f(Q_{12})(x_2 - x)(y - y_1) + f(Q_{22})(x - x_1)(y - y_1)) \quad (6.3)$$

Where x and y corresponds to valence and arousal respectively, and the four $f(Q)$'s are the known extreme values of the sound parameter space. We use this mapping to control four sound processors in the Reactable: a low-pass filter, a delay, a chorus and a distortion (see Figure 6.3). The mapping for each one is defined as follows:

- Filter:
  - Valence-Arousal (-1,-1) to Cut Off-Resonance (10,0)
  - Valence-Arousal (-1,1) to Cut Off-Resonance (10,1)

- Valence-Arousal (1,-1) to Cut Off-Resonance (127,0)
- Valence-Arousal (1,1) to Cut Off-Resonance (127,1)

- Delay:
  - Valence-Arousal (-1,-1) to Delay Time-Feedback Amount (0,0)
  - Valence-Arousal (-1,1) to Delay Time-Feedback Amount (0,1)
  - Valence-Arousal (1,-1) to Delay Time-Feedback Amount (8,0)
  - Valence-Arousal (1,1) to Delay Time-Feedback Amount (8,1)
- Distortion:
  - Valence-Arousal (-1,-1) to Main Parameter-Dry/Wet (0,0)
  - Valence-Arousal (-1,1) to Main Parameter-Dry/Wet (0,1)
  - Valence-Arousal (1,-1) Main Parameter-Dry/Wet (4,0)
  - Valence-Arousal (1,1) to Main Parameter-Dry/Wet (4,1)
- Modulator/Chorus:
  - Valence-Arousal (-1,-1) to Main Parameter-Dry/Wet (0,0)
  - Valence-Arousal (-1,1) to Main Parameter-Dry/Wet (0,1)
  - Valence-Arousal (1,-1) to Main Parameter-Dry/Wet (10,0)
  - Valence-Arousal (1,1) to Main Parameter-Dry/Wet (10,1)

For instance, for the low-pass filter, its two parameters $P1$ (frequency cut-off) and $P2$ (resonance) increase when the player's arousal augments. Accordingly, when the player's valence goes from $-$ to $+$, $P1$ and $P2$ will also increase.

6.3.5. Integration into the Reactable framework

Given the Reactable's modular approach, the integration of the above mentioned physiological control mapping is straightforward. Two new pucks (*Physiopucks*) are added to a standard Reactable, allowing players to deploy implicit physiological interaction without penalizing the standard Reactable performance mechanisms (i.e. controlling sound processes by means of tangible pucks and hand gestures). The control mappings are presented in detail in the following section. The additional software component that bridges the physiological parameters with the main Reactable software was programmed in Pure Data (Pd).

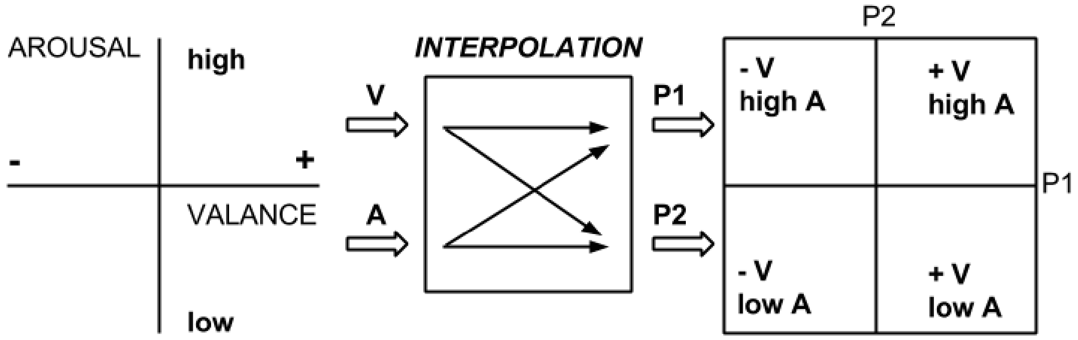


Figure 6.4.: Mapping for indirect physiological control: Player's affective states are transposed to sound parameters via two coupled bi-linear interpolators.

6.3.6. Global and Local implicit interaction

The *b-Reactable* presents two different implicit interaction modes based on affective data: Global and Local. The Global implicit interaction (GII) is represented by a squared-shape *physiopuck* (figure 6.5 A). When the player places it on top of the interface, the parameters of *all* sound processors running in the tabletop at that moment are driven by the player's V/A. Local implicit interaction (LII), on the other hand, is represented by a circular *physiopuck* (Figure 6.5 B), and it only affects the parameters of the processor *closer* to its position. Both *physiopucks* feature a Gain parameter that permits to modify and personalize the extent to which V/A mappings affect sound processors in the Reactable. The Gain matches gesture and V/A inputs according to the interpolation formula:

$$P_{VA} \times G + P_{TAN} \times (1 - G) \tag{6.4}$$

Where P_{VA} is the V/A parameter calculated after the bi-linear interpolation, G is the gain and P_{TAN} is the tangible parameter set by the player. When the Gain is at its maximum, the parameters of sound processors are fully controlled by V/A estimations. With the Gain at zero, sound processors are only controlled by gesture inputs, as in the standard Reactable. For intermediate Gain configurations, the parameter values of sound processors are the result of an interpolation between tangible settings and V/A estimations.

As shown in figure 6.5, the Gain is controlled differently for GII and LII. In the former, the Gain is controlled by rotating the object. When rotating clockwise (G towards 1),

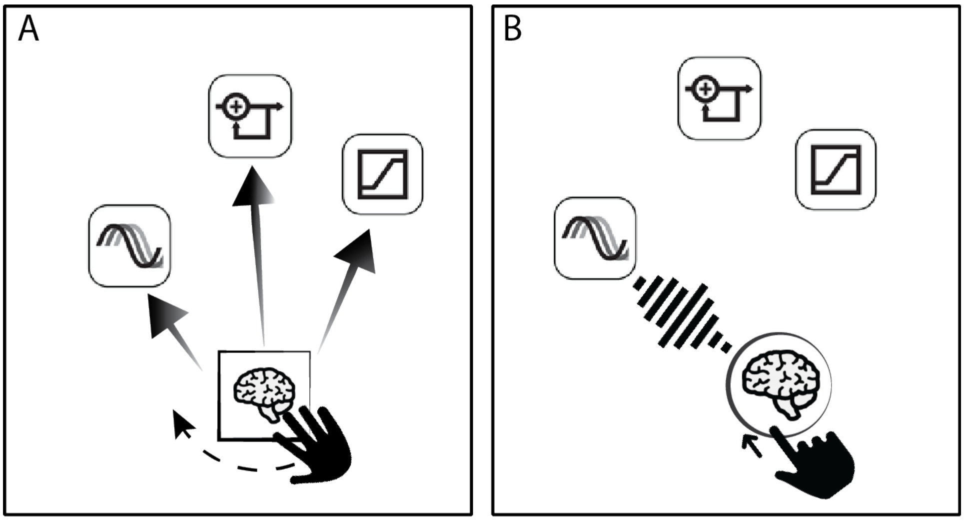


Figure 6.5.: Left panel: Global implicit interaction *physiopuck* (GII) affects all processors. Right panel: Local implicit interaction *physiopuck* (LII) affects only the closest processor

V/A is favored for driving sound processors. When rotating anti-clockwise (G towards 0) gesture control is favored instead. In the case of LII, the Gain amount depends on the distance between the *physiopuck* and its target (see figure 6.5 B), so that the closer the two pucks are, the greater is the Gain. LII works as a relative controller, as V/A estimations control a delta increase or decrease over the instantaneous value set by the user on the target music processor.

This design approach is inspired by the design guidelines of the standard Reactable, whose controllers also respond to global or local parameters (among others), and use the same puck shapes to make them readable and intuitive for users. This design strategy allows a seamless and coherent integration of implicit physiological interaction into the musical interface, without restricting gesture control. It also provides a moderation of over-corrective activation of the feedback loop, which is a key issue in PhyComp design for not constraining players' self-regulation of behavior and emotion to an excessive degree [Fairclough, 2009]. It is also important for expert performers of the conventional Reactable, which are used to control the interface by means of tangible and gesture inputs.

6.3.7. Case study with musicians

To validate our hypotheses, we designed a study with musicians performing improvisation tasks with *b-Reactable* in its two implicit interaction modalities (GII and LII).

Sample, setup and task design

Fifteen participants (12 males, mean age $29, SD = 9$), all of them with a declared experience playing digital musical instruments, performed with our system in its two setups: GII and LII. Each setup implied two music improvisation tasks where participants had to achieve a given V/A target, in terms of how the resulting sound would be perceived by an hypothetical *listener*: (1) high arousal - high valence composition (HAHV), and (2) low arousal - low valence composition (LALV). This V/A target were based on previous studies on emotion analysis using physiological signals [Koelstra et al., 2012]. Both setup and task order were randomized among participants. The study was conducted in accordance with the Declaration of Helsinki.

Measures

During the experiment we collected six types of measures:

1. *Musical experience*: self-reported measures on musical knowledge and previous experience playing the Reactable.
2. *Affective data (EEG)*: V/A estimations based on EEG recordings (as explained in Section 6.3.3)
3. *Affective data (subjective)*: self- reported V/A using the Self-Assessment Manikin (SAM) [Bradley and Lang, 1994].
4. *Behavioral data*: log files corresponding to the usage of the *physiopucks'* Gain in the two setups (GII and LII), including both the *amount* of implicit interaction *Gain* set by the user, and the *Homogeneity* of Gain values along each task (i.e. how much the participant changed the gain value during the task).
5. *System and Performance aspects of the DMI:* as defined by the framework presented in Chapter 5
6. *Open interviews*: performed at the end of each session to gather insights about user experience.

Figure 6.6.: Players performing with the multimodal musical system

Experimental protocol

Each experiment lasted about an hour; Figure 6.6 shows the setup used in each session, and Table 6.1 summarizes the protocol. First, the participant received a description of the experiment, including an explanation about the *b-Reactable* and its interactions with V/A states. Then, the participant was placed in a chair and the experimenter mounted the EEG cap and checked impedance ($< 5kOhms$). A 2-minute EEG baseline was recorded during the presentation of a relaxing auditory stimulus (sea waves) whose playback was triggered by the participant on the *b-Reactable*. This was followed by an affective state (subjective scale) and musical experience assessment gathered through a digital questionnaire. The next step was an open ended hands-on experience with the *b-Reactable*, assisted by the experimenter. Through it, the participant had the chance to learn how the system works, to explore the sounds stored in each generator object, and to try the *physiopucks*.

After re-checking electrodes impedance ($< 5kOhms$) Task 1 began, during which the participant was asked to improvise a musical piece of 5 minutes length, with one of the 2 system setups (GII or LII) following a given V/A target: high arousal - high valence (HAHV) or low arousal - low valence (LALV). After this first task, affective subjective ratings were assessed again and the impedance was checked ($< 5kOhms$). This was followed by Task 2, and once completed, the participant fulfilled again an affective subjective assessment, and answered the questionnaire on system and performance aspects of the DMI.

These two steps were repeated for each system setup (GII and LII), with an additional

Setup	Task 1	Task 2
Global Implicit Interaction	high arousal-high valence	low arousal-low valence
Local Implicit Interaction	high arousal-high valence	low arousal-low valence
Tangible	high arousal-high valence	low arousal-low valence

Table 6.1.: Experimental protocol. Each participant performed two music improvisations (high arousal-high valence, HAHV, and low arousal-low valence, LALV) with the b-Reactable in its two setups (Global implicit interaction, GII, and Local implicit interaction). A tangible only-condition (T) was added for non Reactable experts. The order of both tasks and setups were randomized.

Tangible only-condition (T) for non Reactable experts, which implied to perform the two above mentioned task using the standard Reactable. Each experiment therefore generated 4 or 6 musical improvisations per participant. At the end of each experiment, electrodes were removed and a 10-minute block for open questions was carried out to gather information about user experience.

Data analysis and pre-processing

For analyzing the influence of previous musical experience, we split participants into 3 groups: Novice (N, $N = 5$) with low experience playing electronic music; Knowledgeable (K, $N = 5$), with high electronic music experience but that had never played the Reactable before; and Expert Reactable users (E, $N = 5$). We applied a 3-way mixed MANOVA with *previous musical experience* as a between-groups factor, and *setup* (GII vs. LII) and *task* (HAHV vs LALV) as within-group factors. We combined 6 measures: V/A (subjective scale), V/A (EEG estimation, as described in Section 6.3.3), implicit interaction Gain and Homogeneity. For these multivariate analyses, Wilks´ Lambda Λ was used as the multivariate criterion. Alpha level was fixed at 0.05 for all statistical tests. Greenhouse-Geisser correction was used to correct for unequal variances. All data satisfied the normality criterion as verified by the Kolmogorov-Smirnov test.

For analyzing EEG V/A estimations, we used the estimated values before system calibration (rectification of min and max values and averaging as described in Section 6.3.3). For each recording, we calculated the means of the first and the last 15 seconds of each task, and the difference values between the two were submitted to the analysis. Hence, the greater and positive difference value meant higher arousal or more positive valence.

6.4. Results

6.4.1. Validation of affective estimations

While the implemented V/A estimation algorithms were already validated in previous studies, the baseline recording serves as a good measure to determine the validity of the applied EEG processing. For valence estimation, the means of the first and last 15 seconds were -0.04 and 0.30 respectively ($p = 0.05, t(14) = 2.11$). For arousal estimation, the means changed from 1.73 to 1.38 ($p = 0.01, t(14) = -1.74$). Both estimations showed that, during the relaxation baseline, arousal decreased and valence changed to more positive.

6.4.2. Differences between setups

The overall effect of the setup (GII vs LII) was significant at $\Lambda = 0.09, F(6,7) = 11.22, p < 0.003, \hat{\eta}_P^2 = 0.91$. When looking at individual measures, the difference in subjective arousal was significant at $F(1,12) = 4.57, p < 0.05, \hat{\eta}_P^2 = 0.28$, with the GII setup having a smaller mean of $4.6(SE = 0.3)$ as compared to $5.03(SE = 0.26)$ for the LII setup. The EEG-based arousal estimation also showed a similar trend, with greater changes in the LII setup, with a mean of $0.32(SE = 0.22)$, as compared to the GII setup with mean of $-0.21(SE = 0.16)$. In this regard, a trend for a significant difference was observed with $F(1,12) = 11.73, p = 0.061, \hat{\eta}_P^2 = 0.26$. In addition, behavioral data also showed significant differences. For Gain data, the setups differed at $F(1,12) = 12.07, p < 0.005, \hat{\eta}_P^2 = 0.5$ with means of $0.77(SE = 0.03)$ for GII vs. $0.56(SE = 0.05)$ for LII. A similar pattern emerged for Homogeneity data, where significance was at $F(1,12) = 24.81, p < 0.001, \hat{\eta}_P^2 = 0.67$ with the means of $0.92(SE = 0.01)$ for GII vs. $0.76(SE = 0.03)$ for LII.

6.4.3. Differences between tasks

The overall effect of the task was significant at $\Lambda = 0.09, F(6,7) = 12.05, p < 0.002, \hat{\eta}_P^2 = 0.91$. By looking at each measure, it can be seen that the difference between HAHV and LALV tasks was mainly due to subjective V/A ratings. For arousal it was at $F(1,12) = 80.83, p < 0.001, \hat{\eta}_P^2 = 0.87$ with the means of $3.33(SE = 0.26)$ for LALV vs. $6.3(SE = 0.38)$ for the HAHV tasks. For valence it was at $F(1,12) = 10.62, p < 0.01, \hat{\eta}_P^2 = 0.47$ with the means of $5.23(SE = 0.19)$ for LALV vs. $6.27(SE500 = 0.24)$ for the HAHV tasks. These results confirm the separation of tasks in the Valence-Arousal space from a subjective perspective. Gain data also showed a trend to significance with $F(1,12) = 2.8, p = 0.12, \hat{\eta}_P^2 = 0.19$, with means of $0.69(SE = 0.03)$ for the LALV vs. $0.64(SE = 0.04)$ for the HAHV task.

6.4.4. Interaction between setup and task factors

As shown in Figure 6.7 (right panel), from all measures only EEG-based valence estimation showed a significant interaction between setup and task at $F(1,12) = 6.16, p < 0.05, \hat{\eta}_P^2 = 0.34$. Whereas for the GII setup there was no difference between the LALV and HAHV tasks (0.04 vs. 0.04), a difference was found for LII (-0.11 vs. 0.22). In addition, subjective arousal showed a trend for a significant effect at $F(1,12) = 3.67, p = 0.08, \hat{\eta}_P^2 = 0.23$. In consonance with EEG-based valence, the separation between the LALV and HAHV tasks were higher for LII (3.33 vs 6.73) than for GII setup (3.33 vs. 5.87) (see Figure 6.7, left panel).

6.4.5. Interaction between setup, task and music experience

The triple interaction between setup (GII vs. LII), task (HAHV vs. LALV) and group (N, K and E) was significant for subjective valence ratings and EEG-based valence. In the case of the former, this interaction was significant at $F(2,12) = 6.26, p < 0.05, \hat{\eta}_P^2 = 0.51$. Figure 6.8 shows the nature of this interaction, where novices (N) showed a clear difference between HAHV and LALV tasks in both setups (GII and LII). The knowledgeable group (K) showed this pattern only for local implicit interaction (LII), and the experts group (E) did not show the same level of distinction between tasks.

Similarly, for EEG-based valence the interaction was significant at $F(2,12) = 5.93, p < 0.05, \hat{\eta}_P^2 = 0.5$, showing that the second group (Knowledgeable) demonstrated the strongest difference between the HAHV and LALV tasks in the LII setup (see Figure 6.9).

6.4.6. Analysis of System and Performance aspects of the DMI

Statistical analysis showed no significant differences for System and Performance variables. Table 6.2 describes the means of both GII and LII setups by music knowledge group. A Tangible setup (tangible interaction only) was added as a baseline for Novice and Knowledgeable groups.

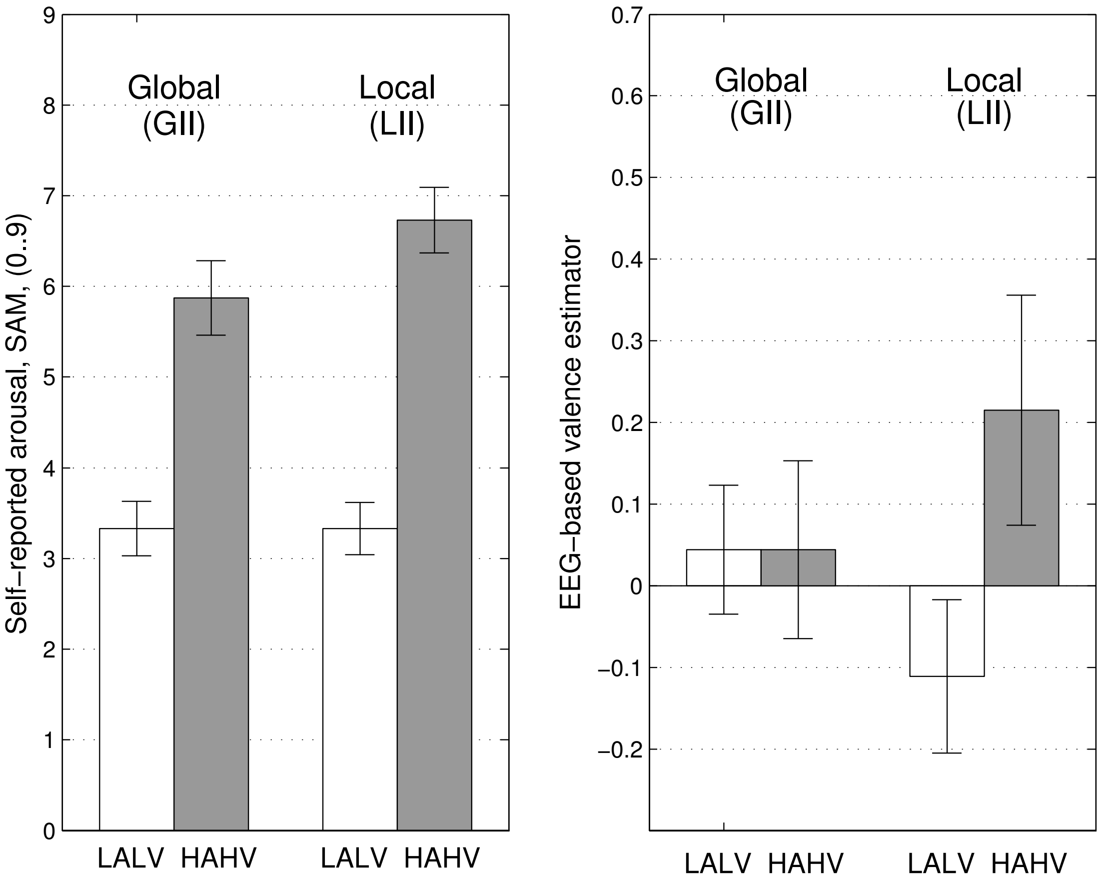


Figure 6.7.: Interactions between setup and task factors for subjective arousal ($p = 0.08$) and EEG-based valence ($p < 0.05$). Low arousal - low valence (LALV), and high arousal - high valence (HAHV). Error bars represent standard error values

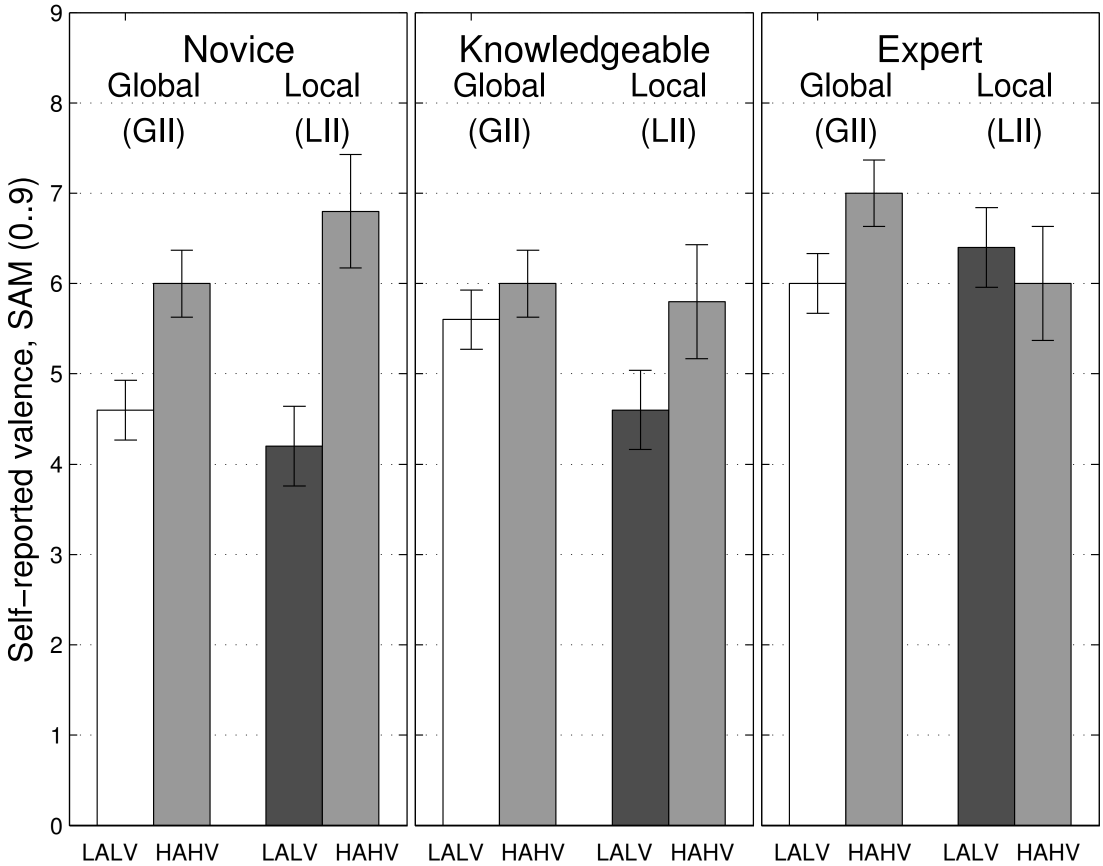


Figure 6.8.: Interaction between setup (GII vs. LII), task (LALV vs. HAHV) and group (Novice, Knowledgeable and Experts) factors for subjective valence ratings ($p < 0.02$). Error bars represent standard error values

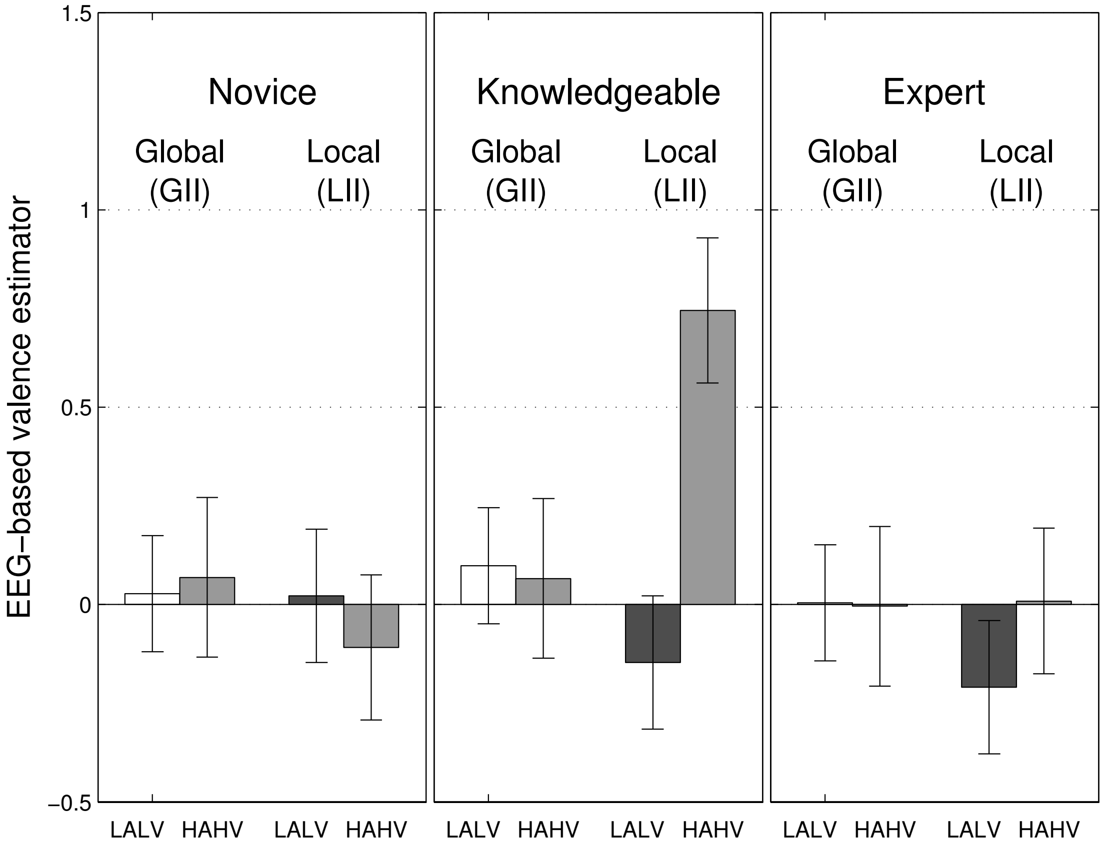


Figure 6.9.: Interaction between setup (GII vs. LII), task (LALV vs. HAHV) and group (Novice, Knowledgeable and Experts) factors for EEG-based valence ($p < 0.02$). Error bars represent standard error values.

Setup	System	Performance
Tangible	$N = 3.60(SE = 0.33)$ $K = 3.7(SE = 0.23)$ $E = -$	$N = 3.00(SE = 0.25)$ $K = 3.4(SE = 0.25)$ $E = -$
Global	$N = 3.60(SE = 0.23)$ $K = 2.95(SE = 0.23)$ $E = 3.10(SE = 0.23)$	$N = 3.20(SE = 0.25)$ $K = 3.00(SE = 0.25)$ $E = 3.10(SE = 0.25)$
Local	$N = 3.90(SE = 0.32)$ $K = 3.50(SE = 0.32)$ $E = 2.80(SE = 0.32)$	$N = 3.10(SE = 0.29)$ $K = 3.60(SE = 0.29)$ $E = 2.80(SE = 0.29)$

Table 6.2.: System and Performance aspects of *b-Reactable* Global and Local setups for the three experimental groups (N= novice; K= knowledgeable; E= Reactable experts). A Tangible condition is included as a baseline for N and K participants.

6.4.7. Insights from user experience interviews

Many Novice and Knowledgeable participants (i.e. with scarce or no previous experience playing Reactable) declared that learning to use the *physiopucks* and understanding their working mechanisms was not more difficult than coping with the standard Reactable itself:

> "After the initial explanation and testing phase, it was easy for me to understand how the *physiopucks* work." (P14, Male, Novice).
>
> "I played with the Reactable before and I found the *physiopucks* working in a very similar way to the rest of the objects. Their labels and shapes were also useful for understanding the difference between general (Global) and Local controllers" (P11, Male, Knowledgeable).

Some Knowledgeable participants found the Local implicit interaction (LII) easier to learn and use compared to the Global implicit interaction (GII):

> "You catch the Local controller very fast, as it affects specific objects on the table. It is more straightforward" (P8, Male, Knowledgeable).
>
> "The Global object spreads through the whole surface and it takes some time to understand how it affects all that is happening in the interface at that moment. This does not happen with the Local object." (P11, Male, Knowledgeable).

In the same manner, many participants showed preference for LII over GII:

> "Global control had less possibilities than Local control. Local is better integrated into the Reactable, and it is less restrictive" (P13, Male, Knowledgeable).
>
> "In Global control music events were happening in an unexpected way (...) in Local control the effect was smooth because it controlled only one object" (P4, Male, Knowledgeable).
>
> "It was easier with Local (implicit interaction) because I could better decide what to control with the EEG (...) With the Global it was a bit tricky to have it working in the way I expected" (P10, Female, Novice).

However, some Expert participants stressed on the expressive possibilities of GII or its combination with LII:

> "If you have more or less an idea of what you want to play, I think Global works better. You just have to dispose the right elements (pucks) on the table, and then the Global (*physiopuck*) works as a steering wheel (...) With

> Local you have to pay more attention and adjust it specifically for one object" (P3, Male, Expert).
>
> "I think the Global mode is more powerful if what you want to do is to use your affective states as an instrument, but I would like to have a mixed controlled system with Local and Global" (P2, Male, Expert).

In terms of multimodality, LII also stood out as the preferred method for combining gestural and implicit physiological interaction:

> "I never played Reactable before, so I wanted to use it all. For that, I found Local more friendly" (P14, Male, Novice).
>
> "Combining hand operations and brain (affective) states was easier in the Local mode. Also you can use the Local (*physiopuck*) as both an EEG input and a manual controller by changing the gain in a specific direction" (P11, Male, Knowledgeable)
>
> "(Local Control) works best if you want to take full advantage of Reactable, meaning using your hands" (P15, Male, Expert).

Comments about feedback perception were mixed. Whereas some participants suggested that the *b-Reactable* "could also work for neurofeedback training" (P3, Male, Expert), or described physiology-driven musical processors as "responding in the direction they were expecting" (P12, Male, Knowledgeable), others did not perceived the feedback in a straightforward manner:

> "It was difficult to see the relation between my affective responses and the music (...) I couldn't plan my performance based *only* on EEG" (P1, Male, Expert).
>
> "I felt the need of stronger changes in the music when using brain objects (*physiopucks*). That would help to better perceive their effect during performance" (P2, Male, Expert).
>
> "The mapping design was right, but it was a bit slow. To have a fast response is important in music performance, otherwise you tend to perceive the brain signals as random" (P8, Male, Knowledgeable).

Interestingly, players did not highlighted differences between GII and LII in terms of feedback perception. In terms of ergonomics and wearability, one player stressed on the intrusiveness of the EEG equipment:

> "The system is responsive, but not for performing on stage. The equipment is uncomfortable" (P10, Female, Novice).

Finally, players suggested a number of system improvements such as a GUI for mapping personalization (”I would include a graphic interface in order to personalize the affective mapping”, P4, Male, Knowledgeable.”), and the addition of multiple LIIs *physiopucks* (”It would be nice to have more Local brain pucks in order to apply them to different Reactable objects”, P2, Male, 610 Expert).

6.5. Discussion

According to our initial hypotheses, the analysis shows five main results:

1. The proposed implicit sonic interactions were distinctly perceived by participants, and used differently during music performance.
2. Participants applied these sonic interactions (both GII and LII) as a meaningful component of music performance.
3. Participant’s perception and preference on implicit sonic interaction varied according to their previous music knowledge.
4. Local implicit interaction (LLI) achieved better results in terms of physiological and subjective perception, and favored multimodality compared to GII.

In the following subsections we discuss these aspects in depth.

6.5.1. Performance of affective estimations and physiology-based implicit interaction setups

One of our main design goals was to achieve a robust real time affective estimations based on EEG. This objective presented a number of technical challenges such as avoiding artifacts caused by the participant’s movement and muscle activity. We aimed at controlling these factors by utilizing professional hardware for EEG monitoring, using relative power estimates, and by restricting user movements around the *b-Reactable*.

The results presented in Section 6.4.1 show that the proposed affective estimations were reliable in a music performance context. The statistical analysis show a significant decrease of arousal and increase of valence after the relaxation induction (which was performed by the participant herself using the *b-Reactable*). Moreover, the between task analysis show that, regardless of the implicit interaction setup (GII or LII), participants were able to perform improvisations in the required V/A target (HAHV or LALV). These findings support our real-time affective estimations and task design. However, we also have to consider that the duration of each task was rather short (5 minutes) and

HAHV and LALV conditions did not represent extreme affective states, which resulted in smaller differences in the EEG recordings.

6.5.2. Meaningful integration of implicit PhyComp in music performance

As the direct assessment of System and Performance aspects did not bring significant effects (see Table 6.2) we have to be cautious on claiming that our implicit sonic interactions were effectively perceived as an *expressive* component of the participants' performance. However, our study shows that all participants understood the operation of the sonic implicit interactions, as they were able to perform music improvisations in the requested affective direction, with both GII and LII setups. In fact, the latter reached better outcomes with greater V/A separations between HAHV and LALV tasks than GII. Together with the reliability of the affective estimations and task design, this shows a better perception of the sonic feedback coming from LII, supported by the analysis of subjective arousal and EEG-based valence, as described in the Subsection 6.4.4. The Performance assessment (although not significant) suggests that LII might be preferred in terms of expressive performance by Novice and Knowledgeable participants. This trend is also consistent with the open interviews (Section 6.4.7) where these kind of participants appreciated the way LII was integrated into the *b-Reactable*, and its potential for being combined with other input methods (e.g. gesture input).

Considering the above mentioned findings, further investigation is needed to specifically determine how physiology-driven implicit sonic interactions enhance music expression. In order to do so, this study could be expanded with a bigger sample, applying longitudinally tests through repeated sessions, and even use other NIME evaluation frameworks like O'Modhrain [2011]. However, the results of this experiment already suggest that the combination of implicit PhyComp with different input methods (i.e. tangible objects and touch) in the musical domain is a promising option for improving multimodality and unsupervised control. Similarly, we can envision these types of indirect interactions complementing other adaptive systems (e.g. gaming or learning applications) that modify their content and interface in real time according to the user's affective states without the need of special training. This approach may have a significant impact in user's learning, experience and exploration of data.

6.5.3. The role of music knowledge in implicit interaction preference

The study also shows a predominant role of previous knowledge in the perception and utilization of physiology-based implicit sonic interaction. Knowledgeable players were the ones reporting higher V/A effects, precisely by means of LII. This is supported by the analysis of EEG-based valence estimations, and also consistent with the post-experimental interviews, where LII was described as “less restrictive”, “better integrated”, “smoother” and “easier to work with”. On the other hand, participants with deeper Reactable expertise tend to favor GII over LII. This is shown in the comments gathered after the experiment, and a similar trend can be seen in the System metrics, (Table 6.2) although they are not significant. This could be explained by the fact that Experts knew well how to operate the Reactable, and were therefore in a better position for arranging music operations that would better fit global implicit interaction. As suggested by several participants, GII required some previous idea of the composition to be performed, in order to reach satisfactory results.

6.5.4. Multimodality and personalization

Together with robust physiological mappings and meaningful interaction, another goal of our study was to assess the possibilities of Global and Local implicit interaction for multimodal control. In this regard, behavioral data resulting from the use of *Physiopucks’s* Gain show that participants used GII and LII in significant different ways.

The use of GII demanded significantly higher Gain than LII (see Section 6.4.2). The homogeneity of GII was also significantly higher than the LII setup. This means that, while using GII, participants prioritized implicit physiological control over gesture input. By doing this, musical processors (filters, delay, distortion and modulators) were driven by the user affective responses, leaving sound generation as the sole gestural operation. As shown by the open interviews, this has been perceived as a restriction and a control impairment by Novice and Knowledgeable users. Experts, on the other hand, did not consider it as a problem, but rather as a global feature that needed to be initialized and disposed in a specific manner to reach a desired musical outcome. Of course, it required understanding and skills on the use of the Reactable, something that Novice and Knowledgeable users did not have.

LII, on the other hand, was operated in a very different manner. It required significantly lower Gain values, with greater customization than GII (i.e. less homogeneity). Through this approach, participants also favored multimodal interaction, as lower Gain accounts for a combination of both physiological implicit and gestural direct control. It is not surprising therefore that LII was preferred by Novice and Knowledgeable users, who saw it as less restrictive, smoother and better integrated to the Reactable framework. This preference is also supported by higher measures of V/A for LII. Finally it is important to note that while all the players had the option of not using implicit sonic interactions by setting the *physiopuck's* Gain to zero, none of the participants did so.

6.5.5. Limitations and future work

The study presented in this Chapter also offers useful insights for guiding future work on physiology-based implicit sonic interaction. Participants' perception was critical for evaluating sonic designs within an expressive context such as NIME. Although participants agreed on the convenience of the proposed implicit control mappings, some were not entirely satisfied with the manner the feedback was presented to them. Concretely, Expert users requested stronger and faster changes in the sound processors when applying physiological control. The challenge of providing the expected sonic feedback for each player could be addressed by improving the current user personalization mechanisms, like adding a graphic user interface (GUI) on the tabletop surface to customize the V/A-to-Sound Parameters interpolation. As mentioned during the open interviews, this feature could significantly improve user experience.

As discussed before, the analysis of System and Performance aspects of the *b-Reactable* did not show significant effects. This is a relevant aspect to be tackled if we aim to determine the contribution of implicit PhyComp to musical expression. One possible explanation for the lack of significance in this regard could be the manner in which we applied the framework, that differed from the first framework deployment, presented in Chapter 5. In that case, we validated the framework from the *listener* perspective. But in the experiment presented in this Chapter we applied the framework to assess the perception of *performer*s. It could be possible that our framework does not reflect the perception of different stakeholders in the way we expected. To further explore this scenario we could have *listeners* evaluating the performances done with the *b-Reactable,* following a given valence/arousal target. Such approach would also allow to assess the congruency between *performer* and *listener* perception of musical improvisations. On the other hand, it is important to consider that the System and Performance assessment, although not significant, was coherent with the insights gathered form participants during the open ended interviews. This could mean that the lack of significant effects is due to the sample size, or the lack of longitudinal studies. In any case, other evaluation frameworks could be applied to verify the consistency of our current results.

Another aspect that requires further improvement is sensor wearability. Although the technical setup used for this study accounts for robust signal acquisition, it is not yet convenient for out-of-the-lab performances. Preparation time is long, moreover participants pointed out the inconvenience of wearing a wired EEG cap for long periods of time. Less invasive and wireless hardware, such as the Starlab Enobio or Emotiv EPOC, could be tested in this regard.

Our physiology-driven implicit sonic interactions could be also expanded beyond EEG. The incorporation of other biopotentials such as EMG, ECG or respiration could be used to estimate other psychophysiological features (e.g. stress), for strengthen our current affective estimations with multimodal physiological data, or for promoting both explicit and implicit physiological control, as it has been done for video games [Nacke et al., 2011]. By following this approach, a future version of the *b-Reactable* could integrate multiple Local and Global *physiopucks* responding to different psychophysiological estimations. Self-regulation and implicit learning could also be further explored by using longer and repeated sessions. In this manner, we could determine whether participants *learn* to control implicit sonic interactions better, in a musical performance context.

As in the case of our first study with *b-Reactable* (Chapter 3), we could also explore collaborative and multiuser music performance. In this regard, we envision a setup with two *Emitters* using the implicit sonic interactions proposed in this Chapter. This approach will allow to study aspects such as multi-user adaptation and synchronization.

Finally, it is evident that several aspects of this work are specific to the musical domain. Whereas music appears as a good candidate for exploring the contributions of implicit interaction in an expressive context beyond goal oriented tasks, it also imposes specific constrains in the way control mappings are defined (i.e. controlling sound processors in the way a modular synthesizer would do) and in users' previous knowledge. Therefore, while our study shows significant differences between Global and Local implicit interaction in music performance, this could be further explored in other domains, ranging from gaming to learning and big data exploration.

6.6. Conclusion & next steps

In this Chapter we have addressed the issue of *meaningfulness* that, together with *perceptualization* and *mapping complexity*, constitutes one of the main aspects of physiology-based implicit sonic interaction analyzed in this dissertation. A meaningful integration in a NIME context implies the perception of implicit PhyComp as an expressive component of the DMI, through which the player can produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to.

In order to tackle this issue, we have created a new version of the *b-Reactable*, which incorporated a number of features informed by our previous experiments, namely a parameter mapping approach, end-user personalization of *physiopucks*, a *complex* implicit physiological input (a two-dimensional representation of user affective responses estimated through EEG, and two different implicit interaction setups (Global and Local).

We tested this new incarnation of the *b-Reactable* in an expressive context (i.e. music performance) involving 15 participants with different levels of music experience (Novice, Knowledgeable and Expert) who performed musical improvisation exercises under two conditions, defined by the aforementioned implicit interaction setups (Global and Local). Four different measures were collected for evaluating user experience and *meaningfulness*: affective data (V/A), behavioral data (based on the use of the *physiopucks*), system and performance aspects of the DMI, and open interviews on user experience.

The main results showed that our affective estimations were valid for the context of music performance, and that participants used these sonic interactions (both Global and Local) in a comprehensive and meaningful manner. However, System and Performance aspects of implicit PhyComp in the realm of DMI have to be further explored by analyzing bigger samples, applying longitudinal studies, and other evaluation frameworks. Subjective, behavioral and physiological data showed that Global and Local implicit interaction were perceived in significantly distinctive ways according to participants' previous musical experience, with preference for the latter.

With this Chapter we conclude our investigation on sonic interaction design for implicit PhyComp within the NIME domain. In order to explore other possible uses of this technology beyond music, in the next Chapter we will apply implicit PhyComp in personal fabrication processes, presenting a do-it-yourself knitting system based on EEG affective estimations, called *NeuroKnitting*.

7. Bonus Track. Implicit Physiological Computing Beyond Sonic Interaction Design

This chapter aims to explore implicit Physiological Computing beyond sonic interaction design and the musical domain. To achieve this goal, we tackle the field of personal fabrication and present *NeuroKnitting*, a system that can be used to create knitted garments according to the users' affective responses estimated from EEG. We deploy this system in two recording sessions, from which we extract preliminary insights and design guidelines. The tests show that *Neuroknitting* can be used for embodying implicit psychophysiological data into unique, customized physical objects. As every human being reacts differently to a given experience, the knitted patterns change according to the user and her context. *NeuroKnitting* thus opens the door to further structured analyses on aspects such as the perception of implicitly generated fabrication patterns, and the use of different stimuli to trigger meaningful user experiences during the fabrication process.

7.1. Introduction

In this Chapter we move away from sonic interaction design (SID) to explore a different use case for implicit physiological computing (PhyComp). We consider this Chapter a *bonus track,* as it does not pursue specific research problems, but rather engages PhyComp technology in a different context with the goal of identifying design guidelines that go beyond the musical domain. In order to do so, we focus on personal fabrication, a field that explores how information relates its physical properties by creating fully functioning systems that include sensing, logic, actuation, and displays. As expressed by Gershenfeld [2008]: "the way the world has evolved, hardware has been separated from software, and channels from their content, but many of the hardest, most challenging, and most interesting problems lie right at this interface".

In this context we develop *NeuroKnitting*, a fabrication system that can be used to create knitted garments according to the users' psychophysiological responses estimated from EEG measures. We present two versions of *NeuroKnitting*, based on different fabrication (i.e. knitting) strategies. The Chapter is structured as follows. We first present our motivation for approaching the field of personal fabrication, together with relevant background information. We then introduce the *NeuroKnitting* system, including all its technical components (i.e. EEG signal acquisition, processing, fabrication pattern definition, and knitting process). This is followed by two deployments for testing the system (patterns through Bach's Goldberg variations, and demos at the Maker Faire Rome). Finally, we discuss the main insights gathered in each deployment.

7.2. Motivation

The recent years have shown significant advances in techniques for do-it-yourself (DIY) digital fabrication such as 3D printing and laser cutting, promoting low-cost, high-speed and high-quality physical production and prototyping. This scenario, together with the expansion of open-hardware platforms such as Arduino[1] and the Raspberry Pi[2], also provoked an increasing interest among different research fields, such as HCI [Willis et al., 2011] and end-user communities as the *Maker* movement, that can now access to a broad set of tools and techniques to create proof of concepts without recurring to specialized and expensive equipment.

As personal fabrication gets more popular and spread, a wide new range of services for designing and manufacturing is emerging. This scenario accounts for new possibilities of making and disseminating, where the phenomena that Lipson and Kurman called "a factory at home" and "one-person industries" is not anymore a vision of the future but a fact [Lipson and Kurman, 2010].

Although personal fabrication has empowered end-users and non-expert communities -encouraging activities around the DIY mindset as in the case of the *Maker Faire*[3]- the production of new content and new fabrication methods do not seem to advance at the same pace. The open nature and affordability of several fabrication tools provides an exciting opportunity to explore new ways to produce home-brew, personal content and to interact with different sources of data. However, most of the creations within these communities are still based on the replication of pre-existent objects, or on the adaptation of methods and patterns previously developed in other fields, such as architecture and

[1] https://www.arduino.cc/ (accessed on October, 2015).
[2] https://www.raspberrypi.org/ (accessed on October, 2015).
[3] http://makerfaire.com/ (accessed on October, 2015).

industrial design. This trend can be clearly observed in the online digital database Thingiverse[4], where most of the user experience is based on sharing and replicating fabrication models, with very little space for customization[5].

In the quest for finding new methods to create alternative fabrication patterns, the field of PhyComp stands out. As described in Chapter 2, the recent advance in human physiological sensing and signal processing techniques have fostered the development of non-invasive systems that allow to measure and interpret brain activity, heart rate and skin conductance, among others, for estimating different psychophysiological states. As in the case of personal fabrication technology, an increasing number of consumer of-the-shelf devices and open source platforms have brought PhyComp out of specialized laboratories and closer to real life conditions. Systems like the Emotiv EPOC[6] and NeuroSky's Mindwave[7], together with open hardware initiatives such as the BITalino[8] are currently bridging the gap between PhyComp and the *maker* community.

In this context we present *NeuroKnitting*, a modified knitting machine, created together with the artists Varvara Guljajeva and Mar Canet, for generating garments according to the users' psychophysiological response to a given experience. In *NeuroKnitting,* PhyComp technology is at the core of the fabrication process. It presents methods to acquire EEG signals in real time and to estimate affective states that are later used for creating personalized garments. Through this approach, we tackle the problem of generating novel and personalized fabrication patterns, creating unique pieces capable of embodying the physiological traits of human experience.

7.3. Background

Although PhyComp offers the unique opportunity of accessing to our affective, perceptive and cognitive states in real time, there are very few initiatives that bring together human physiology and personal fabrication. McGrath and colleagues created the *NeuroMarker* [2012], an artwork that playfully implements the concept of translating designer's ideas into a product. Visitors are invited to use their own raw EEG to fabricate personalized physical objects. However, the *NeuroMarker* do not aim at extracting high level

[4] `https://www.thingiverse.com/` (accessed on October, 2015).

[5] In this regard, personal fabrication platforms have recently put more effort in enabling end-user customization tools, as in the case of Thingiverse's *Customizer,* which lets anyone personalize 3D printable designs created with OpenSCAD.

[6] `https://emotiv.com/epoc.php` (accessed on October, 2015).

[7] `http://store.neurosky.com/products/mindwave-1` (accessed on October, 2015)

[8] `http://www.bitalino.com/` (accessed on October, 2015).

representations of brain activity, thus no link between EEG and psychophysiological states can be traced. More recently, the company Thinker Thing presented *Objects from Thought*[9], a project that interfaces EEG equipment with 3D printing machines using a set of images as emotional stimuli. Although this project utilizes EEG to identify emotional responses, the algorithms for signal processing and mapping are not described in the current documentation.

A remarkable case of the intersection between personal fabrication and PhyComp is Stephen Barrass' *Hypertension*, a project where a singing bowl is shaped from a year of blood pressure readings, producing an individual and unique object [2015]. The bowl, digitally fabricated in stainless steel, is also an acoustic object that produces a tone when rubbed with a wooden *puja* stick. In this way, the object is unique both in terms of shape and in the sounds it generates, directly related to the person from whom the readings were taken. Although *Hypertension* does not make use of real time physiological data, it represents an interesting case of physiology-based fabrication using meaningful data.

7.4. The NeuroKnitting prototype

NeuroKnitting was built with the aim of placing PhyComp at the core of a personal fabrication process. *NeuroKnitting* measures psychophysiological responses to a given stimuli by means of EEG, and produces a knitting pattern for the creation of garments. The system is composed of three main modules, as showed in Figure 7.1: (i) signal acquisition, (ii) signal processing, and (iii) fabrication. The latter includes a pattern generation software and the *Knitic* open-hardware framework for controlling a knitting machine via Arduino.

7.4.1. EEG signal acquisition and toolbox for processing

Leveraging on the methods and techniques applied and validated in Chapter 4, EEG data is acquired using the Emotiv EPOC, a wireless, non-invasive 14-channel EEG headset. The EPOC's electrodes require a saline solution to improve conductivity, and their placement is based on the 10-20 system [Homan et al., 1987]. Data acquired with this device is converted to a digital signal at a sampling rate of 128 Hz. The device also applies internal low and high pass filtering at $85Hz$ and $0.16Hz$ respectively. This data is accessed through the Emotiv EPOC SDK and then fed into Matlab/Simulink[10] for

[9] `http://thinkerthing.com/` (accessed on October, 2015)
[10] `http://es.mathworks.com/products/simulink/` (accessed on October, 2015).

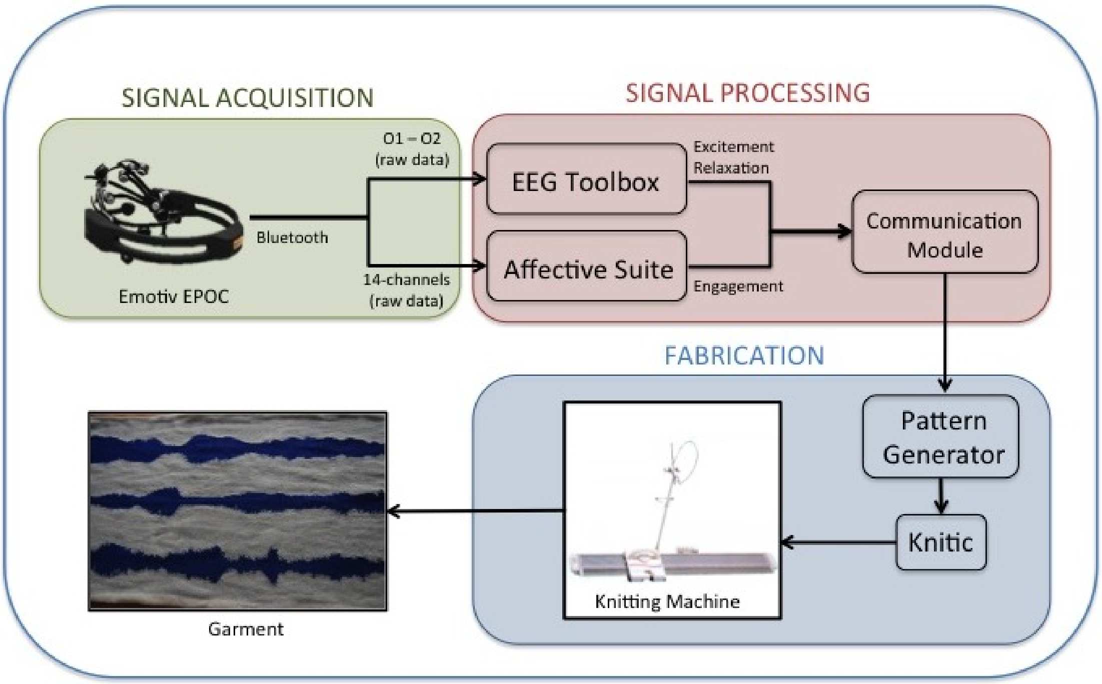


Figure 7.1.: The *NeuroKnitting* system architecture. A signal acquisition module (upper left section) allows measuring brain activity from a EEG headset, delivering raw data for processing. The signal processing module (upper left section) is composed by a toolbox and the EPOC's Affective Suite for extracting high level EEG features. The fabrication module (lower right section) includes pattern generator software and the *Knitic* framework that interfaces the system with a Brother KH-930e knitting machine.

further processing. We use the *EEG-TOOLbox* created for the experiments of Chapter 4, as it has been designed for treating the Emotiv semi-raw data and extracting different EEG features, both online and offline. Its modular approach allows extracting low level EEG features through a number of processing blocks, adapting to different monitoring strategies, as in the case of *NeuroKnitting*. We therefore use the following blocks of the toolbox:

- *Emotiv2Simulink*: based on a Mex S-function and on the drivers provided by the manufacturer, it allows the access to the semi-raw data for all 14 electrodes in real time, plus a built-in gyroscope, time stamp, markers and sync signals.
- *BP filters*: Butterworth IIR discrete band-pass filters that can be tuned to any given custom frequency. Envelope is a block that squares the input signal and then applies an FIR low-pass filter and down- sampling to estimate its envelope. The filter parameters depend on the particular EEG feature to be computed.
- *Envelope*: a block that squares the input signal and then applies a FIR low-pass filter and down-sampling to estimate its envelope.
- *OSCSend*: allows direct communication via UDP through the Open Sound Control protocol (OSC).

As in Chapter 4, the participant's relaxation state is estimated by looking at the EEG alpha and theta bands, and the theta-alpha ratio (t/a). In this regard, the toolbox is configured to extract three main low-level EEG features, according the *a/t* neurofeedback rationale [Gruzelier, 2009]:

- *Alpha relative power*: calculated from the occipital area (electrodes O1 and O2), where this type of activity occurs during close-eye conditions, applying a band-pass filter with a frequency range of $8Hz$ to $13Hz$.
- *Theta relative power*: calculated from the activity of all 14 channels, as cortical theta rhythms are small and diffuse when picked up by scalp electrode, arising almost entirely from the cerebral cortex [Kropotov, 2010], applying a band-pass filter with a range of $4Hz$ to $7.5Hz$.
- *t/a power ratio:* estimated as the main measure of relaxation. A spatial filtering is applied to give more weight to the occipital channels in the calculation of the envelope of the t/a ratio. The envelope is estimated using the envelope block based on FIR-based filter (order 35).

The relative power is obtained by dividing the power of each band by the overall power of the same signal. In this way, the output signal is kept within a range of 0 to 1. A second psychophysiological feature, engagement, is extracted using the Emotiv Affective

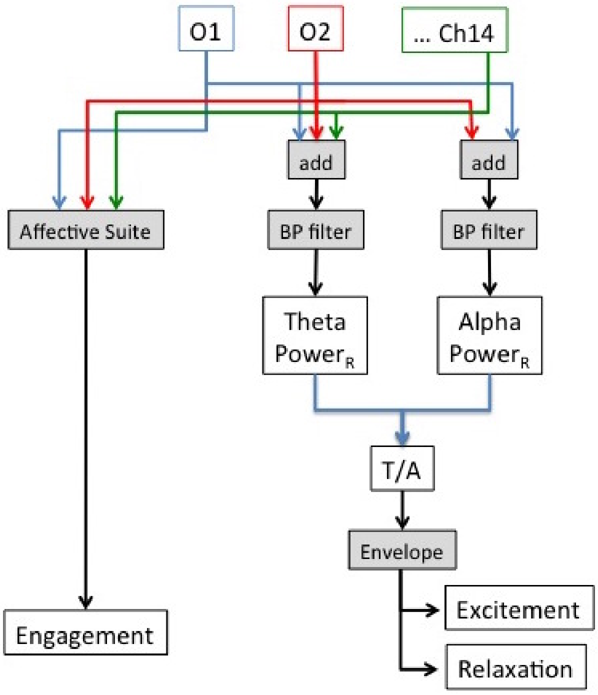


Figure 7.2.: Processing stages of EEG features for the *NeuroKnitting* prototype.

Suite. Figure 7.2 shows the processing stages for both relaxation and engagement.

Each feature is streamed via OSC to a communication module made in Pure Data (Pd) (see Figure 7.1) where they are recorded in a CSV file at a rate of $1Hz$. This downsampling was done calculating the mean of all values within a bin of the same size, in order to achieve a coherent mapping with the knitting pattern (i.e. each stitch made by the knitting machine is around 50mm, so greater temporal resolution will translate to bigger weaves). Therefore, this resampling can be adjusted according to the type of garment to be knitted.

7.4.2. Pattern generator

A pattern generation software was developed using Processing[11], an open source programming language for fast prototyping. The pattern generator automatically reads up the CSV file created by the communication module and displays it by means of three bars with variable width, corresponding to (i) relaxation, (ii) excitement and (iii) engagement activity. The software then creates a PNG file out of this render, where each

[11]https://processing.org/ (accessed on October, 2015).

bar has a maximum width of 40 pixels (i.e. stitches). Width is intentionally limited in order to avoid broad weaves, but this parameters can be easily changed according to the garment that will be created. Given the limitation of most domestic knitting machines for handling multiple yarn streams (they normally support up to two reels) our fabrication pattern is based on two colors (one for the bars and other for the background). The full pattern is therefore 150 stitches width and its length varies depending on the duration of the EEG recording.

The Knitic framework

Once the knitting pattern is created from the aforementioned EEG features, it is sent to *Knitic*[12], an open hardware controller based on Arduino and Processing. This controller was originally developed for the 1980's Brother electronic knitting machines, as an attempt to bring together textile and contemporary digital manufacturing. The electronic knitting machine was arguably the first fabrication tool at home. However, it has not been further developed by the personal fabrication or DIY communities. Hence, *Knitic* aims to update this obsolete technology for the current creation needs. In order to do so, *Knitic* provides hardware and software tools to update the machine pattern-uploading methods. This is a primary technical need, as domestic electronic knitting machines were discontinued in the late 1990s, and the only available method for uploading patterns is via floppy disks or scanned sheets, which are not efficient options for garment production and pattern sharing nowadays.

For overcoming the above mentioned limitations, *Knitic* replaces the original hardware controller of the knitting machine with an Arduino PCB, as showed in Figure 7.3. The PCB reads the inputs of end-of-line sensors and encoders that identify the precise position of the carriage in the machine. The hardware also controls the solenoids that command the movements of each needle based on the input pattern. Once the knitting machine is modified, new knitting patterns can be sent to it using the *Knitic* software, which in its current version utilizes a bitmap file (i.e. PNG) as the source of the pattern. The *Knitic* software provides a front end interface that allows to visualize the pattern in use, and to monitor the operation of the machine (i.e. the position of the carriage, knitting direction, dimensions of the weave, and the rows and stitches already knitted). A black pixel stands for a contrast yarn (needle position D) and a white pixel indicates background yarn (the needle position B).[13]. Currently the *Knitic* project has evolved,

[12] `http://www.knitic.com/` (accessed on October, 2015).

[13] A step-by-step tutorial on how Knitic works at hardware and software levels (including a video) can be found at `http://www.knitic.com/tutorials/` (accessed on October, 2015).

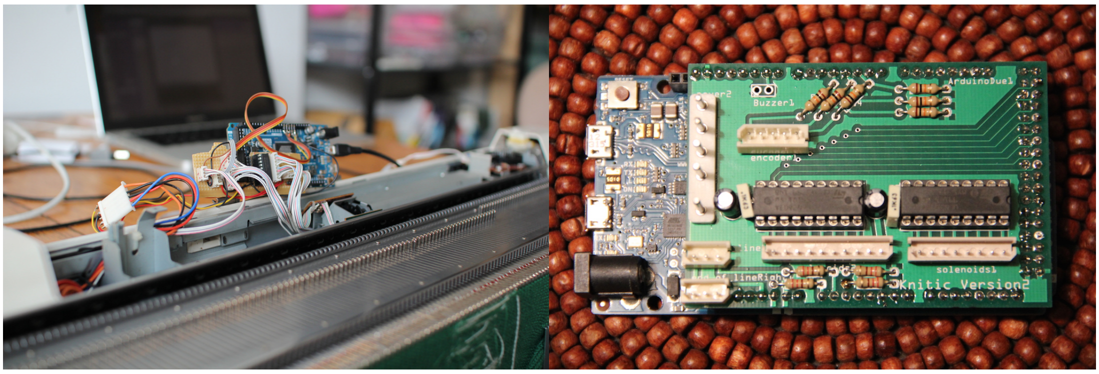


Figure 7.3.: The *Knitic* PCB for Arduino is designed to replace the knitting machine's old brain without any harm. It means the original system can be use anytime if wished.

allowing not only to *recycle* discontinued knitting machines, but also to create open hardware replicable circular knitting machines from scratch, using open hardware and 3D printed pieces.

7.4.3. Deployment 1: patterns through Bach's Goldberg variations

The first test of *NeuroKnitting* in real world conditions was done using music as affective stimulus. We chose music as it is a well-known mood inducer that has been widely used for emotional induction and as an affective reward in different neurofeedback therapies [Egner and Gruzelier, 2003][14]. Bach's Goldberg Variations (concretely the aria and its first seven variations, with length of 10 minutes) were used as musical stimuli. Two subjects, aged 39 and 30 years old, both male, participated in the test, according to the following procedure: (i) EEG headset placement (5 minutes); (ii) calibration phase to assure good conductivity (5 minutes); (iii) musical stimulation phase under closed-eyes condition. EEG activity was collected and sent to the pattern generation system. (10 minutes); (iv) headset removal (3 minutes length).

The generated pattern was used to knit a scarf, as showed in Figure 7.4. We have chosen this garment as it accounts for a straightforward representation of the evolution of EEG activity in the temporal domain. Each stitch of the knitting machine equals 1 second of the stimulus and the EEG recording[15].

[14] Please see Section 6.2.2 for a review on this topic.

[15] A video showing this setup is available at `https://vimeo.com/67714066` (accessed on October, 2015).

Figure 7.4.: Scarves knitted during deployment 1.

Insights from deployment 1

This deployment offered a first insight on the possibilities of implicit PhyComp for personal fabrication . On the one hand, music showed to be a good candidate to stimulate users. In contrast to visual stimuli, sound allowed well-controlled EEG recordings under close-eyes conditions, a fact that reduces significantly the artifacts coming from muscular movement and sensor displacement. The scarves showed to be a good garment to start with. The generated patterns were simple and easy to adapt to the shape of the scarf, but waveform-alike plots were not perceptually intuitive. Participants tended to confuse them with audio waveforms, as they expected sound information to be embedded in the scarves as well.

Although users were able to depict their affective responses in the scarves, they needed a short explanation about the meaning of the pattern and the correspondence of each row with a given psychophysiological state. Another limitation was color. These first patterns used a single color to differentiate the rows from the body of the scarf. Participants expected color to be related to their affective states, rather than an arbitrary element to separate foreground and background . Finally, both participants expressed the desire of selecting "their favorite song" for generating a scarf and for making the fabrication process more engaging and personal.

7.4.4. Deployment 2: Maker Faire Rome

Leveraging on the insights from the first deployment, we decided to update and test *NeuroKnitting* for a second time in a public event. We showed the system at the Maker Faire Rome[16] (MFR), the biggest event around the DIY community and the maker movement in Europe. The Maker Faire celebrated 131 editions worldwide in 2014, and reached over 1.5 million attendees globally since it initial launch in San Mateo, California in 2006.

In MFR we also used *NeuroKnitting* for producing scarves based on musical stimuli, but with two main differences compared to the first deployment: (1) participants were able to choose their preferred song, and (2) the fabrication process involved a new pattern strategy, updated according to the feedback received in deployment 1.

We installed *NeuroKnitting* in a booth open to the general audience. Once there, participants were able to choose a song from a music streaming application (Spotify[17]). We estimated the amplitude envelope of the selected song by including an extra feature in the *NeuroKnitting* communication module (see Figure 7.1). This new information was added to the fabrication pattern for complementing the psychophysiological data embedded in the garment (see Figure 7.5).

As in deployment 1, we used two psychophysiological features (engagement and relaxation) to generate the scarves. Engagement was represented by a dotted pattern that increased and decreased along the garment according to the level of engagement of the participant. Relaxation indices, on the other hand, were displayed through color (red for low, black for middle, and blue for high relaxation) estimating the mean of indices in bins of 30 samples (i.e. 30 seconds of music). The pattern turned blue when relaxation was below 0.3 in average, black when relaxation was between 0.3 and 0.6, and red when relaxation is higher than 0.6. We chose these colors in order to make relaxation information as intuitive as possible, at least in western cultures [Gage, 1999]. On the other hand, the envelope of the chosen music track was estimated and represented as a dotted pattern in the upper part of the scarf (as seen in Figure 7.5).

Twelve participants, 7 female, participated in the recording sessions according to the following procedure: (1) EEG headset placement (5 minutes); (2) calibration phase (5 minutes); (3) stimulation phase under closed-eyes condition (of variable duration, according to the song); (4) headset removal (3 minutes).

[16] `http://www.makerfairerome.eu/en/`(accessed on October, 2015)
[17] `www.spotify.com` (accessed on October, 2015).

Figure 7.5.: Picture of one of the scarves created with *NeuroKnitting* during the second deployment at MFR, based on Stevie Wonder's *For once in my life* (3:50 minutes). The upper section represents the envelope of the selected song. The lower part displays the user's engagement levels (width of the dotted column) and relaxation states (red, blue, and black colors) in bins of 30 seconds.

Insights from Deployment 2

The second deployment of *NeuroKnitting* was useful for testing alternative (and more complex) fabrication patterns, and for trying out the system with a bigger sample of users in a short period of time, avoiding long placing and preparation phases. It has to be said that although we knitted 12 scarves, the system was used by dozens of visitors, which did not received a garment but were able to see a digital simulation of the scarf in a laptop computer. Also, the fact of letting the participant choose their favorite song reinforced participants' attachment to the garments and involvement in the fabrication process. As deployment 1 was restricted to one specific stimulus, the role of the user in the *fabrication loop* was limited to some sort of *perceptive middleware* between the sonic input and the graphical output (i.e. pattern). This second deployment thus allowed participants to occupy a more active role in the process.

As the new patterns included both color (as a representation of relaxation states) and a dotted pattern for revealing sound information, the knitting machine had to be tinkered to handle up to 4 colors. This modification required manual operation (for changing the yarn reels) but together with the possibility of choosing multiple songs, it made *NeuroKnitting* more compelling for the general audience.

We detected, however, two main drawbacks. The first one was related to the recording conditions. Although music showed to be, again, a good affective stimulus, the recording setup was still very sensitive to environmental factors such as noise, light, crowds, etc. Participants were not able to wear headphones as they can cause electrical interference in the EEG recording, so an isolated room was needed for a correct stimulus presentation. The second problem was related to the knitting process. In average, It took up to half an hour to knit a single scarf, as the electronic knitting machine required manual operation to change yarn reels.

7.4.5. Discussion and conclusion

NeuroKnitting represents a step towards the use of PhyComp for personal fabrication beyond replication. The insights presented in this Chapter could lead to structured studies on aspects such as (i) perception of fabrication patterns (e.g. which ones better convey and represent the source data?, which ones provoke more attachment from participants? can participants distinguish their scarf from the ones generated with the psychophysiological responses of others?) (ii) the possible impact of end-user personalization (e.g. letting participants choose colors, elements within the pattern, garment types, etc.), (iii) the use of other stimuli beside music and (iv) the use of other biopotentials beyond

EEG (e.g. heart rate variability or EDA) or other psychophysiological estimations (e.g. cognitive load).

Taking into consideration our current findings, *NeuroKnitting* could be improved in different manners. We chose psychophysiological estimation based on EEG given the convenience of leveraging on the technical tools generated for this dissertation, and because EEG offers a rich physiological input from which several psychophysiological states can be deduced, ranging from emotion to perception and cognition. However, *NeuroKnitting* could be improved by applying a multimodal approach that integrates other biosignals that complement changes in the EEG. This combination, common in the field of affective computing, neuroscience and user experience evaluation, will account for stronger measures and fabrication methods beyond stationary conditions. Also, physiological sensing techniques such as ECG are good alternatives to EEG when working in ambulatory or *out of the lab* scenarios, as it requires less invasive setups without compromising signal quality.

Regarding the stimuli, other experiences beyond music could be explored, such as films, photography or even attending to a live event. In fact, we recently used *NeuroKnitting* to generate personalized scarves for football (soccer) fans, as in the case of the final match of *La Copa del Rey* between FC Barcelona and Real Madrid[18].

So far, *NeuroKnitting* requires a controlled environment for EEG signal acquisition, but as neuroheadsets get cheaper and widespread as consumer devices, users could perform their own recordings and send them along over the Internet to be processed and used digital fabrication in a 3D printing or knitting bureau.

NeuroKnitting also serves as a demonstration of how PhyComp can be used to fabricate a tangible object. The current prototype, however, does not exploit tangibility in depth. We envision other possibilities in this aspect, such as EEG responses mapped to the shape of the object/garment, or to tactile properties such as texture. These approaches might lead to more intuitive representations. For instance, soft or coarse patterns could help to differentiate users' relaxation and excitement, producing more evocative objects.

Neuroknitting is also a tool for embedding personal data into unique, customized physical objects. As every human being reacts differently to a given experience, the knitted patterns change according to the participant and her context. As mentioned before, other methods such as 3D printing can be further explored in this regard. The evocative effect of such objects also has to be taken into account. The patterns generated for *NeuroKnitting* allow us to track psychophysiological changes in time, as it would be

[18] `https://www.flickr.com/photos/mcanet/15072507222/in/photostream/` (accessed October, 2015).

a physical timeline. These patterns can certainly be improved to favor more intuitive readings, especially if the goal is to provoke a response in the *reader* similar to the stimuli (i.e. relaxing music should produce physical objects capable of inducing relaxation). In this regard, we believe that color, texture and shape will be decisive at the moment of defining new fabrication strategies.

Finally, in its current state *NeuroKnitting* might not be consider an interactive system, as pattern generation occurs de-attached from the knitting process (i.e. the affective recordings are used to generate a bitmap that later on is knitted with the help of an operator). This limitation is basically due to the the constraints of working with a restored knitting machine. However, the new *Knitic* circular machine accounts for a completely automatic knitting process. The incorporation of such a tool into *NeuroKnitting* will represent a step toward a full implicit interactive fabrication system.

8. Conclusion

The main goal of this dissertation has been to conceptualize and prototype sonic interaction designs (SID) based on the implicit cues of human psychophysiology, and evaluate them in the context of HCI. For achieving it, we have leveraged on physiological computing (PhyComp) techniques, namely EEG and ECG, to obtain inputs from user's implicit (i.e. perceptive, emotive and cognitive) states in real time, and applied diverse SID methodologies to adapt system responses according to these implicit statuses.

For determining to what extent different sonic strategies aid the perception of implicit states, and improve user experience in specific computer-mediated domains, we have incrementally built up implicit sonic interactions (from direct audification to complex musical mappings) and evaluated them within HCI scenarios (from neurofeedback to music performance), assessing their perceptualization quality, the role of mapping complexity, and their meaningfulness in the musical domain.

This thesis has therefore evolved gradually, with each chapter tackling concrete research problems that emerge from the intersection of PhyComp and SID, providing specific contributions (and questions) that we summarize and discuss, under the scope of their limitations and possible expansions, in the following sections.

8.1. Contribution 1: understanding the context

SID applied to implicit PhyComp is a rather young and emergent field, and this dissertation has contributed to **identify and understand the trends, techniques and problems that emerge from the intersection of PhyComp and SID**. Through a comprehensive review (presented in Chapter 2) which cover theoretical, technical, and design standpoints, we have shown that the interest for exploiting the implicit repertoire of human behavior is at the core of the HCI agenda. Additionally, we have analyzed the main techniques and contexts in which PhyComp is used for creating adaptive systems that directly access to the implicit psychophysiological states of the user, thus promoting personalized responses that can be delivered implicitly or explicitly to improve user experience. In this context, we have studied the role of SID for displaying psychophysiological states, and described how sound can be used to represent the fast temporal dynamics of physiological signals, leveraging on human auditory perception (which provides the highest temporal resolution among the sensory modalities). Moreover, we studied the affordances of both representational, knowledge-driven approaches (i.e. sonifications) and aesthetic, performance-driven approaches (i.e. NIME).

Through this systematic analysis, we have identified specific challenges and problems about SID applied to the electroencephalograph (EEG), which is the main biosignal used in this dissertation:

- Many previous works lack technical details on the SID strategies, EEG equipment, and the processing techniques used for their implementation.
- Almost none of the analyzed works have carried out a systematic validation of their SID methods. In other words, by looking at the current literature, it was uncertain what types of SID strategies are most efficient for representing implicit physiological data in real time, and in different HCI scenarios.
- Very few of the studies that address PhyComp and SID have explored multimodality (e.g. the combination of sonic interactions with visualization). Given that human perception is multi sensorial, combining auditory, visual and tactile information is likely to produce enhanced PhyComp interfaces for both functional and aesthetic purposes.

This assessment work and its contributions have been published in a conference paper [Väljamäe et al., 2013a]

8.1.1. Limitations and possible expansions

It is important to note that, whereas our assessment on implicit interaction (from an HCI perspective), PhyComp and SID have been rather extensive, our literature review on sonic displays applied to human physiology has focused mainly on EEG. This narrow approach is due to the fact of EEG being the main biosignal explored in this dissertation. However, our current work could be further expanded by analyzing SID approaches to other PhyComp techniques, such as electromyography (EMG), electrocardiography (ECG) or electrodermal activity (EDA). Works in this direction are already emerging, like Caramiaux et al. [2015b] and Donnarumma and Tanaka [2014], both from a musical and HCI standpoint. Such efforts will account for a holistic overview of the field, which was out of the scope of this thesis.

8.2. Contribution 2: enhancing user motivation through implicit PhyComp

Through the *b-Reactable* prototype, we have studied how physiology-driven implicit sonic interaction affects user motivation in HCI contexts (i.e. collaborative performance using a digital music instrument) compared to other eminently explicit interfaces (e.g. the Reactable). The experiments presented in Chapter 3 have shown that **motivation aspects of user experience are significantly affected by physiology-driven sonic interactions, concretely in terms of greater confidence, positive affective response (valence) and stronger social affinity between dyads.** Our between-group analyses have also suggested that **the introduction of an incongruent physiological feedback (i.e. sham) significantly affects user experience, deteriorating confidence and communication between participants**.

The results of the first study with the b-Reactable not only offered evidence on the potential of physiology-based implicit interaction for improving single and multi user HCI within the musical domain, it also helped us to detect three relevant aspects of SID applied to implicit PhyComp:

- *Perceptualization*, understood as the process of associating a given display strategy (in this case sound) to the psychophysiological state that acts as the input for its rendering, according to the end-user perception.
- *Mapping complexity* in physiology-based sonic interaction, understood as the number of physiological streams and sound parameters used in a given SID strategy.
- *Meaningfulness* in the NIME context, understood as the potential of physiology-driven implicit interaction for being perceived as an expressive component of the DMI, through which the player can produce musical processes that, being expected or unexpected, contribute to the creative task she/he is committed to.

These aspects have served us as guidelines for our subsequent exploration of implicit, physiology-driven sonic interaction, and the contributions in this regard have resulted in a journal Mealla et al. [2016] and two conference papers [Mealla et al., 2011b,a, b]

8.2.1. Limitations and possible expansions

Although the first version of the *b-Reactable* has shown to have a significant positive effect on participants' motivation, and its sonic designs have been properly perceived by participants, other sonic strategies strategies should be tested to determine whether the effects reported in this regard are specific to our sonic strategies or can be generalized. In the same direction, personalization of mappings by end-users could be also explored in the proposed collaborative setup, as the first version of the *b-Reactable* applied fix, pre-defined mappings. Our subsequent experiments have shown that mapping personalization to play a significant role in perceptualization and user experience.

We should also stress the fact that implicit interaction paradigms like the ones present in the *b-Reactable* can also lead to self-adaptation (i.e. neuro and biofeedback) with participants trying to alter their EEG and ECG activity to match a given sound or tempo. However, this is more likely to happen after a number of training sessions, thus requiring further investigation. Future developments in this regard could determine whether physiology-driven implicit sonic interactions widen multi-user communication, or increase interpersonal synchronization in computer supported cooperative work (CSCW).

The SID strategies implemented in the first version of the *b-Reactable* were simple and straightforward (i.e. audification of EEG and BPM control through ECG) but the fact of using two biosignals simultaneously made difficult to identify the specific impact that each of these had in participants during the experiment. Moreover, it should be noted that the effects found in our study might be of temporal nature, thus future experiments should address the impact of prolonged use of *physiopucks*. In this regard, the literature suggests that this kind of physiology-based interaction is likely to produce subjective experiences different from gesture-based control, as in the case of the BRAAMHS musical interface based on functional near-infrared spectroscopy (fNIRS) Yuksel et al. [2015].

8.3. Contribution 3: perceptualization and mapping complexity

In this dissertation we have also explored two main aspects of SID applied to implicit PhyComp: *perceptualization* (how well a sonic design represents a given implicit physiological state, aiding user perception), and *mapping complexity* (the number of physiological streams and sound parameters used). The studies presented in Chapter 4 have provided empirical evidence on the ***perceptualization*** **quality of parameter mapping sonification and musical mapping for representing implicit physiological states (specifically relaxation), compared to direct sonification techniques**. Parameter mapping sonification has achieved better relaxation effects than direct sonic strategies, but being still comparable with the results of musical mapping.

These experiments have also provided valuable insights about the role that both *mapping complexity* and end-user personalization plays in the *perceptualization* of sonic designs for implicit PhyComp. **Personalized mappings have shown to be more instrumental**

than fixed ones for displaying implicit physiological states through sound (specifically relaxation estimated according to the *a/t* neurofeedback protocol). **Complex physiology-to-sound mappings, based on multiple EEG features, have shown to be more efficient than sonic designs relying on a single EEG feature, for both perceptualizing and relaxation induction**. Finally, the studies have also offered evidence about **personalization becoming less instrumental when multiple physiological features are displayed through sound** (probably due to a decrease on the user attentional resources).

In this dissertation we have additionally contributed with **methodologies for evaluating *perceptualization*, *mapping complexity* and *personalization* of sonic designs for implicit PhyComp**. These methods are based on the combination of objective and subjective measures, and a well-known perception-based scenario (i.e. neurofeedback). This contribution aimed to tackle one of the main issues detected in our literature review, which is that most of the previous works in the field lack a structured validation, making it difficult to determine the efficiency of a particular sonic design for conveying a given implicit psychophysiological state.

Finally, we have proposed a **sonic engine that can be easily configured and expanded for designing a great variety of sonic interactions for different psychophysiological states** (e.g. engagement, cognitive load) and biosignals (e.g. muscle activity, heart rate, etc).

The contributions in this regard have been published in a conference paper [Mealla et al., 2014] and are part of a journal paper currently under review [Mealla et al., in review].

8.3.1. Limitations and possible expansions

In order to corroborate the perceptualization and brain dynamics obtained in our neurofeedback experiments, more training sessions, a larger number of participants, and different EEG equipments could be applied. Such approach will help to detect changes between training sessions and learning curves, both at group and individual levels. Other ways of expanding our current work would be testing different biofeedback protocols (for determining the affordances of SID to perceptualize and induce other psychophysiological states such as cognitive load or engagement) or evaluating perceptualization and mapping complexity in domains such as data mining, diagnosis or entertainment.

Regarding the effect of mapping personalization, the experiments presented in this dissertation have demonstrated that our methods for mapping adjustment were clear and understandable for naive end-users. However, a specific study on interface design for mapping personalization should be performed.

Finally, significant between-groups differences in both subjective and physiological data suggest that our results on *perceptualization* and *mapping complexity* are not obtained merely due to relaxing nature of presented sounds. However, some of the discrepancies that we have encountered between subjective ratings and physiological indices suggest that more experiments have to be carried out, encouraging the use of multimodal measures for correcting and interpreting new results.

8.4. Contribution 4: NIME design and evaluation

In this dissertation we have also presented and tested a **framework for NIME design and evaluation**, focused on *System* and *Performance* aspects of digital music instruments (DMI), that considered music knowledge of different stakeholders, and meant to inform iterative design processes (see Chapter 5).

Through the deployment of this methodology within a master course, we have detected that **all the participants** (some of whom had never performed music, nor programmed computers) **have been able to effectively engage in the creation of DMIs**, following specific constraints imposed during the course. Also, **the assessment tools included in this framework have proven to be useful for evaluating and informing iterative NIME design processes**.

Although these findings have been obtained in the specific context of a master course, we believe that several of these solutions and learnings could be extrapolated to more generic contexts, being other NIME or even HCI courses, and used to inform teachers, designers and practitioners in general.

The contributions in this regard have been published in a conference paper Jordà and Mealla [2014].

8.4.1. Limitation and possible expansions

The study presented in Chapter 5 should be seen as a first step towards the creation of a more general NIME framework. In particular, our current approach could be complemented and expanded with qualitative methods such as interviews and focus groups that will allow to collect valuable insights, particularly from performers and listeners. At Universitat Pompeu Fabra, we have been applying this framework in a master course for the last two years, extending the evaluation from the *listener* perspective to the standpoint of *designers* and *performers,* and including structured interviews and focus groups. In this regard, we are preparing a publication that will report the process and main findings in this direction[1].

Other ways to strengthen our current NIME methodology would be to test it longitudinally with bigger samples and in different NIME scenarios (e.g. workshops) beyond the specific scope of the course presented in this dissertation. For instance, we envision a study where performers could select their favorite DMIs designed by other participants and perform with them for later evaluation as *performers*. Finally, the proposed methods could be also tested with different design constraints (other than the ones applied in our study). In this regard, future work could tackle design guidelines related to touch/force or embodied interaction, as suggested by Zappi and McPherson [2014].

8.5. Contribution 5: Implicit PhyComp supporting musical expression

In this dissertation we have also explored the *meaningful* integration of implicit PhyComp in a NIME context. In Chapter 6 we have addressed to what extent music performers perceived implicit PhyComp as an expressive component of a DMI (i.e. a second version of the *b-Reactable*) through which they could produce musical processes that, being expected or unexpected, contribute to the creative task they were committed to.

Our study has shown that **implicit sonic interactions can be used in a comprehensive and meaningful manner during music performance**, according to the requested affective targets based on a two-dimensional model of valence and arousal. In the same line, subjective, behavioral and physiological data showed that **Global and Local implicit interactions were perceived in significantly distinctive ways according to participants' previous musical experience**, with preference for the Local control.

8.5.1. Limitation and possible expansions

Our studies with the *b-Reactable* have offered useful information about implicit PhyComp supporting musical expression. Nonetheless, there are a number of limitations that have to be considered and eventually tackled by future work. One limitation is related to the proposed physiology-to-sound mappings. Although most participants were able to perceive valence and arousal states through the displayed sounds, some were not entirely satisfied with the manner the feedback was presented to them. A way to meet

[1] Documentation on the developed DMIs within the master course (both videos and written reports) can be found at `http://www.dtic.upf.edu/~smealla/phd_material.html` (accessed on November, 2015).

user expectations in this regard would be improving our current personalization mechanisms by adding a graphic user interface (GUI) for in-depth mapping customization.

In the same way, our NIME framework could be further expanded to cover different stakeholders, *listeners* could evaluate the performances done with the *b-Reactable*. Such approach would allows us to assess the congruency between *performer* and *listener* perceptions on musical improvisations. Also, other evaluation frameworks could be applied to verify the consistency of our current findings.

Future studies on the meaningful integration of implicit PhyComp into NIME could also go beyond EEG. The explorations of other biopotentials such as EMG or respiration will allow to estimate other psychophysiological features (e.g. stress) that could complement implicit affective estimations like the ones presented in this thesis, even promoting multimodal explicit/implicit control, as it has been done for video games [Nacke et al., 2011]. The *b-Reactable* could also integrate multiple Local and Global *physiopucks,* responding to different psychophysiological estimations. Finally, self-regulation and implicit learning could be further explored by applying longer and repeated sessions. In this manner, we could determine whether participants *get better* in perceiving and controlling implicit sonic interactions in a musical performance context.

It is evident that several aspects of this work are specific to music. Whereas music appears as an excellent candidate for exploring the contributions of PhyComp-based implicit interaction in an expressive context, it also imposes specific constrains in the way control mappings are defined and in users' previous knowledge. Therefore, Global and Local implicit interaction modes could be further explored in other domains, ranging from gaming to learning or big data exploration.

8.6. Contribution 6: taking implicit PhyComp beyond SID

At the end of this dissertation we have proposed to explore implicit PhyComp beyond sound and music. In order to do so, we have tackled the field of personal fabrication presenting *NeuroKnitting*, a system for creating knitted garments according to the users' affective responses estimated from EEG. We consider *NeuroKnitting* a design contribution, as it has not pursued specific research problems. However, the insights collected during its testing could certainly lead to structured studies on aspects such as the **perception of implicitly generated fabrication patterns**, the use of different stimuli (e.g. music, films, images) to drive fabrication processes, and the exploration of other biopotentials beyond EEG (e.g. heart rate variability, electrodermal activity).

NeuroKnitting also serves as a **demonstrator on how implicit PhyComp can be used to fabricate tangible objects.** However, with NeuroKnitting we have not exploited tangibility in depth, thus we envision other possibilities in this regard, such as mapping psychophysiological states to the shape of objects and garments, or to tactile properties such as texture. These approaches might lead to more intuitive physical representations of implicit psychophysiological states. For instance, soft or coarse patterns could help to differentiate users' relaxation and excitement, producing more evocative objects.

Methods like 3D printing could also be explored, and fabrication patterns could be improved to provoke a response, similar to the stimuli that originated the object (e.g. relaxing music should produce physical objects capable of inducing or representing relaxation). We believe that color, texture and shape will be decisive at the moment of defining new fabrication strategies in this regard.

Finally, it is important to note that, in its current state, *NeuroKnitting* might not be considered an interactive system, as pattern generation occurs de-attached from the knitting process. In order to overcome this limitation, we could update the knitting system with a different hardware platform (e.g. the *Knitic* circular machine) to achieve a fully automatic knitting process.

Neuroknitting has been showcased in a number of public events and exhibitions such as Maker Faire Rome 2013, Maribor Art Gallery[2], Art Deal Project [3], and Sonar Festival (2013) in Barcelona.

8.7. Non-academic contributions

8.7.1. Workshops and Hackathons

During the realization of this thesis we have also carried knowledge transfer activities aimed to take Physiological Computing technology out of the lab and deploy it in music technology contexts. Most of these actions took place within the Barcelona Music Hack Day (MHD)[4], a 24-hours hacking session in which participants conceptualize, create and present music technology projects (i.e. hacks). The MHD was organized for the first time in July 2009 and since then has proven to be a remarkable way to demonstrate the

[2] `http://mcanet.info/files/catalogue_maribor_25x30_op.pdf` (accessed on November, 2015).
[3] `http://www.artdealproject.com/index.php?lang=esp&mod=03&expos=42` (accessed on November, 2015).
[4] `http://musichackday.upf.edu` (accessed on November, 2015)

creativity around the music technology community, fostering cross-platform and cross-device innovation. The MHD was brought to Barcelona in 2010 by the Music Technology Group of the Universitat Pompeu Fabra.

In 2013, with the support of the EC funded project KiiCS (Knowledge Incubation in Innovation and Creation for Science)[5], we created a special neuroscience track within the MHD, that aimed to provide a set of useful tools and APIs to encourage hacks that bring together music and physiological computing[6]. Through this approach, we encouraged the creation of prototypes that fostered new ideas around music creation and interaction. The initiative gathered 100 hackers from all around the world, and counted with the support of Starlab Barcelona.

In 2015, in the context of the EU funded project RAPID-MIX[7], we organized a second special track on wearable and multimodal technology that brought together experts on physiological and motion sensing, interaction design and interface prototyping. We offered a hands-on workshop on prototyping expressive wearable technology for music performance, where participants had the chance of combining innovative physiological sensing (using the BITalino board[8]), real-time machine learning interfaces (Wekinator[9]) and audio synthesis/processing libraries (Maximilian[10] and JUCE[11]) for prototyping wearable and mobile music interfaces[12].

8.8. Closing remarks

Implicit PhyComp offers us unparalleled advantages for creating a new generation of interactive systems capable of tailoring content and experience to the user perceptive, cognitive and affective states, in a seamless and unobtrusive way. As PhyComp technology becomes more ubiquitous, reliable and affordable, it is expected that these systems will gain greater importance in a wide range of creative industries such as video games, sports, quantified self and, of course, music. We can already see them in the streets,

[5] `http://www.kiics.eu/en/` (accessed on November, 2015).

[6] A video teaser of the event and the presentation of projects developed during the MHD can be found at `https://www.youtube.com/watch?v=SbzkcA4G0Wo&list=PL0EPT_FWD3MlcDoNTKpLBeOETOHJzfLA8` (accessed on November, 2015)

[7] `http://rapidmix.goldsmithsdigital.com/`(accessed on November, 2015)

[8] `http://www.bitalino.com/` (accessed on November, 2015)

[9] `http://www.wekinator.org/`(accessed on November, 2015)

[10] `http://eavi.goldsmithsdigital.com/research/maximilian/` (accessed on November, 2015)

[11] `http://www.juce.com/` (accessed on November, 2015).

[12] A video summarizing the event can be found at `https://www.youtube.com/watch?v=1sqFnHAmAh0` (accessed on November, 2015)

in the shape of wearables or mobile devices. However, most of these interfaces are still limited to activity tracking and monitoring, with very few cases exploring real-time interaction. Differently from direct PhyComp, which faces important constraints such as a limited control bandwidth, accuracy, and limitations of human attention, implicit PhyComp is capable of enhancing human bandwidth by providing adaptive responses perceived as intuitive and timely by users, without consciously generated input commands or training. In this context, we believe sound plays a major role to encompass implicit PhyComp in HCI without burdening the user, and for seamlessly incorporating multimodality when combined with other input methods. Due to the aforementioned reasons, this dissertation represents represents a step towards the meaningful integration of implicit PhyComp in relevant HCI scenarios, that it will hopefully inspire other scholars and practitioners to further explore this path beyond the musical domain.

Bibliography

Edgar Douglas Adrian and K Yamagiwa. The origin of the berger rhythm. *Brain*, 58 (3):323–351, 1935. 2.6, 2.6.3

Jennifer Allanson. Electrophysiologically interactive computer systems. *Computer*, 35 (3):60–65, 2002. 1.2.1

Jennifer Allanson and Stephen H Fairclough. A research agenda for physiological computing. *Interacting with computers*, 16(5):857–878, 2004. 1.1, 2.3

Brendan Allison, Axel Gräser, and Bernhard Graimann. Why Use A BCI If You Are Healthy? *ACE Workshop BrainComputer Interfaces and Games*, pages 7–11, 2007. 1.2.1, 6.2.3

Efthymios Angelakis, Stamatina Stathopoulou, Jennifer L Frymiare, Deborah L Green, Joel F Lubar, and John Kounios. Eeg neurofeedback: a brief overview and an example of peak alpha frequency training for cognitive enhancement in the elderly. *The Clinical Neuropsychologist*, 21(1):110–129, 2007. 4.2, 4.2.1

Daniel Arfib, Jean-Michel Couturier, Loic Kessous, Vincent Verfaille, et al. Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, 7(2):127–144, 2002. 5.3.2

Burak Arslan, Andrew Brouse, Julien Castet, Jean-Julien Filatriau, Rémy Lehembre, Quentin Noirhomme, and Cédric Simon. Biologically-driven musical instrument. In *Proceedings of the Summer Workshop on Multimodal Interfaces (eNTERFACE 05)*. Faculte Polytechnique de Mons Mons, BL, 2005. 5, 7, 4.3.1

Burak Arslan, Andrew Brouse, Julien Castet, Remy Léhembre, Cédric Simon, Jean Julien Filatriau, and Quentin Noirhomme. A real time music synthesis environment driven with biological signals. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 2, pages II–II. IEEE, 2006. 4.3.1

World Medical Association et al. World medical association declaration of helsinki. ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4):373, 2001. 3.3

Gerold Baier, Thomas Hermann, Sven Sahle, and Ulrich Stephani. Sonified epilectic rhythms. 2006. 4, 2.6.4

Gerold Baier, Thomas Hermann, and Ulrich Stephani. Event-based sonification of eeg rhythms in real time. *Clinical Neurophysiology*, 118(6):1377–1386, 2007a. 2.6.4

Gerold Baier, Thomas Hermann, and Ulrich Stephani. Multi-channel sonification of human eeg. In *Proceedings of the 13th International Conference on Auditory Display*, 2007b. 2.6.4

Jerônimo Barbosa, Filipe Calegario, Francisco Magalhães, V Teichrieb, G Cabral, and Geber Ramalho. Towards an evaluation methodology for digital music instruments considering performerś view: a case study. In *Proceedings of 13th Brazilian Symposium on Computer Music*, 2011. 5.2.2

Stephen Barrass. Diagnosing blood pressure with acoustic sonification singing bowls. *International Journal of Human-Computer Studies*, 2015. 7.3

Stephen Barrass and Bjōrn Zehner. Responsive sonification of well-logs. In *Proc. Int. Conf. on Auditory Display (ICAD-00)*, 2000. 2.4.1

Stephen Barrass, Mitchell Whitelaw, and Freya Bailes. Listening to the mind listening: An analysis of sonification reviews, designs and correspondences. *Leonardo Music Journal*, 16:13–19, 2006. 17

Martin J Batty, Samantha Bonnington, Bo-Kim Tang, Malcolm B Hawken, and John H Gruzelier. Relaxation strategies and enhancement of hypnotic susceptibility: Eeg neurofeedback, progressive muscle relaxation and self-hypnosis. *Brain research bulletin*, 71(1):83–90, 2006. 4.2.1

Olivier Bau, Atau Tanaka, and Wendy E Mackay. The a20: Musical metaphors for interface design. In *Proc. NIME*, pages 91–96, 2008. 2.4.2

Victoria Bellotti, Maribeth Back, W Keith Edwards, Rebecca E Grinter, Austin Henderson, and Cristina Lopes. Making sense of sensing systems: five questions for designers and researchers. In *Proc. of CHI*, pages 415–422. ACM, 2002. 5.2.2

Hans Berger. Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 102(1):538–557, 1934. 2.3.1, 2.5.1

Sebastian Bitzer and Patrick van der Smagt. Learning emg control of a robotic hand: towards active prostheses. In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pages 2819–2823. IEEE, 2006. 2.3.2

Alice Bodanzky. Exploring the expressiveness of shape-changing surfaces. In *Proceedings* *of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, pages 403–404. ACM, 2012. 6.2.3

Danny Oude Bos. Eeg-based emotion recognition. *The Influence of Visual and Auditory Stimuli*, pages 1–17, 2006. 6.2.2, 6.3.3

Wolfram Boucsein. *Electrodermal activity.* Springer Science & Business Media, 2012. 2.3.1

Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994. 3.3.5, 4.3.1, 6.3.3, 3

Thomas H Budzynski and Johann M Stoyva. An instrument for producing deep muscle relaxation by means of analog information feedback1. *Journal of Applied Behavior Analysis*, 2(4):231–237, 1969. 4.2.1

William Buxton. Integrating the periphery and context: A new taxonomy of telematics. In *Proceedings of graphics interface*, volume 95, pages 239–246. Citeseer, 1995. 1.1, 2.2.1, 2.1, 2.2.1

Claude Cadoz, Marcelo M Wanderley, et al. Gesture-music. *Trends in gestural control of music*, 12:101, 2000. 5.2

Gemma Calvert, Charles Spence, and Barry E Stein. *The handbook of multisensory processes.* MIT press, 2004. 2.7

Baptiste Caramiaux, Alessandro Altavilla, Scott G Pobiner, and Atau Tanaka. Form follows sound: Designing interactions from sonic memories. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3943–3952. ACM, 2015a. 2.4.2

Baptiste Caramiaux, Marco Donnarumma, and Atau Tanaka. Understanding gesture expressivity through muscle sensing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(6):31, 2015b. 1.2.1, 2.3.1, 8.1.1

John M Carroll. *Human computer interaction-brief intro.* The Interaction Design Foundation, 2013. 1.1

Charles S Carver, Michael F Scheier, M Boekaerts, PR Pintrich, and M Zeidner. On the structure of behavioral self-regulation. *Handbook of self-regulation*, pages 41–84, 2000. 1.2.1, 6.2.1

Ginevra Castellano, Santiago D Villalba, and Antonio Camurri. Recognising human emotions from body movement and gesture dynamics. In *Affective computing and intelligent interaction*, pages 71–82. Springer, 2007. 6.2.2

Joel Chadabe. Interactive composing: An overview. *Computer Music Journal*, 8(1): 22–27, 1984. 5.3.2

John M Chowning. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, 21(7):526–534, 1973. 1

F Cincotti, F Pichiorri, P Aricò, F Aloise, F Leotta, F de Vico Fallani, Jose del R Millan, Marco Molinari, and D Mattia. Eeg-based brain-computer interface to support post-stroke motor rehabilitation of the upper limb. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4112–4115. IEEE, 2012. 1.2.1

Christian Cipriani, Franco Zaccone, Silvestro Micera, and M Chiara Carrozza. On the shared control of an emg-controlled prosthetic hand: analysis of user–prosthesis interaction. *Robotics, IEEE Transactions on*, 24(1):170–184, 2008. 1.2.1

John P Clark and Larry L Tifft. Polygraph and interview validation of self-reported deviant behavior. *American Sociological Review*, pages 516–523, 1966. 2.3.1

SC Collins. *Subjective and autonomic responses to Western classical music.* PhD thesis, University of Manchester, 1989. 4.2.1

Perry Cook. Principles for designing computer music controllers. In *Proc. of NIME*, pages 1–4. National University of Singapore, 2001. 5.2.2

Mihaly Csikszentmihalyi and Mihaly Csikzentmihaly. *Flow: The psychology of optimal experience*, volume 41. HarperPerennial New York, 1991. 1.2.1, 6.2.1

Richard J Davidson, Paul Ekman, Clifford D Saron, Joseph A Senulis, and Wallace V Friesen. Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of personality and social psychology*, 58(2):330, 1990. 6.2.2

Alberto De Campo, Robert Hoeldrich, Gerhard Eckel, and Annette Wallisch. New sonification tools for eeg data screening and monitoring. 2007. 1.1, 2.6.3, 4.3, 4.4

Andrew M Dearden and Michael D Harrison. Abstract models for hci. *International Journal of Human-Computer Studies*, 46(1):151–177, 1997. 6.2.3

Christopher Dobrian and Daniel Koppelman. The 'e' in nime: Musical expression with new computer interfaces. In *Proceedings of the 6th Conference on New Interfaces for Musical Expression*, NIME '06, pages 277–282, Paris, France, 2006. IRCAM; Centre Pompidou. ISBN 2-84426-314-3. 5.2.2

Marco Donnarumma. Music for flesh ii: informing interactive music performance with the viscerality of the body system. In *Proceedings of the NIME Conference, Ann Arbor, MI, USA*, 2012. 2.5.2

Marco Donnarumma and Atau Tanaka. Principles, challenges and future directions of physiological computing for the physical performance of digital musical instruments'. In *Proceedings of the Conference on Interdisciplinary Musicology. Berlin*, 2014. 8.1.1

Paul Dourish and Sara Bly. Portholes: Supporting awareness in a distributed work group. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 541–547. ACM, 1992. 2.2.1

Matthieu Duvinage, Thierry Castermans, Mathieu Petieau, Thomas Hoellinger, Guy Cheron, and Thierry Dutoit. Performance of the emotiv epoc headset for p300-based applications. *Biomed Eng Online*, 12:56, 2013. 4.2.1

Alice H Eagly and Shelly Chaiken. Attitude structure and function. 1998. 3.3.5

Tobias Egner and John H Gruzelier. Ecological validity of neurofeedback: Modulation of slow wave eeg enhances musical performance. *Neuroreport*, 14(9):1221–1224, 2003. 7.4.3

Tobias Egner, Emilie Strawson, and JohnH. Gruzelier. Eeg signature and phenomenology of alpha/theta neurofeedback training versus mock feedback. *Applied Psychophysiology and Biofeedback*, 27(4):261–270, 2002. ISSN 1090-0586. doi: 10.1023/A:1021063416558. 4.3, 4.3.1

Stephen H Fairclough. Fundamentals of physiological computing. *Interacting with computers*, 21(1):133–145, 2009. 1.1, 1.2.1, 14, 2.3, 1, 6.3.6

Stephen H Fairclough and Kiel Gilleade. *Advances in Physiological Computing*. Springer, 2014a. 1.2.1, 2.3.2, 2.3.2, 2.3.2

Stephen H Fairclough and Kiel Gilleade. Meaningful interaction with physiological computing. In *Advances in Physiological Computing*, pages 1–16. Springer, 2014b. 1.1, 2.3.2, 2.3.2, 2.3.2

Rebecca Anne Fiebrink. *Real-time human interaction with supervised learning algorithms for music composition and performance*. Princeton University, 2011. 5.3.2

Jehan-Julien Filatriau, Remy Léhembre, Quentin Noirhomme, Cédric Simon, Burak Arslan, Andrew Brouse, and Julien Castet. Using physiological signals for sound creation. *Proc. SPECOM 06*, 2006. 5, 7, 8, 4.3.1

Tecumseh Fitch and Gregory Kramer. Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system. In *Santa Fe Institute*

Studies In The Sciences Of Complexity-proceedings Volume-, volume 18, pages 307–307. Addison-Wesley Publishing Co, 1994. 1.1

Karmen Franinović and Stefania Serafin. *Sonic interaction design*. Mit Press, 2013. 1.2.2, 1.2.2, 2.4, 2.4.2

Barbara L Fredrickson and Christine Branigan. Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & emotion*, 19(3):313–332, 2005. 6.2.2

Alf Gabrielsson and Patrik N Juslin. Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of music*, 24 (1):68–91, 1996. 4.2.1

Alf Gabrielsson and Erik Lindström. The influence of musical structure on emotional expression. 2001. 4.2.2

John Gage. *Color and culture: Practice and meaning from antiquity to abstraction*. Univ of California Press, 1999. 7.4.4

William W Gaver. What in the world do we hear?: An ecological approach to auditory event perception. *Ecological psychology*, 5(1):1–29, 1993. 1.2.2

Lalya Gaye, Lars Erik Holmquist, Frauke Behrendt, and Atau Tanaka. Mobile music technology: Report on an emerging community. In *Proceedings of the 2006 conference on New interfaces for musical expression*, pages 22–25. IRCAM Centre Pompidou, 2006. 2.4.1

Steven Gelineck and Stefania Serafin. Longitudinal evaluation of the integration of digital musical instruments into existing compositional work processes. *Journal of New Music Research*, 41(3):259–276, 2012. 5.2.2

Neil Gershenfeld. *When things start to think*. Macmillan, 1999. 2.2

Neil Gershenfeld. *Fab: the coming revolution on your desktop–from personal computers to personal fabrication*. Basic Books, 2008. 7.1

Kiel Gilleade, Alan Dix, and Jen Allanson. Affective videogames and modes of affective gaming: assist me, challenge me, emote me. 2005. 1.1, 1.2.1

Kiel M Gilleade and Alan Dix. Using frustration in the design of adaptive videogames. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 228–232. ACM, 2004. 2.3.2

Ross Gillett. *MADDM–Dance-Directed Music*. Ann Arbor, MI: MPublishing, University of Michigan Library, 1985. 2.5.2

Nicholas Gillian, R Benjamin Knapp, and Sile O´ Modhrain. A machine learning toolbox for musician computer interaction. *Proc. of NIME*, 2011. 5.3.2

John Glen. Use of audio signals derived from electroencephalographic recordings as a novel depth of anaesthesia monitor. *Medical hypotheses*, 75(6):547–549, 2010. 2.6.1

Erving Goffman. *Interaction ritual: Essays in face to face behavior.* AldineTransaction, 2005. 2.2.5

Mick Grierson. Composing with brainwaves: minimal trial p300b recognition as an indication of subjective preference for the control of a musical instrument. In *Proceedings of the ICMC*, 2008. 2.5.2, 2.6.1

Florian Grond and Jonathan Berger. Parameter mapping sonification. *The sonification handbook*, pages 363–397, 2011. 2.4.1

John Gruzelier. A theory of alpha/theta neurofeedback, creative performance enhancement, long distance functional connectivity and psychological integration. *Cognitive Processing*, 10(1):101–109, 2009. ISSN 1612-4782. doi: 10.1007/s10339-008-0248-5. 4.3, 4.3.1, 7.4.1

Michael Gurevich and A. Cavan Fyans. Digital musical interactions: Performer - system relationships and their perception by spectators. *Organised Sound*, 16:166–175, 6 2011. ISSN 1469-8153. doi: 10.1017/S1355771811000112. 5.2.2

Michael Gurevich, Ben Knapp, and Sergi Jordà. A Workshop on NIME Education. In *Proc. of NIME*, 2011. 5.2.1

Andreas Haag, Silke Goronzy, Peter Schaich, and Jason Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. In *ADS*, pages 36–48. Springer, 2004. 2.3.1

Zahid Halim, Rauf Baig, and Shariq Bashir. Temporal patterns analysis in eeg data using sonification. In *Information and Emerging Technologies, 2007. ICIET 2007. International Conference on*, pages 1–6. IEEE, 2007. 2.6.3

Brahim Hamadicharef, Mufeng Xu, and Sheel Aditya. Brain-computer interface (bci) based musical composition. In *Cyberworlds (CW), 2010 International Conference on*, pages 282–286. IEEE, 2010. 2.5.2, 6

James V Hardt and Joe Kamiya. Anxiety change through electroencephalographic alpha feedback seen only in high anxiety subjects. *Science*, 201(4350):79–81, 1978. 4.2

Thomas Hermann. Taxonomy and definitions for sonification and auditory display. 2008. (document), 1.2.2, 1.2, 2.10

Thomas Hermann and Andy Hunt. Guest editors' introduction: An introduction to interactive sonification. *IEEE multimedia*, (2):20–24, 2005. 2.4.1, 2.4.1, 2.4.2

Thomas Hermann and Helge Ritter. Model-based sonification revisited—authors' comments on hermann and ritter, icad 2002. *ACM Transactions on Applied Perception (TAP)*, 2(4):559–563, 2005. 2.4.1, 2.6.3

Thomas Hermann, Peter Meinicke, Holger Bekel, Helge Ritter, Horst M Müller, and Sabine Weiss. Sonifications for eeg data analysis. 2002. 1.1, 2.6.3, 2.6.4

Thomas Hermann, Gerold Baier, Ulrich Stephani, and Helge Ritter. Vocal sonification of pathologic eeg features. 2006. 5

Thomas Hermann, Till Bovermann, Eckard Riedenklau, and Helge Ritter. Tangible computing for interactive sonification of multivariate data. In *Proceedings of the 2nd International Workshop on Interactive Sonification*, 2007. 3.5

Thomas Hermann, Gerold Baier, Ulrich Stephan, and Helge Ritter. Kernel regression mapping for vocal eeg sonification. 2008. 8

Lawrence J Hettinger, Pedro Branco, L Miguel Encarnacao, and Paolo Bonato. Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces. *Theoretical Issues in Ergonomics Science*, 4(1-2):220–237, 2003. 1.1, 2.3

Kate Hevner. The affective value of pitch and tempo in music. *The American Journal of Psychology*, pages 621–630, 1937. 4.2.1

Thilo Hinterberger. The sensorium: a multimodal neurofeedback environment. *Advances in Human-Computer Interaction*, 2011:3, 2011. 2.6.4

Thilo Hinterberger and Gerold Baier. Parametric orchestral sonification of eeg in real time. *IEEE MultiMedia*, (2):70–79, 2005. 1.1, 2

Thilo Hinterberger, Jeremy Hill, and Niels Birbaumer. An auditory brain-computer communication device. In *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*, pages S3–6. IEEE, 2004. 2.6.1, 2.6.4

B Hjorth. Time domain descriptors and their relation to a particular model for generation of eeg activity. *CEAN-Computerized EEG analysis*, pages 3–8, 1975. 4.3.1

Richard W Homan, John Herman, and Phillip Purdy. Cerebral location of international 10–20 system electrode placement. *Electroencephalography and clinical neurophysiology*, 66(4):376–382, 1987. 2.6.4, 4.2.1, 6.3.2, 7.4.1

Jonathan Hook, David Green, John McCarthy, Stuart Taylor, Peter Wright, and Patrick Olivier. A vj centered exploration of expressive interaction. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems, pages 1265–1274. ACM, 2011. 6.2.3

Eva Hornecker. The role of physicality in tangible and embodied interactions. *interactions*, 18(2):19–23, 2011. 1.2.2

David Howell. *Fundamental statistics for the behavioral sciences.* Cengage Learning, 2013. 3.4

Andy Hunt, Marcelo M Wanderley, and Matthew Paradis. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research*, 32(4):429–440, 2003. 2.4.2, 5.3.2, 5.3.2

Ann Jones and Kim Issroff. Learning technologies: Affective and social issues in computer-supported collaborative learning. *Computers & Education*, 44(4):395–408, 2005. 3.1, 3.3.5, 3.5

Sergi Jorda. Digital lutherie: Crafting musical computers for new musics performance and improvisation. *PhD diss., Universitat Pompeu Fabra, Departament de Tecnologia*, 2005. 2.11, 2.4.2, 5.2.2, 5.3.2, 5.3.3

Sergi Jordà. On stage: the reactable and other musical tangibles go real. *International Journal of Arts and Technology*, 1(3):268–287, 2008. 5.3.1

Sergi Jordà and Sebastián Mealla. A methodological framework for teaching, evaluating and informing nime design with a focus on mapping and expressiveness. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 233–238, London, United Kingdom, June 30 – July 03 2014. Goldsmiths, University of London. (document), 1.2.1, 5.1, 5.2, 5.3, 6.2.1, 8.4

Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 139–146. ACM, 2007. (document), 1.3, 3.1, 3.2.2, 3.3

Emil Jovanov, Dusan Starcevic, Vlada Radivojevic, Aleksandar Samardzic, and Vladimir Simeunovic. Perceptualization of biomedical data. an experimental environment for visualization and sonification of brain electrical activity. *Engineering in Medicine and Biology Magazine, IEEE*, 18(1):50–55, 1999. 1.1, 3.6, 4.1

Wendy Ju and Larry Leifer. The design of implicit interactions: Making interactive systems less obnoxious. *Design Issues*, 24(3):72–84, 2008. 1.1, 2.2, 2.2.1, 2.2.5, 2.3

Patrik N Juslin. Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae scientiae*, 1(2):225–256, 1997. 4.2.1, 6.3.4

David B Kaber, Carlene M Perry, Noa Segall, Christopher K McClernon, and Lawrence J Prinzel. Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. *International Journal of Industrial Ergonomics*, 36(5):447–462, 2006. 1.1

Martin Kaltenbrunner and Ross Bencina. reactivision: a computer-vision framework for table-based tangible interaction. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 69–74. ACM, 2007. 3.2.2

Eugenijus Kaniusas. *Biomedical signals and sensors I: Linking physiological phenomena and biosignals*. Springer Science & Business Media, 2012. (document), 2.3.1, 2.4, 2.5, 2.3.1, 2.3.1, 2.8, 2.9, 2.5.1

H Khamis, A Mohamed, S Simpson, and A McEwan. Detection of temporal lobe seizures and identification of lateralisation from audified EEG. *Clinical Neurophysiology*, 123 (9):1714–20, September 2012. ISSN 1872-8952. doi: 10.1016/j.clinph.2012.02.073. 2.6.1

Chris Kiefer. *Multiparametric interfaces for fine-grained control of digital music*. PhD thesis, University of Sussex, 2012. 5.2.2, 5.3.2

J Matias Kivikangas, Guillaume Chanel, Ben Cowley, Inger Ekman, Mikko Salminen, Simo Järvelä, and Niklas Ravaja. A review of the use of psychophysiological methods in game research. *journal of gaming & virtual worlds*, 3(3):181–199, 2011. 4

Wolfgang Klimesch. Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2):169–195, 1999. 2.3.2

R Benjamin Knapp and Hugh S Lusted. A bioelectric controller for computer music applications. *Computer music journal*, pages 42–47, 1990. 2.3.1, 2.5.2

Sander Koelstra, Christian Mühl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. Deap: A database for emotion analysis; using physiological signals. *Affective Computing, IEEE Transactions on*, 3(1):18–31, 2012. 6.3.7

Gregory Kramer. *Auditory display: Sonification, audification, and auditory interfaces*. Perseus Publishing, 1993. 1.2.2

Gregory Kramer, Bruce Walker, Terri Bonebright, Perry Cook, John H Flowers, Nadine Miner, and John Neuhoff. Sonification report: Status of the field and research agenda. 2010. 2.4.1, 2.4.1

Juri Kropotov. *Quantitative EEG, event-related potentials and neurotherapy*. Academic Press, 2010. 2.3.1, 2.3.2, 4.2.1, 4.3.1, 7.4.1

Nelson Posse Lago and Fabio Kon. The quest for low latency. In *Proceedings of the International Computer Music Conference*, pages 33–36, 2004. 3.2.5

Peter J Lang, Margaret M Bradley, Bruce N Cuthbert, et al. International affective picture system (iaps): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*, 1999. 6.3.3

PJ Lang and MM Bradley. International affective digitized sounds (iads): stimuli, 802 instruction manual and affective ratings (tech. rep. no. b-2). *Gainsville, FL*, 803, 1999. 6.3.3

Sylvain Le Groux, Jonatas Manzolli, PFMJ Verschure, Marti Sanchez, Andre Luvizotto, Anna Mura, Aleksander Valjamae, Christoph Guger, Robert Prueckl, and Ulysses Bernardet. Disembodied and collaborative musical interaction in the multimodal brain orchestra. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 309–314, 2010. 2.6.3, 6

Paul D. Lehrman and Todd M. Ryan. Bridging the Gap Between Art and Science Education Through Teaching Electronic Musical Instrument Design. In *Proc. of NIME*, pages 136–139, 2005. 5.2.1

Marc Leman. *Embodied music cognition and mediation technology.* Mit Press, 2008. 2.4.2

Yuanqing Li, Cuntai Guan, Huiqi Li, and Zhengyang Chin. A self-training semi-supervised svm algorithm and its application in an eeg-based brain computer interface speller system. *Pattern Recognition Letters*, 29(9):1285–1294, 2008. 2.3.2

Sheng-Fu Liang, Yi-Chun Chen, Yu-Lin Wang, Pin-Tzu Chen, Chia-Hsiang Yang, and Herming Chiueh. A hierarchical approach for online temporal lobe seizure detection in long-term intracranial eeg recordings. *Journal of neural engineering*, 10(4):045004, 2013. 4.3.1

Joseph Carl Robnett Licklider. Man-computer symbiosis. *Human Factors in Electronics, IRE Transactions on*, (1):4–11, 1960. 2.2

Yuan-Pin Lin, Chi-Hong Wang, Tzyy-Ping Jung, Tien-Lin Wu, Shyh-Kang Jeng, Jeng-Ren Duann, and Jyh-Horng Chen. Eeg-based emotion recognition in music listening. *Biomedical Engineering, IEEE Transactions on*, 57(7):1798–1806, 2010. 6.2.2

Donald B Lindsley and JD Wicke. The electroencephalogram: Autonomous electrical activity in man and animals. *Bioelectric recording techniques*, 1(part B):3–83, 1974. 6.2.2

Hod Lipson and Melba Kurman. Factory at home: The emerging economy of personal fabrication. *A report commissioned by the US Office of Science and Technology Policy*, 2010. 7.2

Steven R Livingstone, Ralf Mühlberger, Andrew R Brown, and Andrew Loch. Controlling musical emotionality: an affective computational architecture for influencing musical emotions. *Digital Creativity*, 18(1):43–53, 2007. (document), 6.2.2, 6.1

Michael Lyons and Sidney Fels. Creating new interfaces for musical expression. In *SIGGRAPH Asia 2013 Courses*, page 17. ACM, 2013. 5.2.1

Regan L Mandryk and M Stella Atkins. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International journal of human-computer studies*, 65(4):329–347, 2007. 2.3.2

Steve Mann, James Fung, and Ariel Garten. Deconcert: bathing in the light, sound, and waters of the musical brainbaths. *Proc. ICMC 07*, 2007. 2.6.3

Adnan Marquez-Borbon, Michael Gurevich, A Cavan Fyans, and Paul Stapleton. Designing digital musical interactions in experimental contexts. *contexts*, 16(20):21, 2011. 5.2.2

Karl A McCreadie, Damien H Coyle, and Girijesh Prasad. Learning to modulate sensorimotor rhythms with stereo auditory feedback for a brain-computer interface. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 6711–6714. IEEE, 2012. 5

Karl A McCreadie, Damien H Coyle, and Girijesh Prasad. Sensorimotor learning with stereo auditory feedback for a brain–computer interface. *Medical & biological engineering & computing*, 51(3):285–293, 2013. 2.6.1, 5, 7

Robert E McGrath, Johann Rischau, and Alan B Craig. Transforming creativity: Personalized manufacturing meets embodied computing. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, 4(2):157–173, 2012. 7.3

Sebastián Mealla, Mathieu Bosi, Sergi Jordà, and Aleksander Valjamae. Sonification of brain and body signals in collaborative tasks using a tabletop musical interface. 2011a. 8.2

Sebastián Mealla, Aleksander Väljamäe, Mathieu Bosi, and Sergi Jordà. Listening to your brain: Implicit interaction in collaborative music performances. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 149–154, Oslo, Norway, 2011b. 8.2

Sebastián Mealla, Aluizio Oliveira, Xavier Marimon, Tony Steffert, Sergi Jordà, and Aleksander Väljamäe. The role of personalization and multiple eeg and sound features selection in real time sonification for neurofeedback. In *International Conference on Physiological Computing Systems (PhyCS 2014)*, pages 323–330, Lisbon, Portugal, 2014. SciTePress. 8.3

Sebastián Mealla, Sergi Jordà, and Aleksander Väljamäe. Physiopucks: increasing user motivation by combining tangible and implicit physiological interaction. *Transactions on Computer-Human Interaction (TOCHI)*, in press, 2016. 8.2

Sebastián Mealla, Sergi Jordà, and Aleksander Väljamäe. Implementation, personalization and evaluation of sonic designs for neurofeedback. *Frontiers in Neuroscience, Special Topic on Sonification, Perceptualizing Biological Information*, in review. 8.3

Eduardo Reck Miranda. Brain-computer music interface for composition and performance. *International Journal on Disability and Human Development*, 5(2):119–126, 2006. (document), 2.5.2, 2.16, 2.6.3, 4, 5, 6

Eduardo Reck Miranda and Marcelo M Wanderley. *New digital musical instruments: control and interaction beyond the keyboard*, volume 21. AR Editions, Inc., 2006. 1.2.2, 2.6.1

Melody M Moore. Real-world applications for brain-computer interface technology. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):162–165, 2003. 2.5.2

Th Mulder. Motor imagery and action observation: cognitive tools for rehabilitation. *Journal of neural transmission*, 114(10):1265–1278, 2007. 2.6

Lennart Erik Nacke, Michael Kalyn, Calvin Lough, and Regan Lee Mandryk. Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 103–112. ACM, 2011. 1.2.1, 1.2.1, 2.3.2, 6.2.3, 6.5.5, 8.5.1

Anton Nijholt. Bci for games: A state of the art survey. In *Entertainment Computing-ICEC 2008*, pages 225–228. Springer, 2009. 1.2.1

Anton Nijholt, Danny Plass-Oude Bos, and Boris Reuderink. Turning shortcomings into challenges: Brain–computer interfaces for games. *Entertainment Computing*, 1(2): 85–94, 2009. 1.2.1, 6.2.3

Jesus Olivan, Bob Kemp, and Marco Roessen. Easy listening to sleep recordings: tools and examples. *Sleep medicine*, 5(6):601–603, 2004. 2.6.1

Sile O'Modhrain. A framework for the evaluation of digital musical instruments. *Computer Music Journal*, 35(1):28–42, 2011. 5.2.2, 1, 5.3.2, 5.3.3, 6.5.2

Sandra Pauletto and Andy Hunt. A comparision of audio & visual analysis of complex time-series data sets. 2005. 1.2.2, 3.2.5

Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1285–1294. ACM, 2008. 3.5

Panagiotis C Petrantonakis and Leontios J Hadjileontiadis. Emotion recognition from eeg using higher order crossings. *Information Technology in Biomedicine, IEEE Transactions on*, 14(2):186–197, 2010. 6.2.2, 6.3.3

Rosalind W Picard and Roalind Picard. *Affective computing*, volume 252. MIT press Cambridge, 1997. 2.3.1, 2.3.2

Alan T Pope, Edward H Bogart, and Debbie S Bartolome. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1): 187–195, 1995. 1.1, 1.2.1, 2.3.2, 2.3.2, 2.3.2, 2.6.1

Cristhian Potes, Aysegul Gunduz, Peter Brunner, and Gerwin Schalk. Dynamics of electrocorticographic (ecog) activity in human temporal and frontal cortical areas during music listening. *Neuroimage*, 61(4):841–848, 2012. 2.6.2

Jeff Pressing. Cybernetic issues in interactive performance systems. *Computer music journal*, 14(1):12–25, 1990. 5.3.2

Miller Puckette et al. Pure data: another integrated computer music environment. *Proceedings of the Second Intercollege Computer Music Concerts*, pages 37–41, 1996. 3.2.4, 4.2.1

Sheizaf Rafaeli. Interactivity: From new media to communication. *Sage annual review of communication research: Advancing communication science*, 16:110–134, 1988. 5.3.2

Rafael Ramirez and Zacharias Vamvakousis. Detecting emotion from eeg signals using the emotive epoc device. In *Brain Informatics*, pages 175–184. Springer, 2012. 6.2.2, 6.3.3

Nimesha Ranasinghe, Kasun Karunanayaka, Adrian David Cheok, Owen Noel Newton Fernando, Hideaki Nii, and Ponnampalam Gopalakrishnakone. Digital taste and smell communication. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 78–84. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011. 1

Pramila Rani, Nilanjan Sarkar, and Changchun Liu. Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th International Conference on Human Computer Interaction*, volume 58, 2005. 1.2.1

Davide Rocchesso, Stefania Serafin, Frauke Behrendt, Nicola Bernardini, Roberto Bresin, Gerhard Eckel, Karmen Franinovic, Thomas Hermann, Sandra Pauletto, Patrick Susini, et al. Sonic interaction design: sound, information and experience. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems*, pages 3969–3972. ACM, 2008. 2.4, 2.4.1

A Rodríguez, B Rey, and M Alcañiz. Validation of a low-cost eeg device for mood induction studies. *Studies in health technology and informatics*, 191:43–47, 2013. 4.2.1

D. Rosenboom. Extended musical interface with the human nervous system. *Leonardo Monograph Series International Society for the Arts, Sciences and Technology (ISAST)*, (1), 1997a. 1.2.1, 3, 1.2.2, 2.5.1

D Rosenboom. Extended musical interface with the human nervous system: Assessment and prospectus: Leonardo monograph number 1, 1997b. 2.13, 2.6

James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980. 6.2.2

Tomasz M Rutkowski, Francois Vialatte, Andrzej Cichocki, Danilo P Mandic, and Allan Kardec Barros. Auditory feedback for brain computer interface management–an eeg data sonification approach. In *Knowledge-Based Intelligent Information and Engineering Systems*, pages 1232–1239. Springer, 2006. 5, 6

Joel Ryan. Some remarks on musical instrument design at steim. *Contemporary music review*, 6(1):3–17, 1991. 5.3.2

Javier San Agustin, Henrik Skovsgaard, John Paulin Hansen, and Dan Witzner Hansen. Low-cost gaze interaction: ready to deliver the promises. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 4453–4458. ACM, 2009. 1.1

Jocelyn Scheirer, Raul Fernandez, Jonathan Klein, and Rosalind W Picard. Frustrating the user on purpose: a step toward building an affective computer. *Interacting with computers*, 14(2):93–118, 2002. 2.3.2

Klaus R Scherer, Klaus R Scherer, and Paul Ekman. On the nature and function of emotion: A component process approach. *Approaches to emotion*, 2293:317, 1984. 6.2.2

Timothy Schmele and Imanol Gomez. Exploring 3d audio for brain sonification. 2012. 2.6.2

Albrecht Schmidt. Implicit human computer interaction through context. *Personal technologies*, 4(2-3):191–199, 2000. 1.1, 2.2, 2.2.4

Emery Schubert. Measurement and time series analysis of emotion in music. 1999. 6.2.2

Xavier Serra, M Leman, and G Widmer. A roadmap for sound and music computing. *The S2S Consortium*, pages 1–167, 2007. 2.4

Pradeep Shenoy, Kai J Miller, Beau Crawford, and Rajesh PN Rao. Online electromyographic control of a robotic prosthesis. *Biomedical Engineering, IEEE Transactions on*, 55(3):1128–1135, 2008. 6.2.3

Barbara G Shinn-Cunningham. Object-based auditory and visual attention. *Trends in cognitive sciences*, 12(5):182–186, 2008. 4.4

Hugo Silva. The biosignal c.a.o.s.: Reflections on the usability of physiological sensing for human-computer interaction practitioners and researchers. Technical report, PLUX - Wireless Biosignals, 2015. 2.3.1

Michael Siniatchkin, Peter Kropp, and Wolf-Dieter Gerber. Neurofeedback. the significance of reinforcement and the search for an appropriate strategy for the success of self-regulation. *Applied psychophysiology and biofeedback*, 25(3):167–175, 2000. 4.2.1

Dag Svanaes. *Philosophy of Interaction*. The Interaction Design Foundation, Aarhus, Denmark, 2013. 5.3.2

Koray Tahiroğlu. An exploration on mobile interfaces with adaptive mapping strategies in pure data. In *Proceedings of the 4th International Pure Data Convention*. Bauhaus Universität, Weimar, 2011. 5.3.2

Ana Tajadura-Jiménez, Aleksander Väljamäe, and Daniel Västfjäll. Self-representation in mediated environments: the experience of emotions modulated by auditory-vibrotactile heartbeat. *CyberPsychology & Behavior*, 11(1):33–38, 2008. 2.6.4, 4.4

Kazuhiko Takahashi. Remarks on emotion recognition from multi-modal bio-potential signals. In *Industrial Technology, 2004. IEEE ICIT'04. 2004 IEEE International Conference on*, volume 3, pages 1138–1143. IEEE, 2004. 6.2.2, 6.3.3

Desney Tan and Anton Nijholt. Brain-computer interfaces and human-computer interaction. In *Brain-Computer Interfaces*, pages 3–19. Springer, 2010. 1.1

Atau Tanaka. Musical performance practice on sensor-based instruments. *Trends in Gestural Control of Music*, 13(389-405):284, 2000. 1.2.2, 2.3.1, 2.5.2

Richard Teitelbaum. In tune: Some early experiments in biofeedback music. *Biofeedback and the arts, results of early experiments. Aesthetic Research Centre of Canada Publications*, 1976. 2.5.1

David Tennenhouse. Proactive computing. *Communications of the ACM*, 43(5):43–50, 2000. 2.2.3, 2.2

David L Tennenhouse, Jonathan M Smith, W David Sincoskie, David J Wetherall, and

Gary J Minden. A survey of active network research. *Communications Magazine, IEEE*, 35(1):80–86, 1997. 1.1, 2.2

William Forde Thompson and Brent Robitaille. Can composers express emotions through music? *Empirical Studies of the Arts*, 10(1):79–89, 1992. 4.2.1

Godfried T Toussaint et al. The euclidean algorithm generates traditional musical rhythms. In *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, pages 47–56, 2005. 3

Erin Treacy Solovey, Daniel Afergan, Evan M Peck, Samuel W Hincks, and Robert JK Jacob. Designing implicit interfaces for physiological computing: Guidelines and lessons learned using fnirs. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(6):35, 2015. 1.2.1

Adrian Attard Trevisan and Lewis Jones. Brain music system: Brain music therapy based on real-time sonified brain signals. In *Proc. IET Seminar on Assisted Living*, 2011. 4

Stefano Valenzi, Tanvir Islam, Peter Jurica, and Andrzej Cichocki. Individual classification of emotions using eeg. *Journal of Biomedical Science and Engineering*, 2014, 2014. 6.2.2

A Väljamäe, T Steffert, S Holland, X Marimon, R Benitez, S Mealla, A Oliveira, and S Jordà. A review of real-time eeg sonification research. 2013a. 1.1, 1.2.2, 8.1

Aleksander Väljamäe, Sebastián. Mealla, Tony Steffert, S. Holland, Xavier Marimon, Aluizio. Oliveira, Raul Benitez, and Sergi Jordà. A review of real-time eeg sonification research. In *The 19th International Conference on Auditory Display (ICAD-2013)*, Lodz, Poland, 2013b. 2.3.1, 2.6, 2.17, 2.18, 2.20, 2.21, 2.22, 2.23, 2.24, 2.25

Zacharias Vamvakousis and Rafael Ramirez. P300 harmonies: A brain-computer musical interface. *ICMC/ SMC/ 2014*, 2014. 2.5.2, 2.6.1, 4.2.1

Marjolein D van der Zwaag, Joyce HDM Westerink, and Egon L van den Broek. Emotional and psychophysiological responses to tempo, mode, and percussiveness. *Musicae Scientiae*, 15(2):250–269, 2011. 4.2.1

Rik van Dinteren, Martijn Arns, Marijtje LA Jongsma, and Roy PC Kessels. P300 development across the lifespan: a systematic review and meta-analysis. *PloS one*, 9 (2):e87347, 2014. 2.5.2

Daniel Västfjäll. Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, 5(1 suppl):173–211, 2002. 4.2.2

Daniel Västfjäll. The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. *CyberPsychology & Behavior*, 6(2):181–188, 2003. 4.3.1

Bill Verplank, Craig Sapp, and Max Mathews. A course on controllers. In *Proceedings of the 2001 conference on New interfaces for musical expression*, pages 1–4. National University of Singapore, 2001. 5.2.1

François B Vialatte and Andrzej Cichocki. Sparse bump sonification: a new tool for multichannel eeg diagnosis of mental disorders; application to the detection of the early stage of alzheimer's disease. In *Neural Information Processing*, pages 92–101. Springer, 2006. 5, 2.6.4

Jonna K Vuoskoski and Tuomas Eerola. Measuring music-induced emotion a comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, 15(2):159–173, 2011. 4.2.2

Isabella C Wagner, Ian Daly, and Aleksander Väljamäe. Non-visual and multisensory bci systems: Present and future. In *Towards Practical Brain-Computer Interfaces*, pages 375–393. Springer, 2013. 18

Marcelo Mortensen Wanderley and Nicola Orio. Evaluation of input devices for musical expression: Borrowing tools from hci. *Computer Music Journal*, 26(3):62–76, 2002. 5.2.2

Gil Weinberg and Travis Thatcher. Interactive sonification of neural activity. In *Proceedings of the 2006 conference on New interfaces for musical expression*, pages 246–249. IRCAM Centre Pompidou, 2006. 2.6.2

Mark Weiser and John Seely Brown. Designing calm technology. *PowerGrid Journal*, 1 (1):75–85, 1996. 1.1, 2.2.2

David Wessel and Matthew Wright. Problems and prospects for intimate musical control of computers. *Computer Music Journal*, 26(3):11–22, 2002. 3.2.5

Robert E Wheeler, Richard J Davidson, and Andrew J Tomarken. Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style. *Psychophysiology*, 30(1):82–89, 1993. 3.2.5

Christopher D Wickens. Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3(2):159–177, 2002. 1.2.1

Duncan Williams, Alexis Kirke, Eduardo R Miranda, Etienne Roesch, Ian Daly, and Slawomir Nasuto. Investigating affect in algorithmic composition systems. *Psychology of Music*, 1:24, 2014. 4.2.1

Karl DD Willis, Cheng Xu, Kuan-Ju Wu, Golan Levin, and Mark D Gross. Interactive fabrication: new interfaces for digital fabrication. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, pages 69–72. ACM, 2011. 7.2

Glenn F Wilson and Christopher A Russell. Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human factors: the journal of the human factors and ergonomics society*, 49(6):1005–1018, 2007. 2.3.2

Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002a. 2.3.2

J.R. Wolpaw, N Birbaumer, D.J. McFarland, G Pfurtscheller, and T.M. Vaughan. Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6): 767–791, 2002b. ISSN 1388-2457. 1.2.1

Matthew Wright. Open sound control: an enabling technology for musical networking. *Organised Sound*, 10(03):193–200, 2005. 4.2.1

Dan Wu, Chao-Yi Li, and De-Zhong Yao. Scale-free music of the brain. *PloS one*, 4(6): e5915, 2009. 2.5.2, 2.5.2

Dan Wu, Chaoyi Li, Yu Yin, Changzheng Zhou, and Dezhong Yao. Music composition from the brain signal: representing the mental state by music. *Computational intelligence and neuroscience*, 2010:14, 2010. 2.5.2, 2.6.1, 2.6.4

Beste F Yuksel, Daniel Afergan, Evan M Peck, Garth Griffin, Lane Harrison, Nick WB Chen, Remco Chang, and Robert JK Jacob. Implicit brain-computer interaction applied to a novel adaptive musical interface. Technical report, Technical Report TR-2015-01, Department of Computer Science, Tufts University, 2015. www. cs. tufts. edu/Technical-Reports. html, 2015. 3.5, 8.2.1

Francisco Yus. Misunderstandings and explicit/implicit communication. *Pragmatics*, 9 (4):487–518, 1999. 1.1

Thorsten O Zander, Christian Kothe, Sabine Jatzev, and Matti Gaertner. Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-computer interfaces*, pages 181–199. Springer, 2010. 1.2.1

Victor Zappi and Andrew McPherson. Dimensionality and appropriation in digital musical instrument design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 455–460, 2014. 8.4.1

Marcel Zentner, Didier Grandjean, and Klaus R Scherer. Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, 8(4):494, 2008. 4.2.2

List of Publications

References

[1] Sergi Jordà and Sebastián Mealla. A methodological framework for teaching, evaluating and informing nime design with a focus on mapping and expressiveness. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 233–238, London, United Kingdom, June 30 – July 03 2014. Goldsmiths, University of London.

[2] Sebastián. Mealla. Brain tangible interfaces for multimodal rehabilitation. In *Human Computer Confluence Summer School*, Milan, Italy, 2012.

[3] Sebastián Mealla, Mathieu Bosi, Sergi Jordà, and Aleksander Valjamae. Sonification of brain and body signals in collaborative tasks using a tabletop musical interface. In *Proceedings of the International Conference on Auditory Displays (ICAD)*, Budapest, Hungary, 2011.

[4] Sebastián Mealla, Sergi Jordà, and Aleksander Väljamäe. Physiopucks: increasing user motivation by combining tangible and implicit physiological interaction. *Transactions on Computer-Human Interaction (TOCHI)*, in press, 2016.

[5] Sebastián Mealla, Sergi Jordà, and Aleksander Väljamäe. Implementation, personalization and evaluation of sonic designs for neurofeedback. *Frontiers in Neuroscience, Special Topic on Sonification, Perceptualizing Biological Information*, in review.

[6] Sebastián Mealla, Aluizio Oliveira, Xavier Marimon, Tony Steffert, Sergi Jordà, and Aleksander Väljamäe. The role of personalization and multiple eeg and sound features selection in real time sonification for neurofeedback. In *International Conference on Physiological Computing Systems (PhyCS 2014)*, pages 323–330, Lisbon, Portugal, 2014. SciTePress.

[7] Sebastián Mealla, Aleksander Väljamäe, Mathieu. Bosi, and Sergi Jordà. Let me listen to your brain: Physiology-based interaction in collaborative music composition. In *Brain and Body Interfaces Workshop. ACM CHI Conference on Human Factors in Computing Systems*, Vancouver, BC (Canada), 2011.

[8] Sebastián Mealla, Aleksander Väljamäe, Mathieu Bosi, and Sergi Jordà. Listening to your brain: Implicit interaction in collaborative music performances. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 149–154, Oslo, Norway, 2011.

[9] Aleksander Väljamäe, Sebastián. Mealla, Tony Steffert, S. Holland, Xavier Marimon, Aluizio. Oliveira, Raul Benitez, and Sergi Jordà. A review of real-time eeg sonification research. In *The 19th International Conference on Auditory Display (ICAD-2013)*, Lodz, Poland, 2013.