

---

# EXPLOITING & SHARING CONTEXT

## COMPUTER MEDIATED NONVERBAL COMMUNICATION

---

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik  
der Universität Stuttgart und dem Stuttgart Research Centre for  
Simulation Technology (SRC SimTech) zur Erlangung der Würde eines  
Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

vorgelegt von

**ALIREZA SAHAMI SHIRAZI**

aus Shiraz

Erstgutachter:	Prof. Dr. Albrecht Schmidt
Zweitgutachter:	Prof. Dr. Antonio Krüger
Drittgutachter:	Prof. Dr. Anind Dey

Tag der mündlichen Prüfung: 23.07.2014

Institut für Visualisierung und Interaktive Systeme  
der Universität Stuttgart

2014



*In memory of my father*



# ABSTRACT

Humans are social beings and need to communicate and share their emotions. Communication takes place by exchanging not only verbal information but also nonverbal information. With the development of human civilization, communication is undergoing a constant change. The advances in technologies have led to new communication mediums allowing non-colocated persons to communicate and exchange information. Further, the ubiquity of computers, such as mobile phones, has provided the possibility to use computing technologies for different means in various contexts. Users most often carry the devices with themselves and are even emotionally attached to them. Such computer-mediated communication is generally non face-to-face and communicators are in different contexts. While face-to-face communication consists of verbal and nonverbal information, the lack of nonverbal and contextual information in non-face-to-face communication prevents effective communication and may lead to confusion. Therefore, exploiting and sharing contextual information is essential to enhance the communication between non-colocated persons.

This thesis investigates how to exploit physiological and cognitive information to retrieve awareness about users themselves and their contexts as well as sharing such information using nonverbal modalities through computer-mediated communication channels. It discusses how information about certain user's activities can be obtained using brain signals and user's explicit interactions. Further, it explores nonverbal modalities as a communication channel to express and share context and awareness. The research questions are addressed using empirical methods commonly applied in the human-computer interaction research domain.

In the initial step, we explore two sources as means to obtain the user's context and monitor specific activities. We, first, assess brain signals acquired from commercial brain-computer interfaces (BCI) to determine common activities, i.e., reading, listening, and relaxing. We further assess how the user's emotional state correlates with emotional information provided by the BCIs using videos as stimuli. Second, we investigate how only explicit interactions with mobile applications, instead of using any sensor, can be used to determine the user's physical activities. In particular, we explore how the explicit interaction can be utilized to monitor sleeping as one of the prime everyday activities. Monitoring sleep information shows not only one's daily routines but also indicates the physical state. We assess how exchanging information about one's sleep behavior impacts behavior and awareness in communication. We conduct user studies in the controlled setups and in the wild using application stores to obtain findings with high internal and external validity.

In the next step, we investigate sharing context information using nonverbal channels. We explore rhythm-based tones as a nonverbal mean for communication. We assess how melody composition can be used as a way to express and share emotions. Music, in general, can communicate one's state of mind and it is often characterized as the language of emotion. We use short messages on mobile phones, as one of the most popular services on mobile phones at this time, for sharing emotions. Furthermore, we examine how audio previewing of messages can be used to communicate contents and enable awareness. The current notification approaches such as visual cues and audio tones aim at solely informing the receiver that a message arrived without revealing any further information. We propose an algorithm for audio previewing messages in such a way that content and intention of text messages is additionally conveyed. In the final step, we explore iconic interfaces on mobile phones as a nonverbal modality for sharing sentiments and connect non-colocated users. Through a use case, we assess how the sentiments collected via this channel correlates with moments in a real-time while watching TV. We carry out a study with a large number of users to assess this approach in a realistic context.

The experience gained while conducting several studies in the wild using the application stores allowed us to identify challenges and limitations of this methodology. Further, reviewing prior work that used similar approaches enabled us to have a comprehensive overview about advantages and disadvantages of such studies. Based on the findings, we propose best practices how such user studies can be carried out. We discuss aspects and challenges that should be taken into account during designing such user studies.

The contributions of this thesis provide insights into using physiological and cognitive data to determine activities and emotional states of users and obtain context information. We present how explicit interactions with a mobile application can be leveraged to monitor sleeping behavior of users without using any wearable sensor. It further presents that rhythm-based tones and iconic user interfaces, as nonverbal modalities, can be used to share contextual information. We discuss how sharing context information can affect users awareness and connectedness. The practices for research through the applications stores can be used as a guideline for researchers who want to address their research questions through this research methodology.

# ZUSAMMENFASSUNG

Als soziale Wesen haben Menschen das Bedürfnis zu kommunizieren und ihre Gefühle mitzuteilen. Dies geschieht sowohl mittels verbaler als auch nonverbaler Kommunikation. Wie Menschen kommunizieren, ist einem ständigen Wandel unterworfen. Technische Errungenschaften haben Kommunikationsmedien hervorgebracht, die Kommunikation zwischen Personen, die sich nicht am selben Ort befinden, ermöglichen. Durch die Allgegenwärtigkeit von Computern und Mobiltelefonen ist es möglich geworden, zu jedem Zeitpunkt und in beinahe jedem Kontext zu kommunizieren. Heute haben die meisten Menschen Kommunikationsmittel wie Smartphones ständig bei sich. Hierdurch ermöglichte Kommunikation findet in der Regel nicht im direkten Gegenüber statt, und die Kommunizierenden befinden sich in unterschiedlichen Kontexten. Während direkte Kommunikation verbale und nonverbale Informationen beinhaltet, kann das Fehlen von nonverbalen und kontextbezogenen Informationen in der indirekten Kommunikation zu Missverständnissen führen und erfolgreiche Kommunikation verhindern. Deshalb ist es für die Verbesserung der indirekten Kommunikation wesentlich, kontextbezogene Informationen mitzuteilen.

In dieser Arbeit wird untersucht, wie physiologische und kognitive Daten genutzt werden können, um sowohl Erkenntnisse über die Nutzer und ihr jeweiliges Umfeld zu gewinnen, als auch solche Informationen computervermittelt auf nonverbalem Weg zu teilen. Es wird diskutiert, wie aus der expliziten Interaktion des Nutzers mit digitalen Geräten Wissen abgeleitet werden kann. Hierfür werden Hirnströme und explizite Interaktionen des Nutzers herangezogen. Weiterhin wird untersucht, wie Informationen über das Umfeld durch nonverbale Kommunikation ausgedrückt und geteilt werden können. Die Forschungsfragen werden mit den im Bereich der Mensch-Computer-Interaktion üblichen empirischen Methoden beantwortet.

Im ersten Schritt werden Quellen untersucht, um Informationen über das Umfeld des Nutzers zu gewinnen und Aktivitäten zu beobachten. Es werden zunächst Hirnströme ausgewertet, die mit kommerziellen Brain-Computer-Interfaces (BCI) gewonnen wurden, um allgemeine Aktivitäten, wie Lesen, Zuhören und Entspannen, bestimmen zu können. Weiterhin wird untersucht, wie die durch ein BCI gelieferten Informationen mit emotionalen Zuständen beim betrachten von Filmen korrelieren. Zweitens wird erforscht, wie explizite Interaktion mit mobilen Anwendungen verwendet werden kann, um körperliche Aktivitäten des Nutzers zu ermitteln. Im Speziellen wird untersucht, wie diese Informationen verwertet werden können, um Schlaf als eine unserer Hauptaktivitäten zu detektieren. Aus dem aufgezeichneten Schlafverhalten lassen sich nicht nur tägliche Routinen, sondern auch Hinweise auf den physiologischen Zustand ableiten. Es wird untersucht, wie das Teilen von Informationen über das Schlafverhalten die Achtsamkeit

in der Kommunikation beeinflusst. Hierfür werden sowohl kontrollierte Nutzerstudien im Labor als auch Studien ohne direkte Kontrolle der Probanden durchgeführt, um Ergebnisse mit hoher interner und externer Validität zu erhalten.

Im nächsten Schritt wird untersucht, wie Informationen über das Umfeld nonverbal geteilt werden können. Dabei werden rhythmische Melodien als Kommunikationsmittel verwendet. Es wird untersucht wie selbsterstellte Melodien eingesetzt werden können, um Emotionen auszudrücken. Dabei setzt die entwickelte Anwendung auf SMS, einem der meistverbreiteten Dienste auf Mobiltelefonen, auf. Hierauf aufsetzend wird betrachtet, wie eine akustische Vorschau auf die Nachricht eingesetzt werden kann, um Inhalte zu kommunizieren. Die etablierten Möglichkeiten einer Benachrichtigung, wie visuelle und akustische Hinweise, zielen lediglich darauf ab, auf den Empfang einer Nachricht hinzuweisen. Sie geben jedoch keinen Hinweis auf den Inhalt der Nachricht. Um dies zu ermöglichen, wurde ein Verfahren entwickelt, um aus dem Inhalt einer Nachricht eine akustische Vorschau abzuleiten. Abschließend werden Icon-basierte Benutzungsschnittstellen auf Mobiltelefonen als nonverbaler Kommunikationskanal untersucht. Sie werden verwendet, um räumlich getrennte Nutzer zu verbinden, indem die Nutzer in die Lage versetzt werden, ihre Ansichten auszutauschen. Es wird untersucht, wie die gesammelten Ansichten mit den tatsächlichen Ereignissen korrelieren. Das entwickelte System wird durch eine Feldstudie mit einer großen Teilnehmerzahl evaluiert, um den realisierten Ansatz in einem realistischen Kontext zu betrachten.

Die bei den Feldstudien gewonnenen Erfahrungen, ermöglichen es, die Herausforderungen und Einschränkungen der verwendeten Methode zu identifizieren. Aus der Betrachtung verwandter Arbeiten, die diese Methode verwenden, werden die Vor- und Nachteile solcher Feldstudien abgeleitet. Auf Grundlage dieser Erkenntnisse werden Leitlinien für die Durchführung dieser Art von Studien vorgeschlagen. Hierbei werden insbesondere Aspekte, die beim Design solcher Studien berücksichtigt werden sollten, diskutiert.

Ein Kernbeitrag dieser Arbeit sind tiefere Einsichten über die Nutzung physiologischer und kognitiver Daten. Es wird insbesondere gezeigt, wie sich aus diesen Daten Aktivitäten und Emotionen von Nutzern ableiten lassen. Zusätzlich wird dargestellt, wie sich aus ihnen Informationen über das Umfeld gewinnen lassen. Es wird gezeigt, wie die explizite Interaktion mit mobilen Anwendungen genutzt werden kann, um das Schlafverhalten eines Anwenders zu beobachten. Es wird weiter beschrieben, wie Rhythmus-basierte Töne und Icon-basierte Benutzungsschnittstellen als nonverbale Kommunikationskanäle genutzt werden können, um Informationen auszutauschen. Es wird diskutiert, wie der Austausch von Informationen über das Umfeld das Bewusstsein für andere Menschen und die Verbundenheit mit ihnen beeinflusst. Die vorgeschlagenen Vorgehensweisen für die Forschung mit App-Stores können als Richtlinie für Forscher genutzt werden, die ihre Forschungsfragen mit dieser Methode beantworten wollen.



# ACKNOWLEDGMENTS

Even though only my name appears on the cover of this dissertation, a number of outstanding people have directly and indirectly contributed to it. I owe my gratitude to all those people who have supported and collaborated my research. These acknowledgments are meant to point those who I had the opportunity to work with and I sincerely apologize to everyone I missed to mention.

My deepest gratitude is directed to my supervisor, **Albrecht Schmidt**. I have been amazingly fortunate to have an adviser who gave me the freedom to explore on my own, and at the same time provided guidance to recover when my steps faltered. He taught me being an independent thinker, how to question thoughts, and express ideas. I am grateful for his understanding, patience, support, and friendship. I would also like to thank my examining committee members **Antonio Krüger**, **Anind Dey**, **Stefan Funke**, and **Thomas Ertl** for evaluating the thesis and providing valuable comments and suggestions.

I had the pleasure to work in Albrecht's group with a great number of colleagues. **Florian Alt** who I not only shared my office and my room with, but who became one of my best friends. The trips and sightseeing tours he organized were part of the best experiences I had during my PhD. I am thankful to **Niels Henze** for the motivation during different projects we carried out together. Thanks also to **Bastian Pfleging** (for tips on the German railway system), **Paul Holleis** (for introducing me how to write papers), **Dagmar Kern** (for the great Carnival experiences in Cologne), **Elba del Carmen Valderrama Bahamondez**, **Tanja Döring**, **Stefan Schneegaß** (for the rides home and shopping tours), **Christian Winkler** (for the great times at the Ruhr), **Thomas Kubitza** (for the table soccer matches), **Nicole Recksing**, **Tilman Dingler** (for great discussions during trainings at the gym), **Markus Funk**, **Anja Mebus** (for being an awesome neighbor) as well as to all the students that contributed to my work.

I also had the privilege of collaborating and working with amazing researchers. I am thankful to **Michael Rohs**, **Robert Schleicher**, and **Sven Kratz** for the great collaboration during their times at T-Labs in Berlin. Further thanks go to **Jonna Häkkinä** and her former team (**Ari-Heikki Sarjanoja** and **Arto Puikkonen**) as well as **Kristof van Laerhoven** and his team (**Marko Borazio**) for the great collaboration. I am also grateful to **Enrico Rukzio**, **Ed Chi**, **James Clawson** that in one way or the other contributed to the thesis through their useful comments and feedback.

Finally, and most importantly, I would like to thank my wife, **Marjan Ashrafzadeh**. I praise her support, encouragement, quiet patience, and unwavering love. Her tolerance of my endless travels and late nights working is a testament in itself of her unyielding devotion and love. I am also indebted to my mother **Zohreh Seradj** and my father **Habib Sahami Shirazi** (he may rest in peace) allowing me to be as ambitious as I wanted. I am grateful to my brother and my sister, **Mohammadreza** and **Samira**. Without their faith and support I could not have achieved this accomplishment. Thank you very much.

# TABLE OF CONTENTS

<b>List of Figures</b>	<b>xvii</b>
<b>List of Tables</b>	<b>xix</b>
<b>I INTRODUCTION &amp; MOTIVATION</b>	<b>3</b>
<b>1 Introduction</b>	<b>5</b>
1.1 Research Questions . . . . .	7
1.2 Methodology . . . . .	9
1.3 Research Contributions . . . . .	10
1.4 Research Context . . . . .	14
1.5 Dissertation Outline . . . . .	15
<b>II FOUNDATIONS</b>	<b>23</b>
<b>2 Fundamentals</b>	<b>25</b>
2.1 Communication . . . . .	25
2.2 Computer-Mediated Communication . . . . .	28
2.3 Context-awareness . . . . .	29
2.3.1 What is Context? . . . . .	30
2.3.2 Context-aware Computing . . . . .	30
2.3.3 Acquiring Context . . . . .	31
2.4 Research Methodology . . . . .	32
2.4.1 Research Types . . . . .	32
2.4.2 Research Methods . . . . .	33

2.4.3	From Lab Studies to Research in the Field . . . . .	35
-------	---	----

### **III EXPLOIT CONTEXT IMPLICITLY 43**

#### **3 Mental Task Awareness 47**

3.1	Related Work . . . . .	48
3.2	Data Acquisition . . . . .	50
3.2.1	Task set . . . . .	50
3.2.2	Apparatus . . . . .	51
3.2.3	Dataset . . . . .	52
3.3	User Study . . . . .	53
3.3.1	Procedure & Participants . . . . .	53
3.4	Task Classification . . . . .	53
3.4.1	Feature Set . . . . .	54
3.4.2	Results . . . . .	55
3.4.3	Limitations . . . . .	57
3.5	Implication . . . . .	57
3.6	Summary . . . . .	59

#### **4 Video Annotation with Brain Signals 61**

4.1	Related Work . . . . .	63
4.2	Prototype . . . . .	64
4.2.1	Annotation Application . . . . .	65
4.2.2	Annotation Algorithm . . . . .	65
4.3	User Study . . . . .	66
4.3.1	Procedure . . . . .	67
4.3.2	Participants . . . . .	68
4.3.3	Results . . . . .	69
4.4	Implication . . . . .	70
4.5	Summary . . . . .	71

<b>5</b>	<b>Implicit Sleep Monitoring</b>	<b>73</b>
5.1	Related Work . . . . .	75
5.2	Somnometer: a social alarm clock app . . . . .	77
5.2.1	App Functionality . . . . .	79
5.2.2	Implementation . . . . .	81
5.2.3	Monitoring Sleep Behavior . . . . .	82
5.3	Sleeping Data Ground truth . . . . .	83
5.4	Controlled User Study . . . . .	84
5.4.1	Procedure . . . . .	84
5.4.2	Results . . . . .	85
5.4.3	Discussion . . . . .	89
5.5	In-the-wild Evaluation . . . . .	91
5.5.1	Procedure . . . . .	91
5.5.2	Results . . . . .	91
5.5.3	Discussion . . . . .	92
5.6	Implication . . . . .	93
5.7	Summary . . . . .	94
<b>IV</b>	<b>SHARING CONTEXT NONVERBALLY</b>	<b>97</b>
<b>6</b>	<b>Melody Composition for Sharing Emotion</b>	<b>101</b>
6.1	Related Work . . . . .	103
6.2	Design Rationale . . . . .	104
6.2.1	Sender . . . . .	105
6.2.2	Receiver . . . . .	106
6.3	Prototype . . . . .	108
6.3.1	Composer . . . . .	108
6.3.2	Melody Player Application . . . . .	110
6.4	User Study . . . . .	110
6.4.1	Procedure . . . . .	110
6.4.2	Results . . . . .	111
6.5	Implication . . . . .	112
6.6	Summary . . . . .	113

**7 Sonification Conveys Awareness 115**

- 7.1 Related Work . . . . . 117
- 7.2 Assessment of SMS Usage . . . . . 118
  - 7.2.1 Setup . . . . . 118
  - 7.2.2 Results . . . . . 119
- 7.3 Audio Previewing of SMS . . . . . 121
  - 7.3.1 Approach . . . . . 121
  - 7.3.2 Prototype . . . . . 122
  - 7.3.3 Procedure . . . . . 123
  - 7.3.4 Results . . . . . 124
- 7.4 From a Message to a Melody . . . . . 125
  - 7.4.1 Sonority . . . . . 126
  - 7.4.2 Intention of a message . . . . . 127
  - 7.4.3 Algorithm . . . . . 129
- 7.5 Sonification of Message Intention . . . . . 129
  - 7.5.1 Apparatus . . . . . 131
  - 7.5.2 Procedure . . . . . 131
  - 7.5.3 Results . . . . . 131
- 7.6 Sonification of Instant Messages . . . . . 134
  - 7.6.1 Apparatus . . . . . 134
  - 7.6.2 Procedure . . . . . 134
  - 7.6.3 Participants . . . . . 135
  - 7.6.4 Results . . . . . 135
- 7.7 Implication . . . . . 139
- 7.8 Summary . . . . . 140

**8 Sharing Sentiments with Iconic Interfaces 143**

- 8.1 Related Work . . . . . 145
- 8.2 World Cupinion: a mobile app for sharing opinions . . . . . 147
  - 8.2.1 Design Rationale . . . . . 147
  - 8.2.2 Iconic User Interface . . . . . 148
  - 8.2.3 System Architecture . . . . . 149
- 8.3 User Study in-the-Wild . . . . . 150
  - 8.3.1 Procedure . . . . . 150
  - 8.3.2 Results . . . . . 151

8.4	Implication . . . . .	156
8.5	Summary . . . . .	157

**V CONCLUSION 161**

**9 Guideline for Research in the Wild 163**

9.1	Research in the Wild: 10 Steps Practices . . . . .	164
9.2	Limitations & Challenges . . . . .	173

**10 Conclusion 175**

10.1	Summary of Contributions . . . . .	175
10.1.1	Exploiting Awareness . . . . .	178
10.1.2	Sharing Context Nonverbally . . . . .	179
10.1.3	Practices for Research in-the-wild . . . . .	181
10.2	Future Work . . . . .	181
10.2.1	Physiological Information . . . . .	182
10.2.2	Sharing Mental & Emotional Awareness . . . . .	182
10.2.3	Beyond Text Sonification . . . . .	183
10.2.4	Research in the Wild for other Domains . . . . .	183
10.3	Concluding Remarks . . . . .	184

**VI BIBLIOGRAPHY 187**

**Bibliography 189**





# LIST OF FIGURES

1.1	Selected research prototypes . . . . .	10
3.1	Mental task classification results . . . . .	58
4.1	MediaBrain architecture . . . . .	64
4.2	MediaBrain algorithm output . . . . .	67
4.3	User study setup . . . . .	68
4.4	Output of the updated algorithm . . . . .	70
5.1	Somnometer app screenshots (1/3) . . . . .	78
5.2	Somnometer app screenshots (2/3) . . . . .	80
5.3	Somnometer app screenshots (3/3) . . . . .	81
5.4	HedgeHog device . . . . .	83
5.5	Sleep duration vs. rating . . . . .	86
5.6	Number of comments on Facebook posts . . . . .	88
6.1	EmoShare system architecture . . . . .	108
6.2	Web-based melody composer . . . . .	109
7.1	Situations users do not immediately check the SMS . . . . .	119
7.2	Impact of SMS audio previewing on the user's behavior . . . . .	124
7.3	Quint circle and C major pentatonic scale . . . . .	126
7.4	A sample melody in C major and a minor . . . . .	126
7.5	Algorithm for text sonification . . . . .	130
7.6	Musician vs. non-musician . . . . .	132
8.1	World Cup opinion user interfaces . . . . .	148
8.2	Inputs during 4 matches of the round 16 . . . . .	152
8.3	Sentiments shared during the match England vs. Germany . . . . .	155
9.1	Practices for conducting studies in-the-wild . . . . .	165



# LIST OF TABLES

1.1	Summary of research questions . . . . .	8
1.2	Summary of research prototypes . . . . .	12
2.1	Lab research vs. Research via app stores . . . . .	37
3.1	Brain signal acquisition techniques . . . . .	49
3.2	EEG brainwave frequency bands . . . . .	50
3.3	Task set for data acquisition . . . . .	51
3.4	Data acquired from the NeuroSky BrainBand . . . . .	52
3.5	EEG feature list . . . . .	56
3.6	Mental task classification results . . . . .	57
5.1	Somnometer app characteristics . . . . .	79
7.1	Top frequent emoticons in SMS . . . . .	120
7.2	Most frequent starting keywords in SMS . . . . .	120
7.3	Learnability of audio previewing tones . . . . .	125
7.4	Mapping notes of vowels and constants . . . . .	127
7.5	Example for mapping a sample word . . . . .	128
7.6	Mapping emoticons, punctuations, and keywords to chords . . . . .	128
7.7	Musicians vs. non-musicians understandability . . . . .	133
7.8	Comparison of receive-to-read time based on message intention . . . . .	136
7.9	Comparison of receive-to-read time based on weeks . . . . .	137
7.10	Clustering of message types . . . . .	138
8.1	World Cupinion usage . . . . .	153
8.2	World Cupinion frequency of button clicks . . . . .	154
10.1	Overview of contributions to research questions . . . . .	176







# I

## INTRODUCTION & MOTIVATION





# Chapter 1

---

## Introduction

Communication has been one of main factors in the development of society. Communication is the ability of humans to convey information and share their feelings allows them to understand and cooperate with each other. It takes place for different purposes/reasons. Through contact with others and information accumulated, humans try to understand their own identity [205]. Furthermore, communication also assists in collecting information about others. This information enables to learn about others as well as how to communicate with them. It is further used to develop, maintain, and terminate relationships. Through communication, humans manipulate others, gain compliance, and manage interpersonal conflicts. The social need is yet another purpose for communication. While there might be occasions where humans find comfort in solitude, humans are social creatures in most cases. Therefore, communicating is one of the main ways to fulfill the basic social need humans have.

Through speech, humans convey information and share message. With the development of human civilization and the origin of language, communication has been also revolutionized. In addition to speech, symbols and other mediums have been used for communication and exchange of information. Looking at history of communication, the cave drawing was one of early mediums. Stories and messages were described through series of painted symbols. The invention of writing improved the symbol-based communication. Carving on rocks and later, writing on papers allowed humans to send messages to people in different geographical locations and communicate with them asynchronously. Telecommunication is a new type of communication that emerged from this form of exchanging information. Telecommunication aims at exchanging messages

between non-colocated people. Using drumming patterns or smoke signals is one of early forms of telecommunications for transmitting information nonverbally in distance. In contrast to the writing, for example, this type of communication has been mainly used for synchronous communication.

Telecommunication has been developed with advances in technologies. Through electromagnetic waves and electrical signals, new communication mediums have been invented. Telegraph and telephones are early examples of such communication mediums, allowing non-colocated people to communicate in real-time. With the era of computing in the twenty-first century, we realize by looking around us, that computing technologies permeate our everyday life. Ubiquitous computing provides the opportunity to use computing technologies for different means in various situations. Mobile phones, for example, are one of most ubiquitous technologies. Worldwide mobile phone subscriptions grew to almost 6 billion in 2011 [171]. According to Nielsen, half of the mobile subscribers in the US own smartphones as of February 2012 [136]. Most of the times, users carry their smartphones with them at all times and use it for various purposes in different contexts and are emotionally attached to their phone [193]. While mobile phones are mainly meant for communication and telephony services, advances in technologies has evolved them into more sophisticated devices that have various sensors, and run third-party applications for various purposes apart from communication. This type of communication that occurs through the use of electronic devices is defined as computer-mediated communication (CMC). Here, communication is not necessary face-to-face communication in a shared context, but it is also non face-to-face communication through interaction with a device in different contexts. Therefore, nonverbal information is absent, resulting in a decrease of awareness. The lack of nonverbal contextual information denies effective communication [201] and can lead to confusion [44]. Hence, exploiting and sharing contextual information is essential in enhancing the communication between non-colocated communicators.

Looking at the computer-mediated communication, the interaction is not solely between humans but also between the human and computer. In human-computer interaction (HCI), being aware of the context in which the computer is used results in development of context-aware systems. These systems retrieve contextual information and derive awareness to adapt their behavior for various purposes, e.g., providing different services or increase usability. Sensing technologies such as microphone or GPS sensors available on mobile devices can provide information about the surrounding environment. Additionally, the user himself can be used as a source to derive awareness. Users perceive their surrounded environment using various organs and senses. Thus, information obtained from these sources can be utilized to derive the context. Implicit and explicit interactions with the device can be also leveraged to obtain contextual information [177]. Such interactions can be based on different modalities including command line interfaces

and/or graphical user interfaces (GUI). The implicit interaction particularly has this main advantage that it does not add any additional (mental) workload to users.

The next challenge, after exploiting context, is sharing contextual information between users in the computer-mediated communication. In the face-to-face situation the users both preserve the context the communication takes place. Whereas, in the computer-mediated communication the users are located in different contexts. Lack of contextual information from each side can deny effective communications and lead to confusion. The ubiquitous Internet connectivity available on the devices allows easy sharing of information among non-colocated users. However, means are required to share and represent the information. Textual information such as emoticons is an approach to describe emotional awareness. Using nonverbal information such as tones is another way to represent contextual information. This nonverbal information should be used in such a way that it conveys information and exchange awareness between non-colocated users. This thesis aims at obtaining and sharing context information among non-colocated users. It explores the retrieval of contextual information about users using interactions, particularly implicit interactions, occurring between users and computers. Furthermore, it investigates how awareness and contextual information can be nonverbally exchanged and shared between non-colocated users.

## 1.1 Research Questions

With the increase of ubiquitous computing and its use, determining and sharing context and awareness has been the subject of various research areas. In this dissertation we investigate various possibilities to exploit and share context and awareness by leveraging interaction between humans and computers. Several research questions are investigated (Table 1.1).

Initially, we explore how context can be exploited. While users can explicitly specify their context, we are particularly interested in how information about user's context can be obtained implicitly. The main advantage of the implicit approach is zero additional effort required for users. Users focus on their main task while the computer collects information and obtains awareness and context information respectively. Researchers have explored various approaches to acquire this information. Sensor technologies have been mainly used to monitor the surrounded environment and derive contextual information. On one hand, we assess how context information can be obtained by attaching sensors to the user's body and collecting physiological information. We investigate whether it is feasible to determine mental activities, specifically, reading and relaxing based on information collected from the brain through unobtrusive brain-computer interface (BCI)

**Table 1.1:** Summary of Research Questions.

Research Question	No.	Chapter
<b>I. Exploiting Awareness</b>		
Is it possible to implicitly determine a set of users activities, i.e., reading and relaxing, by using only brain signals?	(R1)	Chapter 3
How does information acquired from brain signals correlate with emotional states of the user by using videos as stimuli?	(R2)	Chapter 4
How is it possible to monitor user's sleep activity based on only explicit interaction with a mobile application without using any sensor?	(R3)	Chapter 5
<b>II. Sharing Awareness</b>		
Is it feasible for users to express and share their emotion through the composition of a melody as a nonverbal mean for communication?	(R4)	Chapter 6
How can audio previewing of a text message convey and share its intention nonverbally?	(R5)	Chapter 7
How can iconic interfaces be used to share sentiments nonverbally in real-time?	(R6)	Chapter 8

systems (R1). Determining cognitive processes is a source for retrieving contextual information. However, certain mental activities, e.g., sleeping or relaxing cannot be recognized from the outside. We particularly investigate the recognition feasibility of reading and relaxing tasks since they correlate with language skills, communication skills, and health. Recognizing and providing feedback about such activities of a person can help to improve their life. We, further, inspect how this information correlates with the emotional state of users using videos as stimuli (R2). On the other hand, we investigate whether it is possible to reliably exploit user's context by leveraging only interactions with the computer instead of using any sensor. We explore how to monitor a user's sleep duration using only their explicit interactions with a mobile phone app instead of any physical or wearable sensors or devices (R3). As one of the prime activities, sleeping shows the availability and routines of a person which is important for communicating with others.

Due to different users' contexts in computer-mediated communication, sharing context information is essential. While sharing this information verbally and explicitly by the user is one approach, nonverbal modalities can be also used to convey the information. In the second part of the dissertation, we examine two nonverbal modalities for communicating and sharing context information, namely, tones and iconic interfaces. Using tones is one of common approaches to inform users about certain events nonverbally. We assess whether it is feasible for users to express and share their emotion through melody composition (R4). Similar to the craft tradition, the melody composition can be a

powerful way to express and share emotions. Further, we investigate if audio previewing of information retrieved from a text message can convey its intention (R5). We assess how this approach impacts the users' behavior. On the other hand, iconic graphical user interfaces are a familiar way for conveying information nonverbally. We examine whether users are able to express and share their sentiments nonverbally in real-time with other non-colocated TV viewers through the use of an iconic interface (R6). The iconic user interface allows users to quickly and with minimum effort share their sentiments at any moment.

It should be mentioned that other approaches exist to exploit and share context information. However, we particularly explore how mobile phones, as one of the most pervasive computing technologies, can be used as a communication channel for exploiting and sharing contextual information. We leverage sophisticated user interfaces and ubiquitous services available on such devices to obtain and share awareness as well as address research questions.

## 1.2 Methodology

The emergence and ubiquity of new computing technologies have encouraged researchers to investigate various approaches to obtain, utilize, and share different contextual information to provide context-aware services and/or enhance communication between humans as well as between humans and computers. Following the same trend, we inspect different resources to exploit and share contexts in this thesis. To answer the research questions, we followed the user-centered design approach. We designed systems and developed prototypes for conducting user studies and assessing hypotheses. The prototypes were either interactive applications or systems designed to solely collect required information. All research prototypes presented in the context of this dissertation are results of collaboration with colleagues and external researchers. Several undergraduates and student assistants also contributed in the development of prototypes as a part of their work. We refer to the Section 1.3 for an overview on the prototypes.

Different user studies were carried out based on the research questions. Some of user studies were conducted in controlled setups with several users. This allowed us to evaluate the hypotheses in very specific contexts. Further, the findings have high internal validity. To increase the external validity of findings, we used a novel approach to conduct user studies. We moved experiments out of the laboratory and conducted in-the-wild studies using available application stores for mobile phones. Such user studies extend classic lab studies and are carried out in more realistic contexts with a large number of users. Prototypes and systems used in these studies should be robust and able to handle



**Figure 1.1:** Two examples of research prototypes implemented to answer research question: (a) **Somnometer**: a mobile application for obtaining contextual information about the sleep activity by leveraging only explicit interactions with the app instead of using any wearable sensor, (b) **World Cupinion**: a mobile application with an iconic interface for expressing and sharing sentiments nonverbally in real-time

large data collected from a large numbers of users. Privacy is an essential concern should be taken into account. In the Section 2.4.3 provides an overview on differences between these two approaches. Further, we provide a guideline on how such user studies can be carried out to answer research questions in Section 9.

### 1.3 Research Contributions

The contribution of this dissertation can be divided into four main parts, with a focus on human-computer interaction: first, we present how contextual information and awareness can be implicitly obtained from certain resources, i.e., brain signals and users explicit interaction; second, we describe how context information and awareness can be shared and conveyed nonverbally using the tones and iconic user interfaces as nonverbal means; third, we report on the development of a set of research prototypes; fourth, we discuss

the lessons learned from conducting studies in-the-wild and propose a guideline on how such a setup should be carried out.

### *Exploiting Context Implicitly*

In the part III of the dissertation, we investigate implicitly exploiting context information based on two resources: (1) *information obtained from the brain signals* and (2) *explicit user interactions with a smartphone*. We present how brain signals retrieved using commercial brain-computer interfaces (BCI) can be used to exploit mental task activities. We describe the recognition feasibility of reading and relaxing tasks out of other daily activities. We, further, show how information provided by commercial BCIs correlates with the emotional state of users. In a case study, we assess how this information can be leveraged to implicitly annotate videos based on emotional information, i.e., excitement information. We propose an algorithm for extracting highlights based on such information. Lastly, we explore obtaining context information without using any sensors. We present the possibility of monitoring users' sleep duration using solely an application on the mobile phone without using any wearable actigraphy devices. We discuss that tracking, visualizing, and sharing sleep information can be used to impact awareness of users about their sleep behaviors.

### *Sharing Context Nonverbally*

In the part IV of the dissertation, we explore sharing context using nonverbal channels. We use *rhythm-based* tones as nonverbal information for sharing contexts and conveying awareness. We investigate the impact of a self-composed melody as a crafted piece of art for sharing emotion. We show that audio previewing on text messages can communicate message's intention. Further, we discuss how this approach affects users behavior in writing and checking text messages. We present an algorithm for the transformation of text messages in to euphonic melodies in such a way that the message's intention can be communicated without reading it. Finally, we assess how *iconic interfaces* can be used to share opinions nonverbally and connect non-located users together. In a case study, we present how TV viewers can nonverbally share sentiments that represent their emotional reactions in real-time using iconic user interfaces. We discuss how this communication channel impacts the experience of watching TV among non-located viewers.

### *Research Prototypes*

To answer the research questions and evaluate the hypotheses, a set of prototypes was developed. The prototypes were used in the user studies to address the questions.

Tables 1.2 includes the list of research prototypes developed and used. Figure 1.1 depicts screenshots of two research prototypes.

Table 1.2: Summary of Research Prototypes.

<b>Prototype</b>	<b>Description</b>	<b>Chapter</b>
Neuroid	is an Android mobile application that uses the NeuroSky BCI to collect the brain signals. The app establishes a Bluetooth connection to the NeuroSky headset. It uses the Android API provided by the headset to retrieve and record data. The data is stored on the SD card of the phone in the Comma Separated Values (CSV) format. When the app is started, it collects demographic information about the user in the first step. Then, it establishes a connection to the headset and start recording data.	Chapter 3
MediaBrain	is a video annotation application. It is a media player and able to fully control the video events such a playback, pause, stop, etc. The application further establishes a connection with the EPOC, a brain-computer interface, and records users emotional information. An algorithm is developed and integrated in the application that extracts the highlights of a movie based on the emotional information acquired and recorded.	Chapter 4
Somnometer	is a social alarm clock application for Android phones that allows users to collect information about their sleep and monitor their sleep behavior. In addition to the conventional alarm clock features, users can define their sleep status, i.e., "gone to bed" or "awake", and can also rate their sleep quality. Users can further share this information with their social network. While the quality of sleep is manually obtained from the users, the app estimates the duration based on tracking user's explicit interactions with the app instead of using any sensor (Figure 1.1(a)).	Chapter 5
EmosShare	consists of two components: the Composer and the Music Player Application. The composer is a web-based melody composer that allows users to create a melody of 32-quarter notes. Users can compose a melody by selecting notes individually or smoothly moving the mouse cursor on single notes. After finishing the composition, the melody is encoded in Midi format and sent as an SMS to a mobile phone. On the mobile phone, a music player application monitors incoming messages. On the receive of messages containing notes, the notes are extracted, a melody is generated and directly played.	Chapter 6

*Continued on next page ...*



Table 1.2 – . . . Continued from previous page

Prototype	Description	Chapter
EmoDetector	is a mobile phone application that runs as a background process without having any impact on other functionalities of the phone. The application searches for certain sets of characters in incoming messages and plays a corresponding notes in case of finding a positive match. By playing a corresponding tone, we attempt to audio preview the message and convey its intention without the need to read the message. The character sets are chosen based on analysis of more than three thousands short messages.	Chapter 7
skypeMelody	is a Skype plug-in application that reads incoming text messages, transforms them to MIDI files, and plays back the melodies. An algorithm is developed that creates a melodic representation from arbitrary message strings. It separates a message into sentences based on punctuation marks. For each sentence, it analyzes key strings such as emoticons, keywords, and punctuation marks to extract its intention. Based on the intention, a corresponding pentatonic scale is chosen. Further, each single character of each word is mapped to the corresponding note. Through this sonification approach, the message context is presented and conveyed.	Chapter 7
World Cupinion	is a mobile application that allows users to share their opinions about events happens during a soccer match in real-time. The app aims to connect non-colocated soccer fan viewers. It consists of an iconic user interface for expressing and sharing of opinions nonverbally (Figure 1.1(b)). The iconic interface contains a set of sentiments related to events happen in a soccer match. The sentiments are presented nonverbally using proper icons. The interface lowers the cost of interaction and allows users to share their opinion quickly and in short burst interactions. It further decreases space between active users and lurkers. The feedback visualization conveys the current opinion of users about the ongoing match.	Chapter 8

### *Research in-the-wild Methodology*

In order to answer several research questions in this thesis, we conducted studies in the field with a large number of users using application stores for mobile phones. Based on the experiences gained during the time the studies were conducted as well as the review of prior work that used similar approach, we provide practices on how to conduct such studies. We provide a guideline and highlight the limitations of such studies in comparison with the classic lab studies. We report the challenges faced during these

studies and discuss how they can be addressed. The findings can be very valuable and useful for other researchers who want to use the similar approach to conduct user studies and assess hypotheses.

## 1.4 Research Context

The research leading to this dissertation was carried out over the course of five years at the *University of Duisburg-Essen* (User Interface Engineering Group) and at the *University of Stuttgart* (Human-Computer Interaction Group). Different cooperations with experts in the context of various projects resulted in publications that contributed in this thesis.

### **Nokia Research Center**

During the five years, two joint projects were conducted together with Dr. Jonna Häkkinä from the Nokia Research Center in Oulu, Finland. The projects investigated nonverbal means for expressing and sharing awareness between non-located users. In one project, we investigated the use of melody composition as a mean for expressing and conveying emotions. In the second project, sonification of textual information for communicating awareness was assessed. The results of the projects are published in the MobileHCI 2009 [160] and CHI 2010 [168] conferences.

### **Telekom Innovation Laboratories, TU Berlin**

Through a joint project with Prof. Dr. Michael Rohs and Dr. Robert Schleicher from the Telekom Innovation Laboratories at the TU Berlin, we investigated how iconic interfaces can be used for sharing sentiments in real-time using mobile phones. A mobile application was developed and released in Android markets to carry out a large-scale user study and collect data. The results of the cooperation is published in the CHI 2011 conference [167] and the IJMHCI journal in 2011 [174].

### **Technische Universität Darmstadt**

A part of this thesis investigates an approach on implicitly exploiting context information about sleep activity and its effect on social awareness. In the context of a collaboration with Prof. Dr. Kristof Van Laerhoven, TU Darmstadt, an expert in wearable computing, we used the HedgeHoge sensor developed by his group for monitoring sleep behavior. Furthermore, we closely cooperated with Dr. James Clawson from Georgia Institute of

Technology and Dr. Ed H. Chi from Google. The result of the collaboration is published in the IJHCS journal in 2013 [161].

## 1.5 Dissertation Outline

This thesis consists of ten chapters, which are distributed into five parts. The first part of the thesis includes the motivation behind the research described in this dissertation. It introduces the research questions and the methodologies used to address them. It also provides an overview on research contributions and the context in which research was carried out. Part II reviews the fundamentals considered in context of this thesis. Then, it is followed by the two main parts of the thesis. The *Exploit Context Implicitly* part (Part III) explores possible approaches to implicitly obtain context information about mental activities as well as sleeping behavior. Part IV, *Sharing Context Nonverbally*, focuses on nonverbal means to convey and share awareness and context between non-colocated users. The final part (Part V) introduces a guideline on conducting research in-the-wild. It further contains a summary of the research contributions and future work. Related work is integrated individually into the chapters.

### Part II: Background

**Chapter 2 – Fundamentals:** This chapter provides an in-depth introduction to the foundations. It starts with components involved in communication. Further, we look into computer-mediated communication and compare it with face-to-face communication. We discuss implicit and explicit interaction with such a medium. To identify contextual information, we look at definitions and review prior work to understand what context is and how it can be acquired. Finally, we review research methodologies to gain a better understanding of how research questions can be addressed and hypotheses can be examined. We compare classic lab studies conducted in controlled settings with studies conducted through deployment in-the-wild with a large number of users to identify advantages and disadvantages of each methodology.

## Part II: Exploit Context Implicitly

**Chapter 3 – Mental Task Awareness:** Determining the user’s context is central for ubiquitous computing. Being able to determine what a user is doing enables numerous use cases. Context are determined by either equipping the environment with sensors (e.g., cameras) or attaching sensors to the user’s body (e.g., accelerometers and gyroscopes). Certain mental activities, however, cannot easily be recognized from the outside. Differentiating between sleeping, relaxing, listening to music, or thinking hard about a problem can look exactly the same from the outside. However, the cognitive processes assigned to these tasks differ. In this chapter, we explore the feasibility of using commercial Brain-Computer Interfaces (BCI) to obtain contextual information. We are particularly interested in contextual information related to learning, especially reading and relaxing. The amount we read directly influences the size of our vocabulary, language skills, and general knowledge. In addition, regular relaxation, breaks and naps are correlated with more effective skill acquisition and learning. Through a user study, we assess and discuss the feasibility of exploiting mental awareness using a commercial BCI.

**Chapter 4 – Video Annotation with Brain Signals:** The advances in signal processing have enabled commercial BCI headsets to retrieve and provide certain information such as emotional state (excitement, frustration, etc.) in real-time. This allows the usage of this information in real-time and the development of context-aware systems. However, the question is, how does this information correlates with user’s state. In this chapter, we investigate the correlation between the emotional information provided by the commercial BCIs and emotional information that users explicitly provide using videos as stimuli. We assess the feasibility of implicitly annotating videos based on the nonverbal information obtained using the BCI. Brain signals can reveal different information such as facial expressions or the level of excitement. It can further reveal different information that correlates with scenes users watch in a video. This information can be used for annotating a video and generating a summary. Adding annotations to time segments on a video timeline makes it easier to search, find, and playback important segments of the video. We present an annotation tool that allows implicit annotation of videos based on information acquired from BCIs. We, further, propose an algorithm that can be used to extract highlights based on the excitement information obtained using a commercial BCI.

**Chapter 5 – Implicit Sleep Monitoring:** The proliferation of mobile devices in everyday life has led to an increasing amount of information about users' personal contexts. With sensors embedded on the smart phones, different contextual information about the user's activities and the surrounding environment can be obtained. Implicit and explicit interactions with applications, in contrast, are other sources for retrieval of contextual information instead of using any sensor. In this chapter of the dissertation, we aim to investigate how solely using explicit interactions with a mobile application for implicit monitoring sleep patterns. Sleep information, e.g., whether a person has gone to bed or is awake, shows not only one's daily routines but also indicates physical state, and the availability of a person. On the other hand, sleeping has been identified as one of the prime activities that contributes significantly to the state of an individual's mental and physical health. We explore how monitoring and sharing sleep information as another type of activity that can impact awareness, connectedness, and sleeping behavior. This information can be valuable not only to the person themselves, but also to others.

### Part III: Sharing Context Nonverbally

**Chapter 6 – Melody Composition for Sharing Emotion:** Many interactive technologies designed for other purposes have been adapted for use within intimate relationships. People use numerous techniques and technologies to maintain an emotional connection. Webcams, emails, instant messaging, and blogs are examples of such technologies used as mediators of human interaction for sharing and maintaining emotion. In computer-mediated text-based communication means for explicitly expressing emotions and feelings by abbreviations and symbols have been developed. We, in contrast, are interested in how emotions can be shared using non-textual information in the non face-to-face communication. In this chapter we assess how a melody as nonverbal information can be used for expressing and sharing emotions. Sharing emotions is one of the main purposes of communication. Users are emotionally attached to their phones. Whereas it seems on one hand it is natural to use this device as a mediator for sharing emotions, currently it only states very basic and simple ways of deploying these devices for sharing emotional feelings. Apart from telephony service, the short messaging service is one of the most popular services available on mobile phones. We use this service as a communication channel and leverage its capabilities to easily,

quickly, and cheaply share emotional feeling and create awareness between friends or partners.

**Chapter 7 – Sonification Conveys Awareness:** The emergence and advances in services and applications on mobile phones allow users to communicate through different communication channels such as text messaging, emails, etc. Visual clues or tones are common notification mechanisms used on the mobile phone to make receivers aware of incoming messages. Synchronous communication tools (e.g., chat clients) mainly use visual clues (highlighting the application’s window). In addition, asynchronous communication tools (e.g., email clients) often make use of audio notifications. However, such notifications neither convey the content nor the intention of a message. In this chapter we investigate how such notifications can be provided in such a way that it conveys information based on the content of messages. Sonification is an approach that uses non-speech audios to convey information. It aims to translate relationships or information in data into sounds that exploit the auditory perceptual abilities of human beings such that the data relationships are comprehensible. We explore how the sonification of text messages can be achieved for conveying message’s intention and content nonverbally.

**Chapter 8 – Sharing Sentiments with Iconic Interfaces:** The ubiquity of mobile phones and the Internet connectivity on them provides this possibility to share information among non-colocated users in real-time and connect them together. The goal of the work presented in this chapter is to investigate how iconic interfaces can be used to share sentiments nonverbally among non-colocated users. Through a user case we investigate how an iconic user interface can be used not only for exchanging information that represents emotional reactions to events shown live on TV, but also for visualizing and conveying reactions of other non-colocated users in real-time. Smartphones are indeed used as a second screen for social networking, chatting, and web browsing while watching television. They can serve as standalone platforms for collecting and sharing the user’s emotional responses to TV-related experiences.

## Part IV: Conclusion

**Chapter 9 – Guideline for Research in the Wild:** While studies in laboratories with a controlled setup is one way of conducting evaluations, researchers

have tried to increase the external validity of findings by carrying out studies in more realistic context. Early research indeed shows the importance to conduct in-situ experiments when analyzing mobile usage behavior. The emergence of application stores provides the opportunity to conduct studies in the wild with a large number of users. We also used this approach to answer our research questions. Prior work has highlighted certain challenges of large-scale studies conducted in the wild based on their experiences. They mainly focus on either specific aspects or are limited to the respective authors' experience. What is still missing is a comprehensive overview of the lessons learned from all the different studies conducted in the wild and how research can address the identified challenges. Based on the user studies reported in this dissertation as well as the analysis of related work, we provide practices how to conduct studies through applications stores. We identify challenges and limitations of such studies. The guideline can help other researchers who want to use this methodology to answer their research questions.

**Chapter 10 – Conclusion:** The conclusion summarizes the contributions made in this thesis. It reviews the research questions addressed within this dissertation. It, furthermore, identifies and discusses potential directions for future work.









# II

FOUNDATIONS



# Chapter 2

---

## Fundamentals

In this dissertation, it is investigated how to exploit contextual information and share it in computer-mediated communication. In the initial step, it is essential to understand the foundations of communication. The following chapter of the thesis provides an overview of the foundation in detail. We, first, discuss goals involved in communication as one of humans' main abilities to exchange information. Specifically, we describe components involved in communication. Second, we look into computer-mediated communication. We compare such communication with the face-to-face communication. Moreover, implicit and explicit interactions with such a medium is explored. To identify contextual information, we look at definitions and review related work to understand the *context* and how it can be acquired. Finally, we overview research methodologies to realize how research questions can be addressed and hypotheses can be examined. We compare classic lab studies conducted in controlled settings with studies conducted in-the-field with a large number of users to identify advantages and disadvantages of each setup.

### 2.1 Communication

The meaningful exchange of information, or in other words *communication*, is when information is successfully conveyed. Communication as a human ability to share messages and feelings has an important role in the development of the society. Humans as social creatures need to express and receive love. They desire to socialize and be in

the company of other people. Communicating with others satisfies these needs. Further, humans communicate for different goals. The goals for communication can be divided to three general types: self-presentation, relational, and instrumental goals [29]. Self-presentation goals involve communicating identity and presenting oneself as one wants to be seen by others. Through relational goals, humans develop and maintain relationships. With instrumental goals, humans attempt to shape the behavior of others and as such, their actions involve persuading someone to do something in a certain way. Being aware of the goals for communication is important for aligning interpersonal interactions with specific goals.

Communication consists of a sender (source), a communication channel, and a receiver. Samovar et al. [170] define the human-to-human communication as:

*“a dynamic process in which people attempt to share their thoughts with other people through the use of symbols in particular settings”* [170]

The definition describes that communication has several characteristics. It is a dynamic process indicating it is an ongoing activity. When a word is produced it cannot be retracted. Further, the communication is symbolic. The communication relies on the interpretation of information heard and seen, as direct access to others' thoughts and feelings is almost impossible. The communication is, therefore, contextual. Setting and environment (context) help to determine and to understand the words and actions communicated.

In human-to-human communication three components are mainly involved: (1) Thought/Message: information exists in the sender's mind, (2) Encoding: the message is sent in symbols/words/cues, (3) Decoding: the receiver translates the symbols/words/cue into information. These three components together establish a successful communication. Further, the communication is either synchronous or asynchronous. In the synchronous situation the communication takes place in real-time. Whereas in asynchronous communication, the interaction between the sender and receiver is delayed. Such delays may result in confusion in the communication. On the other hand, communication consists of verbal and nonverbal information, which we discuss in the following part.

### *Verbal & Nonverbal Communication*

Communication, in general, is through exchange of verbal and nonverbal information. The verbal communication refers to the spoken part of communication, or in the other word, speech. It is a medium for communication that entails talk, which uses languages to exchange messages. A language is a system of symbols that are combined and

manipulated using grammar. Grammar, further, is sets of rules that specify how the symbols (words) should be combined together to convey messages. Tone and pitch are other important elements in speech and a language.

Messages, ideas, and opinions are communicated not only through words, but also through nonverbal signals such as gestures and body postures. When humans talk to each other, facial expressions, gaze, eye contact, vocal qualities, body movement, etc. are nonverbal behaviors used to communicate. Darwin is one of the very early scientists who argued that humans reliably show emotions in their faces [49]. In non face-to-face communication, however, vocal qualities (volume, rate, pitch, etc.) and vocal characteristics (laughing, crying, whining, etc.) are predominant. Further, culture and space have an important role also in understanding nonverbal information [75]. Direct eye-to-eye contact, for example, is positive in Western cultures. Whereas avoiding eye contact to show respect is common in Japan. Further, nonverbal communication represents two-thirds of all communication [91]. Samovar et al. [170] define nonverbal communication as follows:

*“nonverbal communication involves all those nonverbal stimuli in a communication setting that are generated by both the source and his or her use of the environment, and that have potential message value for the source and/or receiver”* [170]

The definition includes that nonverbal communication can be both intentional and unintentional. Shaking the head, for example, is an intentional behavior to confirm a statement. Yawning during a conversation is unconscious and may have different meanings for others. Furthermore, the nonverbal messages can serve as substitutes for verbal messages. Verbal and nonverbal messages often work in unison. Separation of verbal and nonverbal behavior into distinct categories is virtually impossible [103].

Both definitions reveal that context, indeed, impacts and adapts verbal and nonverbal communication. It helps to specify an appropriate language as well as how it should be used for communication. In the face-to-face situation the sender and the receiver both perceive context the communication is taken place. However, situational information is not completely clear in a non face-to-face situation. Lack of contextual information can result in confusion during the communication. Sharing contextual information, e.g., location, time, emotion, etc. between the sender and the receiver can enhance the communication.

## 2.2 Computer-Mediated Communication

In face-to-face communication humans analyze nonverbal information in addition to words being spoken in order to interpret the communicator's intention. Lack of such cues can lead to confusion and incoherence [44]. The advances in technologies resulted in new mediums for communication. Further, communication can take place through devices such as (mobile) phones and computers, called computer-mediated communication (CMC). Such communication can be established regardless of sender and receiver locations. However, nonverbal information is partially omitted. The absence of nonverbal contextual cues denies important information and effective communication [201].

The information transmits during communication are divided to *codifiable* and *noncodifiable* information [155]. Codifiable information can be described through symbols. Therefore, it can be transmitted through computer-mediated channels. However, communicating noncodifiable information is a challenge. In video-based and voice-based CMC, for example, more nonverbal information is available and exchanged even though the sender and receiver have different contexts. On the other hand, text-based CMC lacks nonverbal information. Users, hence, adapt their language, style, and other cues to enhance communication. Using punctuations or emoticons (a set of characters represents facial expression of emotion) are approaches that users have adapted.

Specifying and sharing status also conveys one's current situational information. Interactions with such mediums have been used as resources for retrieving contextual information. Therefore, it is essential to get insights on interactions between the human and computer. This allows us to understand which and how information can be obtained based on the interaction between the user and computer.

### *Implicit & Explicit Interaction*

Humans interact with their environment though information being received and sent. They use their senses to collect information as well as the effectors to respond and interact. In an interaction with a computer outputs from users is used as inputs for the computer and vice versa. In traditional computer applications the interaction was purposeful and direct. The results were also explicitly attended. The Norman's execution-evaluation loop, indeed, expresses such interactions [139].

As computers have become more ubiquitous, systems and applications have been developed in which users attention and intention are lower. The user interaction with such systems has been redirected towards being more implicit. Information is gathered not only through explicit user's inputs, but also implicitly through sense-based interactions.



Hence, the human-computer interaction can be divided into *explicit* and *implicit* interactions. In the explicit interaction the users intentionally provided certain information to the computer through various channels such as the command-line or graphical user interface, gesture, etc. They also expect to get certain feedback/results from this explicit interaction; whereas in the implicit interaction the user is not intentionally providing information for interaction. Schmidt defines the implicit human computer interaction as follows:

*“an action performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input” [175]*

As the definition describes, the computer leverages and makes use of the user interaction. Systems can use implicit interactions in addition to explicit interactions. However, facilitating solely implicit interaction is also achievable. The implicit interaction consists of two main steps: perception and interpretation [175]. In the perception a system senses and collects information. Information can be acquired from the user or the context. In Section 2.3 we discuss what context is and how contextual information can be acquired. In the interpretation step the system uses the information obtained. Based on the results, relevant feedback can be provided or proper actions can be triggered. Therefore, it is possible to implicitly exploit context and awareness by leveraging implicit interaction. In contrast to explicit interactions, implicit interactions do not add any additional cost or cognitive load to users.

## 2.3 Context-awareness

As previously mentioned, context impacts communication. Except in face-to-face, in other types of communication being aware of the context requires users to sense and share situational information. In human-to-human communication the sender and the receiver can explicitly specify and share their context. In computer-mediated communication, however, the computer can be used to obtain context implicitly. Such systems that sense and utilize contextual information are called computer-aware systems. To retrieve contextual information and develop context-aware services, we need to understand what context is.

### 2.3.1 What is Context?

The term context is widely used with different meanings. In the Cambridge dictionary, the term context (as a noun) is defined as the situation within which something exists or happens and that can help explain it. Computer scientists have attempted to define and elucidate context [159, 172, 178]. The early definitions refer to context as location, environment, identity, time, etc. Dey and Abowd provide a more collective and generic definition for context:

*“any information that can be used to characterize the situation of an entity. An entity is a person, place, or object, considered relevant to the interaction between a user and an application, including the user and applications themselves.”* [1]

Based on this definition, a piece of information can be used to characterize a situation is context. Location, identity, activity, and time are the prominent information for characterizing a situation. This information answers the questions of who, what, when, and where. Based on this data, other contextual information can be inferred. For example, when a system is able to identify activities of the user, it can further measure the calories they burned daily. This information can be used to assess the user’s physical activities and health status, respectively.

### 2.3.2 Context-aware Computing

In human-computer interaction the contextual information is used to understand who, where, when, and for what interacts with applications. Such information can be utilized in designing a system, called context-aware computing. Schilit et. al. are the first persons introduced the concept of context-aware computing [172]. They describe the term as *“... context-aware software adapts according to the location of use, the collection of nearby people, hosts, and accessible devices, as well as to changes to such things over time. A system with these capabilities can examine the computing environment and react to changes to the environment”*. Based on this description, the context is strongly related to the location of a system. However, Dey and Abowd provide a generic definition for context-aware computing:

*“A system is context-aware if it uses context to provide relevant information and/or services to the users, where relevancy depends on the user’s task”* [1].

They categorize context-aware systems into three different groups. First, applications that either present contextual information or suggest selections of services to the user, e.g., applications that are presented in Chapter 6 and Chapter 7. Second, applications that trigger an action or reconfigure the system on behalf of the user based on in context changes, e.g., the system which is presented in Chapter 5. Finally, systems that annotate captured data with context information(e.g., the application is used in Chapter 4).

For developing context-aware systems, researchers have proposed different models, frameworks, and architectures [52, 178, 173]. All frameworks and models have three steps in common. In the first step, acquiring context information; afterward, interpreting and extracting features; and, finally, utilizing obtained information. While the first step requires sensing and getting inputs from context, the next two steps are concealed within the system design.

Numerous context-aware systems have been developed for various purposes. In the human-computer interaction domain, such systems mainly aim to ease the interaction and increase the usability. The main challenge, however, is how and what information can be obtained from context.

### 2.3.3 Acquiring Context

Context-aware applications require information from the context in which they are used. In a naive approach users can explicitly provide information relevant to a given situation and crucial for the application. However, this approach adds additional load to the user. Alternatively, information can be obtained implicitly by sensing the situation.

Obtaining information from the same system on which the application run is straightforward. On the other hand, collecting information from contexts that lie outside of the system requires other approaches. Two approaches can be used [178]. In the first approach, environments are smart and provide infrastructures for obtaining contextual information. This requires that sensors are embedded in the environment and monitor the surrounding area. Thus, such systems do not work in environments without such infrastructures. Active Badge is one of the first systems that use this approach [202]. Embedding the system (respectively the application) with sensors that enable the acquisition of information is the second approach. This approach does not rely on any infrastructure and can be used in any context, particularly in mobile computing.

Systems are able to use various sensors to sense environment and retrieve information. GPS and light sensors, for example, provide information about a physical location. Accelerometer and gyro sensors reveal the orientation and acceleration of a device. With

advances in sensing technologies, more sophisticated sensors are developed by sensing humans status. Eye trackers monitor human eyes and provide gaze information. Biosensors and brain-computer interfaces supply information about the brain and physiological status of the user. Sensors obtaining information from the physical environment monitor unobtrusively, whereas sensors that monitor a human's body can be obtrusive for users. With available sensors it possible to obtain contextual information available in the environment in which the user is as well as from the user himself. This information can be leveraged in the development of services and applications to increase usability.

## 2.4 Research Methodology

Answering research questions is a main goal of scientific work. Through a hypothesis, researchers describe a phenomenon. However, the hypothesis should be still evaluated to assure its validity. In comparison with a scientific theory, a hypothesis is a more focused statement that can be validated with a single experiment [157]. In the human-computer interaction (HCI) research domain different types of research can be conducted to examine hypotheses. Various techniques are available to conduct such experiments. However, these methodologies are not exclusive for the HCI domain and are applicable to other research domains. In this section we provide an overview on different research methodologies for conducting studies and examining hypotheses. The hypotheses proposed in this dissertation are assessed by conducting different types of research. Different methods are used for carrying out the user studies.

### 2.4.1 Research Types

After hypothesizing a phenomena or specifying research questions, investigations should be conducted to examine the phenomena and answer the questions. Rosenthal and Rosnow categorize investigations into three groups: descriptive, relational, and experimental [157].

#### *Descriptive Research*

The descriptive (observational) investigation aims to construct an accurate description of what is happening. In such research, no hypothesis is examined and no variable is manipulated. Through unobtrusive approaches and without interfering with it, a phenomenon in the real word is observed and information is collected. The description can be qualitative or quantitative. Interviews, observations, and focus groups are methods

for conducting these types of studies. The descriptive approach provides neither the relationship between factors nor insights for why this happens.

### *Relational Research*

Relational (correlation) research focuses on the identification of relations between two or more variables. However, it can rarely determine the causality between factors [42]. Hence, it is unknown which variables cause changes in others or if the changes are due to hidden factors that have not been considered and examined. Statistical correlation analyses can be run to assess relations between factors.

### *Experimental Research*

An experimental (causal) study investigates in order to determine relations between multiple variables as well as causality. The causality describes how variables influence each other. Experimental studies are based on hypotheses. The hypotheses are examined in at least two conditions. In each condition variables are measured quantitatively. Through statistical analyses measurements are tested and compared. The results are used to assess the hypotheses. Such investigations can be conducted in the lab with a more controlled setup or in the field with higher ecological validity. Randomization is essential for elimination of biases. Strict control of factors influencing dependent variables is required.

These three research types can be used in different phases of a research program. In the early phase of a research program, the descriptive research is often the first step for identifying interesting phenomena, allowing the researcher to specify future research directions. Relational studies can be used to identify relations between variables. Further, fundamental causal relations can be explored through experimental research. Independent of research type, ethical considerations are essential and should be taken into account. Participants' privacy should be definitely preserved.

## 2.4.2 Research Methods

After choosing an appropriate research type for answering research questions, the next step is carrying out the user study. Various methods are available to conduct an investigation and collect required information. In the following part we provide a short overview on methods used in projects and user studies presented in this thesis.

### *Interviews & Focus groups*

Interviews and focus groups are methods for collecting qualitative data. These methods allow direct discussions with participants. The discussion's structure can be anywhere from free form to semi- and fully-structured discussions. An interview is conducted with one single participant. It allows exploration of a wide range of concerns about a problem. It further gives interviewees the freedom to individually provide detailed responses. Whereas, the focus group is with several individuals in a group. Such a group discussion allows the researchers to easily and effectively gathering a broad range of opinions. Robson suggests between eight and twelve participants for a focus group [156]. Conflicts, disagreements, and debates in a focus group might also reveal new areas for further assessment [24]. Both methods are powerful for understanding user's opinions, concerns, and views.

### *Questionnaire*

A questionnaire (survey) is a fully structured and well-defined set of questions to which an individual is asked to respond [107]. The questions can be defined in such ways that users respond qualitatively (using free texts) or quantitatively (e.g., using Likert scales [112]). By using a survey, it is easily possible to get a large number of responses from users geographically dispersed. Hence, it allows researchers to get a quick overview of the users. Standard questionnaires are also available to assess system perspectives. NasaTLX (NASA Task Load Index) is, for example, a standard questionnaire for assessment of user's perceived workload [81]. The SUS (System Usability Scale) questionnaire is another questionnaire for examining system's usability [23].

### *Observation*

The observation method requires researchers to observe users behavior in a realistic context. The observation can be with or without intervention. Without intervention the observation is more useful as individuals are observed in natural settings. Hence, the findings are ecologically valid. However, observations with some components of intervention may be essential in some situations due to privacy and/or ethic concerns. The observation can be done manually or automated. In manual observations, observers collect data. For example, data collection can be fulfilled by writing notes, taking photos, or recording video from users behavior. For automated observations, a system should be used to automatically collect data. From example, Sahami Shirazi et al. use a script to automatically observe interaction behavior on users' social networks, i.e., Facebook and collect information [161].

### *Data Logging*

In this method data required to answer research questions is logged. A data collection system is usually used for recording data. Furthermore, the system collects data automatically and implicitly. Hence, it does not add any additional (mental) cost to users. Task completion time or activity logging, for examples, can be monitored and collected through this approach. Based on the data required, additional hardware might be used. Eye trackers or brain-computer interfaces, for example, can be used to obtain information from the eyes [6] and the brain [162]. This method is particularly helpful when conducted over a longer period of time. Choosing the appropriate data granularity and proper data management are crucial components of any automated data collection system [107].

Each method discussed has its own advantages and disadvantages. Based on the research type, researchers can choose one or more methods to conduct a user study. It should be mentioned that the methods discussed above are not all available methods. Diaries and ethnography, for example, are other methods to conduct a user study [107].

### 2.4.3 From Lab Studies to Research in the Field

Conducting studies in laboratories with a controlled setup and a specific population is a common approach for answering research questions. Such a setup allows researchers to evaluate hypotheses in a very specific context. External influences are minimized. Findings have high internal validity. However, simulation of the realistic context is not always achievable. Researchers have tried to increase the external validity of findings by conducting studies in a more realistic context and with a larger number of users.

Studying human behavior with a sufficiently large and representative sample has been an apparent challenge. In a first attempt to find a solution, living labs have been set up to simulate more realistic contexts and open up more study opportunities [3, 137]. But living labs, nevertheless, are essentially controlled settings and studying the usage of mobile devices and home technologies in their natural habitat is not really possible. Further, collecting input from only a small set of participants can be problematic in many design situations. A comprehensive evaluation would be to assess hypotheses with a diverse set of users in all possible contexts the devices might be used.

Another approach to conduct studies with a fairly large number of participants is using crowdsourcing Internet marketplace, such as Amazon Mechanical Turk [102]. Participants complete tasks for a monetary or a non-monetary (e.g., reputation) reward [82, 102, 181]. Using Mechanical Turk researchers can submit tasks that are completed

by Mechanical Turk workers, so called ‘Turkers’, that receive money in return. However, it has been shown that the Turker population shifted to be mostly from India with a low income [158]. This results in studies with highly biased samples.

Researchers have started to explore a similar direction and conduct large-scale studies. They started to leverage mobile application stores, such as Apple’s App Store or Google Play, to organize human subject studies. Mobile applications and games are designed to observe the users’ behavior and are published in application stores. In contrast to commercial apps, the apparatus is specially designed to answer research questions. As the apparatus is freely available, it can be installed and used by thousands of users.

First attempts simply published their existing research prototypes in available stores to showcase novel ideas and use the large number of users as a validation of the work [212]. Researchers also investigated how the large-scale deployment of mobile applications can be used as a research tool. McMillan et al., for example, used the game *Hungry Yoshi* to investigate how to collect subjective feedback [118] and Pielot et al. also used a game to compare different approaches to ask participants for consent [152]. Another direction investigates aspects that are mainly relevant for a particular type of applications. Girardello and Michahelles, for example, studied how users rate mobile applications by publishing *AppAware*, a recommender app for mobile applications [67]. Finally, mobile application stores have been used to distribute apps that observe more general aspects of human behavior. Ferreira et al. published a widget for mobile phones that provides information about battery usage [60]. Thereby, they analyzed when mobile users use and charge their devices. Using *appazaar*, another app recommender, Böhmer et al. collected data about app usage from a large number of participants [15]. They analyzed which apps participants used and when participants used them on a global scale. Sahami Shirazi et al. assessed the mobile device orientation when users interact with them, in particular when surfing the Web [164]. Further, they investigated notifications on the mobile phones [165].

Large-scale research in the field – we also refer to it as “in-the-wild” in this thesis – through app stores extends classic lab studies by moving the studies from the lab to the app store. In comparison to lab studies (or controlled field studies), large-scale studies through app stores have several fundamental differences. Table 2.1 depicts these differences. In the following part, we discuss these differences in details. We use the term lab or lab study in an inclusive way that spans studies where the researcher has control over the participants. Thus, it also includes controlled field studies, such as testing navigation systems where participants follow predefined routes.



**Table 2.1:** The comparison of research in labs vs. app stores

	<b>Lab</b>	<b>App Store</b>
<b>Control</b>	high	low
<b>User Recruitment</b>	based on invitation	based on self-selection
<b>Data Collection</b>	full access	partial access
<b>Costs/Effort</b>	limited to experiment design	danger of consumerization cycles
<b>Public Visibility</b>	none/low	potential of high attention
<b>Disruption</b>	any new technology available in the lab can be investigated	bound to adopted technology
<b>External Validity</b>	low	high

### **Uncontrolled study setup**

While in lab studies the setup is completely controlled, studies through app stores are essentially conducted in an uncontrolled environment. There is no direct contact between instructors and participants. Hence, it is not possible to directly describe the procedure of the study. Further, there is no control over participants to ensure that the app is used in the intended way or used at all.

### **Self-selection of participants**

Instead of actively recruiting participants, researchers release apps in app stores and ask users who intrinsically choose to download and use the app. Thus, the background (gender, language, profession, nationality etc.) of the subjects is completely out of the control of the researcher. Users may be much larger in numbers, but still demographically biased. However, the type of adoption may be also an important insight itself: finding a specific target group of the app and testing whether everyday users are ready for this kind of service.

**Data-logging captures user behavior**

Whereas in lab experiments the instructor takes measurements and also observes the experiment with his senses, research apps have to log all user activities and send off the data for remote analysis. This requires users' consent and sufficient bandwidth to collect the data from massive number of users. Whereas app interactions can be recorded, the actual user information, e.g., user context, intention, interests, may be hard to reconstruct from sensor measurements. Research in the field can continuously update and add logging mechanisms to achieve more fine-grain results. Furthermore, a single user may also be approached individually to collect more qualitative feedback and clarify situations that are not becoming clear from sensor data.

**Higher implementation costs**

Research in lab settings allows researchers to precisely calculate the costs of equipment, user recruitment, the experiments themselves, and data analysis. Instead, for in-the-wild research the resulting costs are less clear. First, the app has to be designed in such a way that to be enough attractive to volunteers find, install, and use it . Furthermore, the research app might be competing with similar apps in the market. Thus, costs of implementation may be significantly higher than for lab prototypes. App users normally expect quick reactions to feature requests and bug reports. Ignoring user feedback may quickly result in bad ratings and low retention and adoption of the app.

**Research becomes public**

As research apps are publicly available for studies in the large, comments and ratings they receive are also publicly visible. Successful apps might be picked up by blogs, forums, and other multipliers that could quickly result in a larger number of users and high public attention. Thus, experiments in-the-wild have to be able to cope with both positive and negative public reception instantly.

**Radical innovation vs. incremental development**

As in-the-wild research acts on public grounds and is conducted publicly, it is firstly limited by the technology adopted in the market. Research apps rely upon standard technology in the possession of real-world users versus experimental setups and emerging technologies available in labs. Secondly, research radically countering public practice, e.g., privacy critical topics, might trigger user opposition and denial of using such a service.

### **Higher validity with many users**

Despite all the risks and efforts of large-scale research in-the-wild may trigger, its main merits lead to obtain results with high external validity. Successful research tests hypotheses with hundreds, thousands, or many more users who voluntarily use the app. A larger number of users can underline the validity of lab studies with limited resources. In addition to the confirmation of lab results with real users, research apps also bear the potential of detecting specific insights on app usage that might only be triggered by the diverse variety in the real user context.







# III

EXPLOIT CONTEXT  
IMPLICITLY





# Outline

Obtaining contextual information and sharing this nonverbal information between non-colocated users in different contexts can compensate for the absence of nonverbal information in computer-mediated communication and result in the enhancement of communication. Being aware of context can reduce confusion and incoherence in such communication. In this part of the dissertation we investigate how contextual information can be obtained. Prior work has used different sensing technologies to monitor surrounding environments and develop context-aware systems. We, in particular, use two sources for exploiting contextual information: *human brain* and *explicit interactions with a mobile application*.

Information acquired from organs in the human body system provide valuable insights about the body state. The breath rhythm and heartbeat rate, for example, reveal whether the human has stress. The brain, as the center of the nervous system, provides information about mental activities. With advances in sensing technologies, commercial brain-computer interfaces allow researchers to monitor brain signals out of clinical setups. We investigate the feasibility of using information provided by such interfaces to determine certain mental activities.

On the other hand, we investigate to obtain contextual information by utilizing users interaction instead of any sensor. We assess how it is possible to obtain context information based on the explicit interactions with a mobile application. Over the last decade, mobile phones have evolved from simple telephones to sophisticated devices with powerful user interfaces. Smartphone users are no longer limited to the applications provided by the phone's manufacturer. Users are able to install and use third-party applications on their smartphones and transform them into important tools fulfilling diverse functions, from communication means to entertainment. We leverage these advances and assess how a user's sleeping behavior can be monitored solely based on interactions with user interfaces of a mobile application instead of using any wearable sensors.

This part of the dissertation includes following chapters:

- **Chapter 3 – Mental Task Awareness.** We investigate the feasibility of classifying mental tasks and obtain awareness about the user’s current activity using a single electrode brain-computer interface (BCI). Being able to determine what a user does enables numerous capabilities. Activities can be recognized either by embedding sensors in the environment or attaching sensors to the user’s body. However, certain mental activities cannot easily be recognized from the outside. We present using information acquired from the brain using a commercial off-the-shelf BCI to determine reading and relaxing tasks. Reading and regular relaxation are correlated with effective skill acquisition and learning. Being aware of these activities and providing feedback enable users to improve their life.
- **Chapter 4 – Video Annotation with Brain Signals.** This chapter assesses the reliability of contextual information, i.e., emotional information retrieved from a commercial BCI headset. We investigate the correlation between the emotional information that users explicitly provide and the emotional information extracted from the BCI headset through a case study. We present an annotation tool that automatically annotates a video based on excitement information acquired from the BCI headset. Through a user study we examine whether the implicit automatic annotation does correlate with explicit annotation the user manually does. Further, an algorithm is proposed for extracting highlights of a video based on the excitement information.
- **Chapter 5 – Implicit Sleep Monitoring.** In this chapter we explore the possibility of monitoring the sleeping duration based on the use of an alarm clock application for mobile phones. Sleep, as one of main human daily activities, reveals the routines and availability of a person. Increasing individuals’ awareness of their own and others’ sleep habits has the potential to motivate changes in behavior. While using actigraphy sensors is a traditional approach, here, we discuss how iconic interfaces of the app are used to obtain inputs from a user and monitor his sleep behavior. We, further, investigate how sharing sleep information impacts user’s connectedness and awareness.

# Chapter 3

---

## Mental Task Awareness

Determining the user's activities is central for ubiquitous computing. Being able to determine what a user is doing enables numerous use cases. For example, the recent quantified self-movement shows that there is an increasing interest in monitoring ourselves what we are doing. Prior research shows how to recognize diverse activities, including walking, sleeping and driving. Activities are determined either by equipping the environment with sensors (e.g., cameras) or attaching sensors to the user's body (e.g., accelerometers and gyroscopes). Certain mental activities, however, cannot easily be recognized from the outside. Differentiating between sleeping, relaxing, listening to music, or thinking hard about a problem can look exactly the same from the outside.

Recognizing and monitoring cognitive processes have gained momentum as novel sources for contextual information [74, 208]. It usually requires expensive and bulky hardware (functional magnetic resonance imaging, eye trackers) with few notable exceptions [26, 74]. Still most of the systems are cumbersome to wear. We explore to what extent an off-the-shelf, single electrode Brain Computer Interface (BCI) system can be used to recognize mental activities. The system itself is relatively unobtrusive, lightweight (a head band and a mobile phone) and can be used for long-term deployments (battery life of 6-8 hours).

We are particularly interested in contextual information related to learning, especially reading. The amount we read directly influences the size of our vocabulary and language skills [184] that respectively effects communication skills. Additionally, the more people read throughout the day the higher their general knowledge and critical thinking skills

are [184, 190]. Being able to just count the minutes we read daily would help to assess the general knowledge of a person, as there are strong correlations between these two [190]. In addition, regular relaxation, breaks and naps are correlated with more effective skill acquisition and learning [55, 203]. Relaxing sufficiently often and long enough can improve ones health, mood and fitness [180]. Thus, recognizing when a user reads and relaxes would enable and extend applications in order to improve users' lives.

The contributions of this chapter can be summarized as follows:

- We present using information implicitly acquired from a wearable, unobtrusive off-the-shelf single electrode brain computer interface system (BCI) to determine mental tasks and obtain contextual awareness.
- We describe and report the recognition feasibility of reading and relaxing tasks out of four other daily activities dependent and independent of the user.

*This chapter is based on the following publication:*

- A. Sahami Shirazi, M. Hassib, N. Henze, K. Kunze, and A. Schmidt. What's up in your mind? mental task awareness using single electrode brain computer interfaces. In *5th Augmented Human International Conference (Kobe, Japan), AH'14*, page 4. ACM, March 2014

## 3.1 Related Work

The human brain continuously generates electric signals. Table 3.1 describes different techniques available to record the signals [108]. The MEG (Magnetoencephalography) technique, for example, maps the brain activity by recording magnetic fields produced by electrical currents occurring naturally in the brain. On the other hand, the EEG (Electroencephalograms) technique measures brain voltage fluctuations resulting from ionic current flows within the neurons. The EEG approach is the most widespread method for signal acquisition due to its temporal and spatial resolution. Commercial BCI sets such as EPOC and NeuroSky use this technique to obtain brain's signals. In the EEG method, the electrodes attached to the surface of scalp record the brain signal. The signals are measured as voltage levels in the time domain. By using signal processing techniques, the recorded signals can be split into several frequency bands (Table 3.2). Analysis of the frequency domain provides insights into brain states and activities.

**Table 3.1:** Different techniques for recoding brain signals [108]

<b>Technology</b>	<b>Shortcoming</b>
Electrocorticogram (ECOG)	invasive, Surgery
Magnetoencephalography (MEG)	expensive
Computer Tomography (CT)	only anatomical data
Single Photon Emission Computed Tomography (SPECT)	radiation exposure
Positron Emission Tomography (PET)	radiation exposure
Magnetic Resonance Imaging (MRI)	only anatomical data
Functional Magnetic Resonance Imaging (fMRI)	expensive
Functional Near-Infrared (fNIRS)	expensive, infancy

Researchers have investigated task classification using EEG signals. Hosni et al. used the Kerin & Aunon dataset [99] and compared three different feature extraction techniques using Radial Basic Function and Support Vector Machine classification [92]. The best accuracy classification reported was 70%. del R Millan et al. proposed a neural classifier to recognize three mental tasks from online spontaneous EEG signals with 70% accuracy [50]. The tasks were relax, left/right movement, cube rotation, and subtraction. All the classifiers are user-dependent. Petersen et al. [149] used the EPOC and attempted to distinguish among emotional responses reflected in different scalp potentials when viewing pleasant and unpleasant pictures compared to neutral content. Crowley et al. [47] assessed and reported the suitability of the NeuroSky MindSet to measure and categorize a user's level of attention and mediation. Yasui [207] proposed a technique for measuring the psychophysiological status of the human and associated applications based on brain signals. He analyzed the mental state of a car driver and showed that the pattern while driving was modified by a specific activity such as talking on a mobile phone. We refer to Lotte et al. [115] for an overview on classification algorithms for EEG-based brain-computer interfaces.

Regarding reading activity, researchers also utilized EEG signals or eye movements to classify text comprehension, reading skills, and reading techniques. The EEG headset is, for example, used to collect and assess cognitive information from students while reading different texts [131]. Mostow and Beck used neuro-feedback to discover reading problems in children, being able to discriminate between reading easy and hard

**Table 3.2:** EEG Brainwave frequency bands and mental states

<b>Brainwave</b>	<b>Frequency band</b>	<b>Mental state</b>
Delta	0.1–3 Hz	Deep dreamless sleep, unconscious
Theta	4–7 Hz	Intuitive, creative, imaginary, dream
Alpha	8–12 Hz	Relax but not drowsy, conscious
Low Beta	12–15 Hz	relaxed yet focused, integrated
Midrange Beta	16–20 Hz	Thinking, be aware of self and surroundings
High Beta	21–30 Hz	Alertness, agitation
Gamma	30–100 Hz	Motor functions, high mental activity

sentences [130]. Neuro-feedback Training is a type of biofeedback that uses real time EEG or fMRI to illustrate brain activities with a goal of controlling central nervous system and training certain brain functions. Bulling et. al proposed a method for reading segmentation recognizing eye movement by electrooculography [26].

In contrast to previous research, we investigate the feasibility of classifying reading and relaxing tasks based on EEG signals retrieved from a single dry electrode BCI. In the following we describe how the data is collected and used for the classification of mental activities.

## 3.2 Data Acquisition

In the initial step a dataset is required. We developed a prototype to record brain signals and conducted a user study to collect data. We chose simple daily tasks that contain auditory and visual stimuli as well as thinking.

### 3.2.1 Task set

As our goal was recognizing reading and relaxing tasks, we chose three additional common mental tasks: watching, listening, and problem solving tasks. Table 3.3 describes

the task set. For reading we had a short story and for problem solving we used a Sudoku game at the medium level difficulty level. We recorded the audio from a popular radio station for the listening task. Finally, a short documentary video was used for the watching task. The task set includes auditory and visual sensory tasks that occur in different brain lobes.

**Table 3.3:** Five different tasks used for recoding and collecting brain signals.

Task	Type
reading a short story	visual
listening to music	auditory
watching a short video	auditory & visual
relaxing	auditory & visual
playing sudoku	problem solving

### 3.2.2 Apparatus

We used the NeuroSky BrainBand brain-computer interface for data acquisition. The device is a commercial BCI equipped with a single dry electrode placed on the subject's forehead. It has one reference electrode on the left ear. The device includes a chip which filters and preprocesses the EEG signal and transmits it via Bluetooth to the application (1 Hz). The EEG processing protocols are not open source. As stated in the NeuroSky white papers [135], an FFT is performed on the raw signal giving the band powers that are then scaled using a proprietary algorithm to produce outputs. The outputs are only relative to each other.

We developed an Android mobile application, called *Neuroid*, that uses the NeuroSky BCI to collect data. The app establishes a Bluetooth connection to the NeuroSky headset. It uses the Android API provided by the headset to retrieve and record data. The data is stored on the SD card of the phone in the Comma Separated Values (CSV) format. When the app is started, it collects demographic information about the user in the first step. Then, it establishes a connection to the headset and starts retrieving and recording data. Since the BCI headset is portable, the mobile application gives more freedom of movement during the data collection process.

### 3.2.3 Dataset

The NeuroSky headset provides different information. Following data collected from the BCI headset using (Table 3.4):

**eSense Values.** Attention and Mediation values ranging from 1 to 100, at a sampling rate of 1 Hz. These values are determined via Neurosky proprietary algorithms. Values between 40 and 60 are considered ‘*neutral*’ or baseline, between 60 and 80 mean slightly elevated eSense levels, and between 80 to 100 refer to strongly elevated attention/meditation levels. Values below 40 are interpreted as (slightly/strongly) lowered levels. A zero eSense value means the signal cannot be calculated reliably due to background noise.

**Neurosky Power Values.** A series of eight 3-byte long values ranging from 0 to 224 provided at 1 Hz. These values are: delta (0.5–2.75 Hz), theta (3.5–6.75 Hz), low-alpha (7.5–9.25 Hz), high-alpha (10–11.75 Hz), low-beta (13–16.75 Hz), high-beta (18–29.75 Hz), low-gamma (31–39.75 Hz), and mid-gamma (41–49.75 Hz). These values do not have a unit. Therefore, they can only be interpreted by comparing them with each other and to themselves to consider relative quantity and temporal fluctuations.

**Blink.** A one byte value ranging between 1 and 255 provided whenever a blink is detected. The value has no unit and only indicates the relative strength of the blink.

**Raw Wave.** A 16-bit value provided at 512 Hz sampling rate. Values for the communications protocol lie in the interval between  $-/+2048$ . Typically in EEGs, time-frequency transforms are used to change the raw signal to the frequency domain, to extract the EEG power values.

**Table 3.4:** Data recorded during a user study using the NeuroSky BrainBand BCI.

Data	Range	Sampling Rate	Description
<b>eSense Values</b>	1 to 100	1 Hz	attention & mediation values
<b>Power Values</b>	0 to 224	1 Hz	a series of eight 3-byte long values
<b>Blink</b>	1 to 255	when detected	one byte value presents the blink
<b>Raw wave</b>	$-/+2048$	512 Hz	a 16-bit value



## 3.3 User Study

We conducted a controlled user study and recorded brain signals during the activities from participants to collect the dataset required for achieving the goal.

### 3.3.1 Procedure & Participants

After a short introduction, the participant was asked to fill in a demographic questionnaire. Then, they performed the five tasks one after the other. Each task took five minutes for the data to be collected. We counterbalanced the order of the tasks to reduce sequence effects. The study was conducted in a quiet university laboratory with normal lighting conditions and minimal noise from other electronic equipment. We attempted to minimize the distraction and respectively noise in the data due to the surrounding environment. The study consisted of the five aforementioned tasks (Table 3.3). We recorded five minutes of brain signals during each task. At the end of the study the participants answered another questionnaire and provided qualitative feedback.

We recruited 20 participants (8 female) with an average age of 23.3 years ( $SD=2.2$ ). The participants were recruited through the university mailing lists and social network. All participants were students in different majors such as electrical engineering, mechanical engineer, etc. Only three participants had experience using a BCI device. Each session took approximately 35 minutes.

## 3.4 Task Classification

The initial analysis revealed that the data collected from five participants was corrupted due to the corruption in the Bluetooth connection or saturation of the Neurosky sensor. Thus, the data from these five participants was removed from the dataset. We used the annotated data collected from 15 participants using the NeuroSky BrainBand device to derive features. Further, the power values were clipped for some seconds because some of the subjects reached the maximum value. Since this noise was minimal, it was rectified by taking the average of surrounding signals to compensate for the signal clipped at particular points.

### 3.4.1 Feature Set

We derived spectral and time-domain features from the collected data. In addition, we use the signals preprocessed by the NeuroSky development kit. The features are determined for one second jumping windows using Matlab.

Spectral features are computed by applying a fast Fourier transform (FFT) on the raw signal and bandpassing the delta, theta, alpha, beta and gamma frequency bands, the average of each band is used as the feature. In addition, the ratios between all pairs of frequency bands are calculated. The mean FFT value and the variance of the FFT are also used. In prior work, the cepstral coefficients are also suggested for feature extraction on EEG signals [143, 189]. The coefficients are originally used for Automatic Speech Recognition (ASR) to extract features from speech. A power cepstrum is calculated the squared magnitude of the inverse Fourier transform of the logarithm of the squared magnitude of the Fourier transform of a signal. It is defined in the equation 3.1.

$$PowerCepstrum(x) = \left| \mathcal{F}^{-1} \left\{ \log(|\mathcal{F}\{f(t)\}|^2) \right\} \right|^2 \quad (3.1)$$

On the other hand, seven time-domain features are extracted from the raw time-domain EEG signal. These include the maximum positive, minimum negative and average amplitude of the raw signal per segment, and the Root Mean Squared (RMS) value of the raw signal. In addition, four features are extracted from the NeuroSky signals: the average attention and meditation values as well as the average NeuroSky power band values for the five frequency bands. Further, we considered *Hjorth* parameters derived from Hjorth [89] as features. These parameters have been used prior work and proved to be successful in EEG classification especially in the field of emotion recognition [8]. Three main parameters are included: signal activity, mobility, and complexity. These parameters are explained in the equations 3.2, 3.3, and 3.4. The Hjorth activity refers to the variance of the signal, whereas the signal mobility calculates the mean of the signal frequency. The complexity parameter measures the deviation of the signal from the sine shape.

$$Activity(x) = \frac{\sum_{n=1}^N (x(n) - \bar{x})^2}{N} \quad (3.2)$$

$$Mobility(x) = \sqrt{\frac{var(x')}{var(x)}}, \text{ Where } x' \text{ donates the first derivative} \quad (3.3)$$

$$Complexity(x) = \frac{Mobility(x')}{Mobility(x)} \quad (3.4)$$

In total 32 features were extracted. Table 3.5 includes the list of the features. These features were used to train classifiers for recognizing relaxing and reading. We developed a user-independent classifier that determines the mental activity without prior training and user-dependent classifier specific for individual participants.

### 3.4.2 Results

We used the features as input to train Bayesian networks that recognize relaxing and reading versus the respective other tasks. The Bayesian Networks has been used in prior work. It considers dependencies between various attributes. Experimental results reported in the following are obtained using WEKA [76]. All learning parameters use the default values in WEKA unless otherwise stated. We investigated two types of classification:

- **User Independent Classification.** We trained user independent classifiers using the leave-one-out cross-validation to train a classifier and test its performance. That means that we trained the classifier with data from 14 participants and evaluated the performance using the data from the remaining participant. The process was repeated for all participants resulting in 15 runs that were aggregated afterward.
- **User Dependent Classification.** In addition, we trained user dependent classifiers. Four minutes of each activity were used to train the Bayesian networks leaving one minute for evaluation. For both classifiers we used the feature selection option WEKA provides.

The results revealed that the user independent classification between reading and the other tasks was on average 68.2% and between relaxing and others was 53.5% of all cases. The user dependent classification determined reading vs. other tasks with 74.4% and relax vs. others with 79% on average.

Table 3.5: The features extracted for training classifier.

No.	Feature	Description	No.	Feature	Description
1	Attention Mean	Avg.NeuroSky Attention/segment	17	Theta to Beta	Mean Delta to Theta ratio/segment
2	Meditation Mean	Avg.NeuroSky Meditation/segment	18	Theta to Gamma	Mean Delta to Theta ratio/segment
3	Avg. Raw	Avg. of signal raw value/segment	19	Alpha to Beta	Mean Delta to Theta ratio/segment
4	Min. Negative Amplitude	Min. raw amplitude/segment	20	Alpha to Gamma	Mean Delta to Theta ratio/segment
5	Max. Positive Amplitude	Max. raw amplitude/segment	21	Beta to Gamma	Mean Delta to Theta ratio/segment
6	Raw RMS	RMS of raw signal/segment	22	FFT Mean	Mean FFT signal/segment
7	Mean Delta	Mean Delta signal/segment (FFT)	23	FFT Variance	Variance of the FFT signal/segment
8	Mean Theta	Mean Theta signal/segment (FFT)	24	NS Delta Mean	Mean NeuroSky Delta/segment
9	Mean Alpha	Mean Alpha signal/segment (FFT)	25	NS Theta Mean	Mean NeuroSky Theta/segment
10	Mean Beta	Mean Beta signal/segment (FFT)	26	NS Alpha Mean	Mean NeuroSky Alpha/segment
11	Mean Gamma	Mean Gamma signal/segment(FFT)	27	NS Beta Mean	Mean NeuroSky Beta/segment
12	Delta to Theta	Mean Delta to Theta ratio/segment	28	NS Gamma Mean	Mean NeuroSky Gamma/segment
13	Delta to Alpha	Mean Delta to Theta ratio/segment	29	Cepstral Coefficient	Cepstral coefficient/segment
14	Delta to Beta	Mean Delta to Theta ratio/segment	30	Hjorth Activity	Variance of raw signal eq. 3.2
15	Delta to Gamma	Mean Delta to Theta ratio/segment	31	Hjorth Mobility	Mobility of raw signal eq. 3.3
16	Theta to Alpha	Mean Delta to Theta ratio/segment	32	Hjorth Complexity	Complexity of raw signal eq. 3.4

Since the classification performance left room for improvement, we pairwise classified reading and relaxing with other tasks using the same classifiers. The result revealed that the classification performance between reading or relaxing and all the other tasks on average was more than 75% except for the reading vs. watching movie tasks (64%). Therefore, we excluded the movie task and repeated the classification using the same classifiers. The result showed that the user independent classification between reading and the other tasks excluding the watching movie was on average 68.5% and between relaxing and others was 58.2% of all cases. The user dependent classification determined reading vs. other tasks with 97.2% and relax vs. others with 73.2% on average. As expected, excluding the data from the watching task increased the performance of the classifications in total. Figures 3.1 depicts the results of user dependent and independent classification for all 15 participants.

### 3.4.3 Limitations

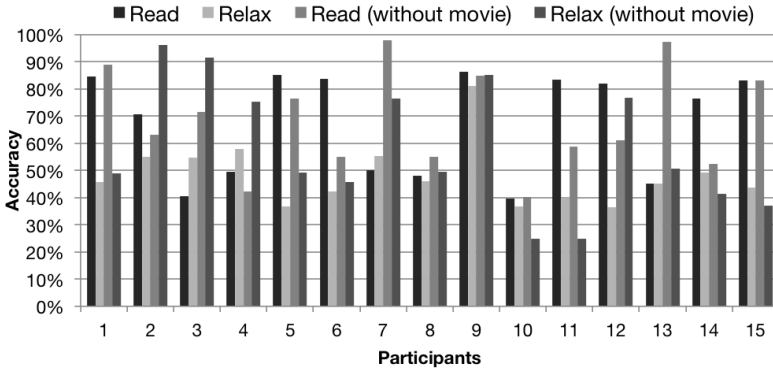
The BCI headset used for this study had a single dry electrode placed in front side of the head. This means that the device mainly collects signals from the frontal lobe. Different brain lobes are responsible for certain tasks. Thus, collecting data from other lobes is essential to increase the accuracy. Further, the task set included only five tasks. However, there were common daily tasks that included different sensory and mental activities. The data acquisition process was also conducted on the same day and in one session for each user. Separation sessions could minimize and normalize user's effects on the data.

## 3.5 Implication

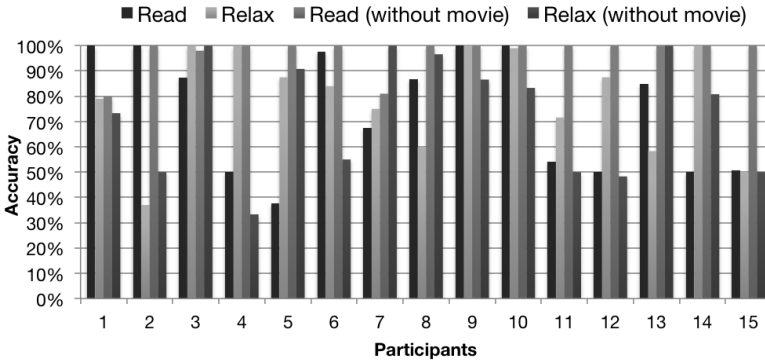
The results show that the mental task classification has a higher accuracy if it is preformed in a subject dependent manner. With independent reading classification we achieve

**Table 3.6:** The result of independent and dependent classification.

<b>Classification</b>	<b>Read</b>	<b>Relax</b>	<b>Read (without movie)</b>	<b>Relax (without movie)</b>
<b>user independent</b>	68.2%	53.5%	68.5%	58.2%
<b>user dependent</b>	74.4%	79%	97.2%	73.2%



(a) Result of the user-independent classification



(b) Result of the user-dependent classification

**Figure 3.1:** The result of task classification for all 15 participants.

accuracy between 80 % to 100 % for 8 of the 15 participants. Although the single electrode BCIs have a crude spatial and temporal resolution, the results interestingly reveal that user-independent classification could be possible. Qualitative feedback shows that all participants can imagine carrying the BCI for a longer period of time during everyday life. The results suggest that commercial BCI headset can be used outside of a clinical setup for exploiting context i.e., mental task contextual information.

Furthermore, the results reveal that the EEG signals from the front side of the brain seem similar for the reading and watching movie tasks. Although there is no one-to-one mapping between different mental tasks and certain brain lobes, certain sensory tasks can be majorly associated to particular brain lobes. For example, listening, as an auditory sensory task, is associated with temporal lobe activity. In contrast, listening to a speech involves language understanding which is associated with frontal lobe activity [185]. During mediating/relaxing the concentration in the relaxation process itself leads to high frontal lobe activity [13]. Brain puzzles such as Sudoku require memory, concentration, and high cognitive load which are all functions of the frontal lobe [182]. Watching movies and reading are associated with different parts of the brain. While the occipital lobe is responsible for vision, which is in the rear part of the brain, the frontal lobe is responsible for interpreting. As watching a movie is mainly a visual task, the mental activity is also in the rear lobe. Since the BrainBrand's electrode is placed on the front part of the brain, it mainly retrieves EEG signals from the frontal lobe. Hence, it is assumed that the BrainBrand BCI is suitable for classifying the activities that mainly happen in the front side of the brain.

## 3.6 Summary

In this chapter, we investigated whether it is possible to retrieve contextual information by determining and classifying certain tasks based on only brain signals obtained from an off-the-shelf BCI. We were particularly interested in the reading and relaxing activities. The reading activity influences the general knowledge. Reading has a direct correlation with the size of vocabulary, language skills, and communication skills. On the other hand, getting enough relaxation is essential for one's health. Being able to count the minutes one reads daily and is able to relax can be used to develop context-aware systems, provide feedback, and increase their awareness. To achieve the goal, a prototype was developed, a user study conducted, and the required data was collected. We recorded brain signals for five common daily tasks, i.e., reading, relaxing, watching, listening, and playing from 20 participants. However, only data from 15 participants were used for evaluation due to technical issues raised during the study. A set of features were extracted from the data and used to train classifiers that were both dependent and independent of users.

The analysis revealed that it is feasible to determine reading and relaxing activities using the off-the-shelf single dry electrode BCI system. The findings suggest that such BCIs can be used to determine activities that mainly occur in the frontal lobe of the brain due to the fact that the sensor is located in the frontal part of the brain. The distinction of reading with 97.2% and relaxing with 73.2% in the user-dependent cases is a first step to

implement an application that can count the minutes read or relaxed during the day. Such an application helps in assessing one's language skills, general knowledge and learning progress [184].

More interestingly, we achieve between 80 % to 100 % for 8 of the 15 participants using our independent reading classification. The results also reveal that user independent classification is possible, whereas all computer brain interfaces mainly need training before they can be used. Regarding the importance of reading as a knowledge acquisition task, this is an important insight. The results suggest that the brain can be used as a source to exploit contextual information for the development of context-aware systems. This information can be retrieved implicitly without adding any additional cognitive load to the user. Such systems can be used on a daily basis and outside of clinical setups. Furthermore, collecting more data can help to increase the accuracy of the classification.



# Chapter 4

---

## Video Annotation with Brain Signals

The brain, as the center of the human nervous and intelligent systems, is the most complex organ in the human body. Neuroscientists have studied the brain from all aspects - from how it structured and works to how it develops and malfunctions. Various invasive and non-invasive techniques (Table 3.1), e.g., MEG (Magnetoencephalography) or EEG (Electroencephalograms), have been proposed to monitor and record the brain signal [108]. Computer scientists have investigated systems that acquire and utilize brain signals for direct interaction with a computer instead of using normal pathways [191, 194, 195], called *Brain-Computer Interface* (BCI). Such interfaces allow the computer to monitor the brain and utilize its signals as nonverbal information for implicit or explicit interaction.

Brain Computer Interface (BCI) systems mainly use in medical and clinical research. This is due to their potential to assist patients with severe motor disabilities and brain disorders such as Amyotrophic Lateral Sclerosis (ALS) or Alzheimers. For example, there is a semi-autonomous wheelchair that uses a BCI to retrieve certain mental signals to move the chair [71]. Grag et al. used brain signals to assist patients with sleeping disorders [63]. However, with the advances in technology, off-the-shelf EEG BCI headsets are readily available and can be used in other research domains. Such commercial BCIs allow researchers to use the system in daily life and also with healthy individuals.

Further, the advances in signal processing enables the commercial BCI headset to retrieve and provide certain information such as emotional states (e.g., excitement, frustration, etc.) in real-time. This allows researchers to use this information in real-time and to develop context-aware systems. The question is how this information correlates with user's state. In this chapter we investigate the correlation between the emotional information provided by the commercial BCIs and emotional information that users explicitly provide. We consider a use case to achieve our goal. We investigate the feasibility of implicitly annotating videos based on the nonverbal information obtained by the BCI. The brain, as the center of the nervous system, has different neural activities that are a function of the mental and cognitive activities. Brain signals can reveal different information, such as facial expressions or the level of excitement. Further, it can reveal different information that correlates with the scenes the users watch in a video. This information can be used for annotating a video and generating a summary. Adding annotations to time segments on a video timeline makes it easier to search, find, and playback important segments of the video.

Various approaches have been explored to annotate videos (semi) automatically in order to summarize videos. Annotations are either defined automatically by analyzing and processing videos or explicitly by users/annotators. We utilize brain signals as implicit inputs for annotating video time segments and extracting a set of highlights. Conducting a user study, we examine whether the implicit, automatic annotation does indeed correlate with the explicit annotation the user manually does. We develop an annotation tool and propose an algorithm for extracting highlights of a video based on excitement information.

The contributions of this section are as follows:

- We assess the correlation between excitement information implicitly obtained from a commercial BCI headset and manually provided by users.
- We present an annotation tool that allows users to implicitly annotate videos based on nonverbal information, i.e., excitement information acquired from brain-computer interfaces.
- We propose an algorithm that is used for extracting highlights in a video based on the excitement information obtained from the brain using a commercial BCI.

*This chapter is based on the following publication:*

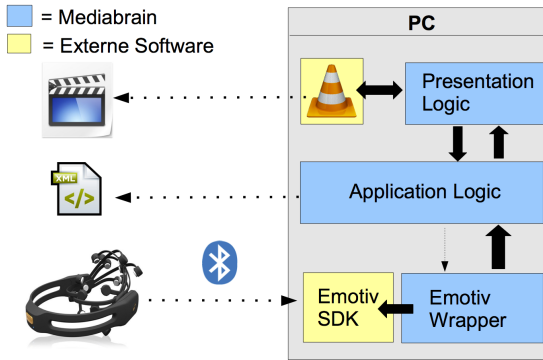
- A. Sahami Shirazi, M. Funk, F. Pfeleiderer, H. Glück, and A. Schmidt. Mediabrain: Annotating videos based on brain-computer interaction. In *Mensch & Computer 2012: interaktiv informiert—allgegenwärtig und allumfassend!?*, pages 263–272. Oldenbourg Wissenschaftsverlag, 2012

## 4.1 Related Work

The EPOC BCI by Emotiv and NeuroSky devices (MindSet, MindWave, BrainBand) are the two most popular commercial BCI sets. The EPOC has 14 saline sensors and two reference electrodes. It detects various facial expressions, level of engagement, frustration, mediation, and excitement. Comparatively, the NeuroSky BCIs have a single dry electrode. The NeuroSky devices have two electrodes and distinguish neutral and attentive mental states with 86% accuracy [135].

Researchers have started using BCIs for designing and implementing interactive systems used for different contexts in daily life. ThinkContacts is an application that allows users to call a contact in an address book by using brain signals as inputs. It uses the NeuroSky MindSet to measure the degree of attention each contact gets in an address book to find out which contact to call [147]. Neurowander is also a BCI game using brainwaves as inputs for a game [209]. In particular, researchers have utilized the ERP (Event Related Potential) wave for interaction with a system. The ERP wave is a brain response that is directly the result of a thought or perception. The P300 wave is the famous ERP elicited in the process of decision-making. Kanoh et al. used the P300 signal for controlling the mouse course. It works by cycling through the eight possible directions around the current cursor position. When the signal is triggered, the mouse moves into the desired direction [96]. Li et al. developed a P300-based keyboard that basically works by cycling through all letters until the desired one is reached [110]. NeuroPhone is a system that uses ERP signals obtained from the EPOC headset to select a contact from an address book on an iPhone and dial the number [28].

Various automatic or semiautomatic approaches for adding meta data to videos have been investigated. The meta data is used to extract specific part of videos for various reasons such as to generate a summary or to extract particular scenes. Yamamoto et al., for example, used the social activity, i.e., users' comments and weblogs for annotating videos [206]. Nagao et al. provided an annotation tool that allowed users to easily create annotations including voice transcripts, video scene descriptions, and visual/auditory



**Figure 4.1:** The architecture of the *MediaBrain* annotation application.

object descriptions [132]. Nakamura et al. explored affective response in order to understand video commenting systems [133]. Various algorithms also tried to automatically annotate videos [106, 202]. Saur et al. developed a tool, which automatically annotated basketball videos based on their content [188]. Sahami Shirazi et al. presented the use of an iconic interface on the mobile phone for sharing opinions during sport events, annotating the events, and detecting highlights [167].

None of the previous work utilized emotional information for annotating a video. Emotional reactions to scenes in a video of the video viewer are correlated with what happens in scenes. In this research we use the brain signals acquired from the Emotiv EPOC headset to annotate a video and find highlights. We develop an annotation prototype described in the next section.

## 4.2 Prototype

To investigate the feasibility of video annotation using information acquired from a brain-computer interface, the *MediaBrain* annotation tool, was developed. Figure 4.1 shows the architecture of the application.

We used the Emotiv EPOC headset to obtain the brain signals while watching a video. Similar to the NeuroSky headset used in the case study presented in Chapter 3, this headset also uses the EEG technique for the signal acquisition. But it has 14 electrodes and two reference electrodes. The electrodes are located all around the scalp and permits researchers to record signals from different brain's lobes. Hence, the data resolution of

this sensor in comparison to the NeuroSky headset is higher. The BCI headset transmits data to the computer via a Bluetooth connection. The SDK (Software Development Kit) that comes with the headset provides various measurements: facial expressions, level of engagement, frustration, mediation, and excitement. The headset has also a built-in gyroscope that detects the user's head orientation. The sampling rate is 128 samples/second.

### 4.2.1 Annotation Application

A video annotation application called *MediaBrain* is developed as an software for the personal computer. The application includes the open source VLC<sup>1</sup> media player for fully controlling the video events such as playback, pause, or stop a movie. It is implemented in Visual C++. *MediaBrain* and consists of three different layers (Figure 4.1):

- Emotive Wrapper layer: Establishes a Bluetooth connection with the EEG headset and handles the user's brain signals acquired. It uses the EPOC's SDK to retrieve the brain information.
- Application Logic layer: Records and stores the brain signals acquired in an XML file. It also tags the data with the video timestamp.
- Presentation Logic layer: includes a wrapper around the VLC media player and is used to control and play the video.

After gathering the information, the tool uses the XML file to identify, extract, and play scenes that were highlighted based on the excitement information. In the current version just the excitement values are recorded and used to annotate the video. However, it is easily a possibility to extend the tool and use other parameters for the annotation.

### 4.2.2 Annotation Algorithm

An algorithm is developed to extract the highlights based on the information recorded (Algorithm 1). The algorithm requires two parameters:

---

<sup>1</sup> VLC media player: <http://www.videolan.org/vlc/index.html>, last accessed August 27, 2014

- **L**: length of a highlight in seconds
- **N**: maximum number of highlights

In the first step all highlights are sorted in descending order based on the excitement value (see Figure 4.2(a)). Then, a highlight segment is detected. To achieve this, the scene that has the maximum excitement value together with the  $\pm L/2$  seconds is extracted. The other excitement values in this time segment are excluded for further calculation. This procedure is continued until the maximum number of highlights (**N**) is calculated or no more data is available. If **N** is not provided, all available points are extracted. The pseudo code of the algorithm is described in Algorithm 1. Figure 4.2(b) depicts the excitement graph with calculated highlights.

---

**Algorithm 1** The pseudo code describes the algorithm for annotating the scenes with the excitement values

---

```

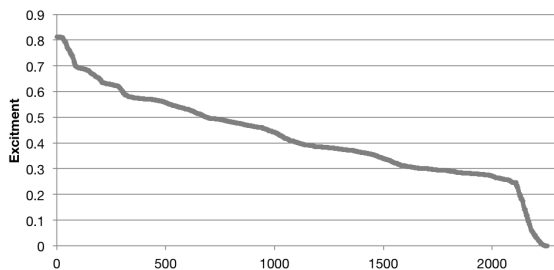
L ← length of highlight;
N ← maximum number of highlights;
excitement_array ← sort the excitement data in a descending; order
for 1 till N do
    item ← select first item in excitement_array;
    highlight_start_time = item.timestamp − L/2;
    highlight_end_time = item.timestamp + L/2;
    for highlight_start_time till highlight_end_time do
        remove data from excitement_array ;
    end for
end for

```

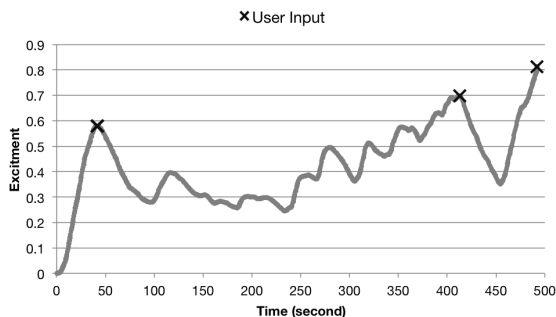
---

### 4.3 User Study

A user study was conducted to evaluate the *MediaBrain* tool and assess the feasibility of annotating the video implicitly based on the excitement information provided by the Emotiv EPOC headset. The study investigated the correlation of implicit excitement information obtained from the EPOC headset and the information users manually and explicitly provided.



(a)



(b)

**Figure 4.2:** (a) sorted user excitement values, (b) excitement graph with calculated highlights

### 4.3.1 Procedure

We invited the participants to our lab at the University of Stuttgart. After briefing, in the first step each participant was asked to answer a questionnaire about the demographics. Then, we continued by watching a video. The participant wore the headset and started watching a short animation movie, called *Big Buck Bunny*<sup>2</sup>. We assured that all the 16 electrodes had a good connection to the scalp. We selected this movie as it had few funny scenes that should result in a distinct excitement graph. The movie's length was 10 minutes and shown on a 40" display. Figure 4.3 shows the setup of the user study.

<sup>2</sup> *Big Buck Bunny* movie: <http://www.bigbuckbunny.org/>, last accessed August 27, 2014



**Figure 4.3:** The user study conducted in a quiet room in the laboratory. The participant wore the EPOC headset and watched a movie for 10 minutes.

The participant wore the EPOC headset while watching the animated movie on the big screen. The experimenter was able to check the live data on the second monitor.

Along with the excitement data obtained implicitly from the EPOC headset, the users were asked to explicitly specify their excitement while watching the movie. Hence, we extended the *MediaBrain* application in a way that users could state their excitement by pressing a button. During the study, we asked the participants to press the button every time they believed they were excited about a scene. This information was stored together with the video timestamp in an XML file and used later for the evaluation. At the end of the study, the participants filled in another questionnaire and provided qualitative feedback about their experience during the study. The study took approximately 30 minutes for each participant.

### 4.3.2 Participants

We recruited eleven participants (seven male average age 23.2,  $SD=1.02$ ) for the user study. All participants were students recruited via mailing lists and forums at the university. The study was conducted in a calm laboratory environment to minimize the distraction of the participants.



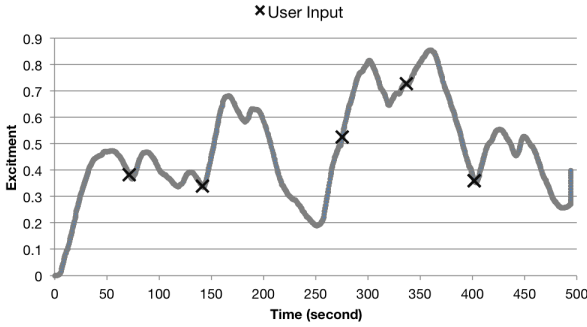
### 4.3.3 Results

Based on the demographic questionnaire, 70% of the participants used their computer daily for watching videos. None of the participants took part in any other user study related to the BCI or previously used any type of BCI headsets. Only one participant previously saw the animation shown in the study.

The results revealed that the participants pressed the button (specified their excitement explicitly) 13 times on average. There were six scenes where 85% of the users pressed the button. The rest of the explicit highlights were widely distributed. The six scenes included mainly unexpected actions in the movie, which led to surprise and excitement.

We also investigated the correlation between the explicit and implicit excitement information from users. We took each explicit input from users and checked whether this input matched with a highlight detected by the algorithm. The results showed that with  $L=5$  seconds only 27% of explicit inputs matched with the implicit excitements. With  $L=10$  seconds the result was 36%. Further investigation revealed that the user inputs were on average 10 seconds earlier than the local maximum excitement values. Interestingly, the user inputs matched with the points where the level of excitement started increasing (changes in gradient). However, we expected that the explicit inputs located on the local maximums (peaks) in the excitement graph (see Figure 4.4). Based on the Model Human Processor [30] the total cycle time of processors in humans' cognitive system, namely the perceptual, the cognitive, and the motor processor is approximately 300 milliseconds. On the other hand, the delay might be related to the headset. Nonetheless, based on the headset manufacture documents, no delay is reported. Therefore, we updated our algorithm in a way that the points where the level of excitement started increasing were considered to be highlights (see Algorithm 2). Based on the updated algorithm we analyzed the data again. The results showed that with the new algorithm 65% of the users inputs overlapped with the highlights extracted via the algorithm.

The qualitative feedback revealed that all users were relaxed during the study and enjoyed watching the movie. None of the users found the interaction with the EPOC headset inconvenient or disturbing. Also, all users mentioned that needing to explicitly identify their excitement did not distract them from concentrating on the movie. In total 77% of the users stated that they could imagine using the system in daily situations, i.e., in front of the TV.



**Figure 4.4:** User-Highlights mapped to the excitement graph using the updated algorithm.

## 4.4 Implication

The results reveal correlations between the scenes in the movie and the excitement level acquired from the brain-computer interface. With the algorithm, proposed videos can be implicitly annotated with excitement information obtained from the EPOC headset and highlights can be extracted. The study shows that the local maximums in the excitement graph correlate with the highlights in the video, but these are not the moments users believe they are excited. The comparison of algorithm output with users' explicit inputs depicts that moments which users think they get excited are, interestingly, the points

---

**Algorithm 2** The updated algorithm for annotating the scenes with the excitement values after the user study analysis

---

```

L ← length of highlight;
N ← maximum number of highlights;
excitement_array ← point where the gradient changes;
for 1 till N do
    item ← select first item in excitement_array;
    highlight_start_time = item.timestamp - L/2;
    highlight_end_time = item.timestamp + L/2;
    for highlight_start_time till highlight_end_time do
        remove data from excitement_array ;
    end for
end for

```

---

where the excitement value starts increasing (gradient changes), as indicated in the excitement graph. This led us to update the algorithm and propose a new approach. It is important to consider points that the gradient is changed, instead of the peaks in the excitement graph for moments that users believed they were excited.

The study showed that this information could be used to annotate a video. Users emotional reactions while watching a video are rich resources for annotating and extracting various scenes in a video. The annotation can be used to automatically generate a summary of a video. Annotating a video with different emotional information gives us the opportunity to create a variety of summaries based on different criteria. We only utilized and investigated the excitement information. However, we assume other emotional information such as frustration or facial expressions can be similarly used for annotation or any other purposes. Additionally, sharing the emotional reactions between non-colocated viewers might result in an increase in the connectedness and awareness they experience.

On the other hand, it is presented that the annotation can be preformed implicitly and transparently - without any additional cost. Such implicit interaction is passive and does not apply any additional cognitive load to the users. In contrast to manual annotation, the cost of implicit annotation is very low. The qualitative feedback shows wearing the BCI headset while watching the video did not disturb the users. It should be mentioned that our participants wore the headset for a short period of time. For longer usage, further investigation is essential.

In general, the study unveils that emotional information provided by commercial BCIs are valuable sources to obtain context information about users. This nonverbal information can be used to developed context-aware systems. The information can be shared in computer-communicated communication to enhance such communication.

## 4.5 Summary

In this chapter we investigated whether the contextual information, i.e., excitement information, implicitly obtained from a commercial BCI correlates with the emotional state that the user explicitly provided. The brain, as the center of human intelligence, provides valuable information about users' mental states. Brain-computer interfaces (BCI) provide the opportunity to acquire brain signals and establish a direct communication channel between the brain and external devices. We used commercial BCIs and conducted a feasibility study to investigate this correlation. In the case study, we considered video annotation based on this information. This information is a rich resource and provides details about the scenes in a video.

To achieve the goal, we developed an annotation tool called *MediaBrain* that uses the EPOC headset to acquire brain information and annotate a video. An algorithm was proposed that extracts highlighted moments in a video and generates a summary based on the information acquired. The annotation is performed implicitly. In contrast to manual interactions, the implicit interaction does not add any additional cognitive load or require additional actions from users. During a user study, we assessed this feasibility and compared the annotation with the explicit inputs from users.

The results reveal that implicit video annotation based on information retrieved from a BCI headset is possible. The emotional information obtained from the BCI correlates with the scenes in the video. Further, this information correlates with explicit inputs from users. Interestingly, it is not the local maximums, but rather the points in which the gradient changes that are the moments that users believe they got excited. Using this information, it is confirmed that highlights can be extracted and a summary can be automatically generated. This study only assessed the excitement information. However, we believe that the same approach can be applied to utilize other information for annotation. Further, the results suggest that emotional information obtained from BCIs is rich resource to retrieve context information. It can be used in computer-mediated communication and the development of context-aware systems.

Further, the results suggest that the commercial brain-computer interfaces can be used to obtain emotional states of users. This information indeed correlates with the emotional state of users. As sharing emotions is one of the main purposes of human communication, such emotional information can be obtained from BCIs and implicitly shared between users in computer-mediated communication. It can also be use for developing new context-aware systems.

# Chapter 5

---

## Implicit Sleep Monitoring

The proliferation of mobile devices in everyday life has led to an increasing amount of information about users' personal contexts. With sensors embedded in the smart phones different contextual information about the user's activities and the surrounding environment can be obtained. Implicit and explicit interactions with applications are other sources for retrieve of contextual information. In this chapter of the dissertation we aim to investigate how one can use solely explicit interactions with a mobile application for implicitly monitoring certain physical activities, i.e., sleep patterns. Sleeping has been identified as one of the prime daily activities that reveals the availability of a person. It a contributes significantly to the state of an individual's mental and physical health [16, 58]. There is a mutual relationship between sleep and daily life where problems in one often impacts the other [210]. There is also a correlation between a lack of sleep and an increase in the number of diseases a person is prone to contracting, e.g., heart disease [11] and diabetes [70]. Lack of sleep can affect memory [117, 197], cognitive functioning [59], and alertness [17], which can lead to poor work performance and put individuals at an increased risk for injury. As such, increasing individuals' awareness of their own and others' sleep habits has the potential to motivate changes in behavior that result in healthier daily practices [61].

On the other hand, many individuals share details of their lives with fellow friends [4]. Looking at the usage of Google or Facebook, it is apparent that keeping friends updated with one's current status through the use of social media has become a popular pastime. It is common for many posts to contain context information about a user. This information is either posted automatically by a third-party service or is posted intentionally by the

user. From the perspective of the reader, this information is useful in that, for example, couples in long-distance relationships are kept updated on what their partner is currently doing, family members are kept up-to-date with each other's activities, and friends might be triggered to meet each other based on posts containing location information.

We explore how monitoring and sharing sleep information as another type of activity can impact awareness, connectedness, and sleeping behavior. Sleep information, e.g., whether a person has gone to bed or is awake, shows not only one's daily routines, but also indicates one's physical state, reveals one's sleep patterns, and reflects a sense of wellness. This information can be valuable not only to oneself, but also to others. On one hand, knowledge of one's sleeping habits might explicitly trigger healthier sleeping behavior (e.g., if an individual realizes that she did not sleep enough during the past couple of days, she might attempt to catch up on sleep in the near future), on the other hand, being aware that most friends are already asleep might implicitly lead one to go to bed as well. Similarly, sharing this information with one's social network can facilitate social interaction and impact awareness and connectedness, as it indexes one's presence, absence, and availability which is essential for communication.

Three high-level research questions (RQs) are investigated in this project:

1. Is it possible to reliably monitor the sleep duration of users using only their interactions with a mobile phone application instead of using any physical or wearable sensors or devices?
2. Does providing a mechanism to track sleep information impact individuals' awareness of their own sleep habits and, if so, does this increased awareness inspire them to think of starting to engage in healthier sleep behaviors?
3. Does using an alarm clock that enables the sharing of sleep information through a social network impact users' feelings of connectedness and awareness?

To achieve our goal, we implemented a social alarm clock app for Android mobile phones, called *Somnometer*. The app allows users to rate their sleep quality and specify their sleep status, i.e., gone to bed, snoozed the alarm, and awake. Users can also share their sleep status and quality with their social network. While the quality of sleep is manually obtained from the user, the duration is estimated based on tracking a user's explicit interactions with the mobile application. We conducted two user studies in an attempt to address our research questions and evaluate the prototype. In a controlled study, we recruited eight participants to use the app for six weeks and provide the research team with qualitative feedback on their experiences. As we were also interested in observing emergent user behavior, we conducted a second parallel study in the wild. To recruit

broadly, we distributed the app on Google Play, the official Google Android marketplace, for free over the duration of the study. During the in-the-wild evaluation, of the 725 users who downloaded *Somnometer*, 173 used it actively over the course of six weeks.

This chapter demonstrates the following contributions:

- It is possible to monitor users' sleep duration using only an application on the mobile phone instead of using wearable actigraphy devices.
- Tracking and visualizing one's sleep habits impacts knowledge of sleep activity that, in turn, can be used to encourage healthier sleep behaviors.
- Sharing sleep information with social networks impacts feelings of awareness and connectedness among friends.

*This chapter is based on the following publication:*

- A. Sahami Shirazi, J. Clawson, Y. Hassanpour, M. J. Tourian, A. Schmidt, E. H. Chi, M. Borazio, and K. Van Laerhoven. Already up? using mobile phones to track & share sleep behavior. *International Journal of Human-Computer Studies (IJHCS)*, 71(9):878 – 888, 2013

## 5.1 Related Work

When it comes to tracking sleep and helping people wake up or fall asleep at appropriate times, the alarm clock has long been the subject investigated by various researchers. Ozneć et al. designed the Reverse Alarm Clock for improving children's sleeping behavior [145]. The goal of this project was to help children know whether or not it is a good time to get out of bed. Landry et al. used an alarm clock for supporting personal, routine-based decision-making [105]. The basic functionality of an alarm clock was challenged in [176] and a networked alarm clock was designed that uses other's presence information as a source for setting up the wake-up time. Hemmert et al. designed the Digital Hourglass to enable users to set a desired wake up time by the number of hours the user wants to sleep [83]. With this approach users are more focused on the amount of sleep.

Audio is the modality most frequently used to wake up an individual using their mobile phone as an alarm clock. Some non-phone-based alarm clocks employ more creative

and sophisticated strategies for waking individuals such as using tactile feedback, e.g., vibrating beds (mainly for persons with hearing impairments), visual feedback, e.g., through increasing brightening of lights gradually as with Philips *Wake-up Light*<sup>3</sup>, or simply forcing users to get out of bed to turn off the alarm by jumping from the nightstand and rolling away, e.g., the Tocky<sup>4</sup> alarm clock. Furthermore, there are various products on the market today that can be used to track different aspects of sleep, such as duration, frequency, or quality using non-invasive methods of monitoring, called actigraphy devices. Actigraphs are used to record full circadian rhythm data over the course of multiple, successive days. Accordingly, they have the ability to produce insight in the user's sleep habits and rhythms. Actigraphy devices have been validated and used in the medical community as well as are available as consumer information tools (e.g., the ActiWatch<sup>5</sup>).

Several commercial mobile applications are also available that use the phone as a device to manually track users' sleep (e. g., Sleep Tracker TYLENOL<sup>®</sup> PM<sup>6</sup>). Some apps employ automatic sensing via accelerometer and orientation sensors to track sleep (e.g., Sleep Cycle<sup>7</sup> and Sleep as an Droid<sup>8</sup>). FitBit<sup>TM9</sup> monitors how many times and how long the user wake up during the night using a three-dimensional accelerometer and demystifies the user's sleep cycle. Other products have been targeting sleep phase detection specifically in order to wake up users at a more convenient sleep stage, using wearable units such as a headset (e.g., Lark<sup>TM10</sup>) or a wristband (e.g., WakeMate<sup>11</sup>). Several researches also have investigated using body posture detection and movements during sleep as means to measure sleep quality [192, 111]. In contrast, our app does not use any actigraphy device to monitor sleep. We refer to Choe et al. for a comprehensive overview on design considerations and challenges of using computing to support healthy sleep habits [35]. Schmidt et al. discussed and reported the current state technologies for tracking sleep behavior [179].

---

<sup>3</sup> Philips Wake-up Light: <http://www.wakeuplight.philips.com/>, last accessed August 27, 2014

<sup>4</sup> Nanda Home: <http://www.nandahome.com/>, last accessed August 27, 2014

<sup>5</sup> Actiwatch, Philips Respionics, Andover, MA, USA

<sup>6</sup> Sleep Tracker TYLENOL<sup>®</sup> app: <http://itunes.apple.com/app/id317459304>, last accessed August 27, 2014

<sup>7</sup> Sleep Cycle app: <http://www.sleepcycle.com/>, last accessed August 27, 2014

<sup>8</sup> Sleep Droid app: <https://market.android.com/details?id=com.urbandroid.sleep>, last accessed August 27, 2014

<sup>9</sup> FitBit: <http://www.fitbit.com/>, last accessed August 27, 2014

<sup>10</sup> Lark<sup>TM</sup>: <http://lark.com/>, last accessed August 27, 2014

<sup>11</sup> WakeMate: <http://www.wakemate.com/>, last accessed August 27, 2014



Sleep patterns seem to have a tight correlation with awareness and connectedness. Interestingly, the bed, as a medium for intimate communication and as a tool for bridging the distance between remotely located individuals, has already been the subject of several projects, e.g., [54, 68]. Mhóráin and Agamanolis employed an augmented eye mask to monitor eye movements and transmit muscle signals of sleeping pattern to a remote device and map them to music [121]. The aim was to increase awareness between non-colocated partners. However, the system was not evaluated through any formal study. BuddyClock [100] allows users in a small social network to automatically exchange sleep information with each other. It is reported that the alarm clock affected participant behaviors and allowed them to feel more connected to those with whom they shared their sleeping behaviors.

In contrast to previous projects, we only focus on sleep as a daily activity. Our research investigates the potential of monitoring sleep behaviors using only users' explicit interactions with the mobile phone app instead of using any sensor or actigraph device. The interaction with the app can reveal different information about the user's sleep status. We aim at exploring the feasibility and reliability of using such sleep information to provide feedback and impact their awareness on users' sleep habits. Previous work, further, shows that users perceive awareness and connectedness as being meaningful and important. This inspired us to investigate how sharing sleep patterns can impact awareness and connectedness. We study how sharing only sleep activity information with social networks can be used as a way to impact awareness and feelings of connectedness. The app uses existing social networks, i.e., Facebook for sharing the information and does not convey any information about other users.

## 5.2 Somnometer: a social alarm clock app

To answer our research questions we developed a social alarm clock app for Android phones, called *Somnometer*. The app allows users to collect information about their sleep and monitor their sleep behavior. Besides the conventional alarm clock features, users can define their sleep status, i.e., gone to bed or awake, and rate their sleep quality. *Somnometer* allows users to share their sleep status (gone to bed, snoozed the alarm, and awake) and quality with their friends through social networks and track their sleep behaviors. While the quality of sleep is manually obtained from the user, the app estimates the duration based on tracking a user's interactions with the app. Based on the design framework proposed for sleep technologies [35], the *Somnometer* app's characteristics is described in the Table 5.1.



**Figure 5.1:** The Somnometer app screenshots (a) the user is awake and no alarm is set, (b) setting an alarm in 2 ways: by time and by countdown timer, (c) an alarm is set and the user has gone to bed, (d) when the alarm is triggered, it can be snoozed or deactivated, (e) after the alarm is stopped the user rates their sleep.

**Table 5.1:** The *Somnometer* app’s characteristics based on the design framework proposed by Choe et al. [35]

<b>Goal</b>	Monitoring	<b>Technology Platform</b>	Mobile
<b>Feature</b>	Tracking, Persuasion, Social	<b>Stakeholders</b>	Without Disorder
<b>Source</b>	Peer-reviewed literature, Folk wisdom	<b>Input Mechanism</b>	Input by user, Automatic

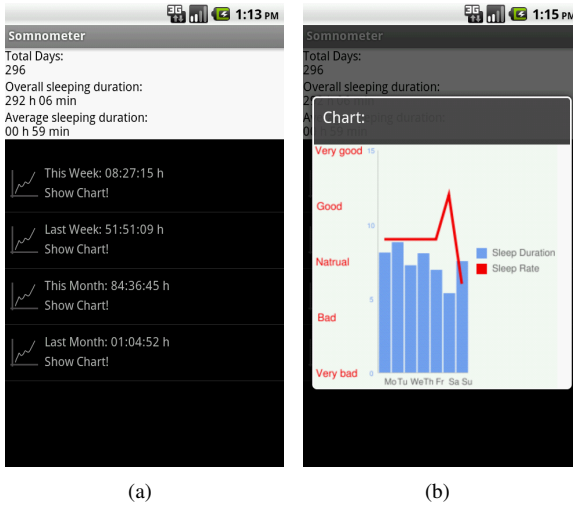
### 5.2.1 App Functionality

The *Somnometer* app has several functionalities. The basic functionality is the alarm clock. Users are able to set an alarm in two ways: (1) by entering a specific time or (2) by defining a countdown timer, which allows users to set the number of hours they want to sleep (see Figure 5.1(b)). Users need to set their alarm daily at any time in order to use this application. Users can also specify their sleep status via a button on the top of the interface (Figure 5.1(a) & Figure 5.1(c)). The button has two states: “awake” and “sleeping”. As soon as an alarm is switched off, the app assumes that the user’s status has changed and they are awake, thus the status button displays “awake” (Figure 5.1(a)). On the other hand, when an alarm is set and the user explicitly presses this button, then the button toggles from “awake” to “sleeping” and it is assumed that the user has gone to bed (Figure 5.1(c)). However, this does not mean that the user falls asleep immediately.

When an alarm is triggered a dialog box pops up allowing users to deactivate or snooze the alarm (Figure 5.1(d)). The snooze default duration is 5 minutes and can be customized. Furthermore, when an alarm is deactivated a dialog pops up, asking the user to rate her sleep quality (Figure 5.1(e)). Reporting sleep quality is optional. The app collects, keeps track of the user’s sleep behavior, and provides feedback about one’s sleep behavior (Figure 5.2) if enough data is available.

Further, users can associate a message to each button state. The default messages in the app are “Good Night!” when the user goes to bed, “Sleep a bit more” when an alarm is snoozed, and “Good Morning!” when the alarm is deactivated. Users can modify the default messages (Figure 5.1(b)).

When launching the app for the first time, the user grants the *Somnometer* access to her Facebook account with a limited set of permissions, enabling *Somnometer* to post



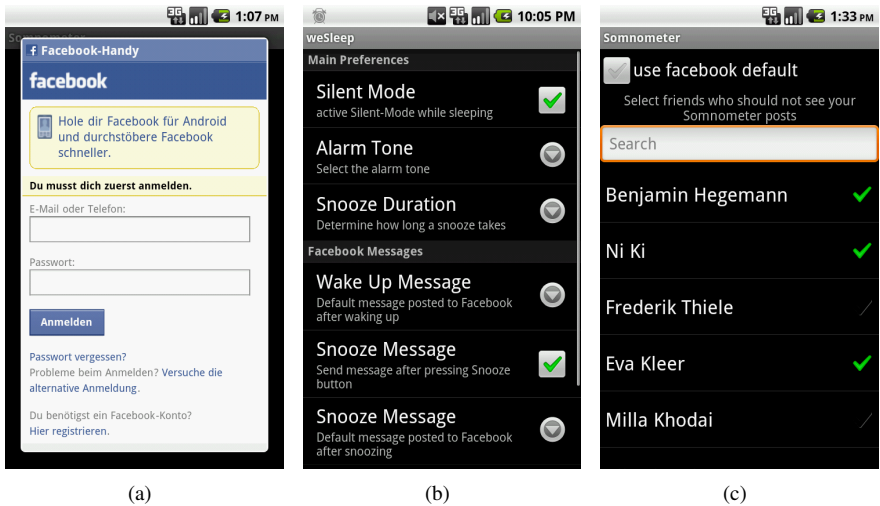
**Figure 5.2:** (a) Sleep duration is shown within different time frames, (b) a week visualizes of the sleep behavior and rating.

the status messages to the user’s Facebook Wall (Figure 5.3(a)). The button “Post to Facebook” in Figure 5.1(b) enables/disables sharing messages in Facebook.

There are separate privacy controls available for sharing sleep state (the “awake”/“sleeping” status), the sleep rating, and snoozing (Figure 5.3(b)). The default sharing setting is no info posted to Facebook. If users choose to share their sleep state with Facebook, their messages are automatically posted to the Facebook directly after their assigned states have occurred. Users have the ability to also choose to share their sleep ratings with Facebook. Sleep ratings are shared together with the “awake” message in a single status message.

Since sharing sleep information might create privacy issues, it is crucial that users have a full control over what information is shared with whom. Thus, we provide a feature that enables users to customize the privacy setting of the *Somnometer* sharing posts and choose friends with whom they want to share their information (Figure 5.3(c)). If the user does not use this feature, then the default privacy setting of her Facebook account is used.

Some users prefer to keep mobile phones in silent mode in the bedroom. To minimize disruption, a feature is provided which puts the mobile phone on silent mode when the user toggles the status button to “sleeping”.



**Figure 5.3:** (a) the user grants the app access to her Facebook account with a limited set of permissions, (b) separate privacy controls available for sharing sleep state, (c) visualization of a week of sleep behavior and rating.

To gather insight into app usage and understanding as to how users interact with the app, an analytics platform for mobile devices, called Flurry<sup>12</sup>, was used. Moreover, to obtain demographic information about our users, we ask the users to enter their gender and age when the app is installed and launched for the first time. We also ask how often and for which purposes they use their phone alarm clock. Answering these questions is optional.

## 5.2.2 Implementation

The app is implemented for Android smartphones and compatible with the Android platform 1.6 and above. When it is run for the first time, it initializes a local SQLite database to store the data collected from users. The app uses the information to monitor sleep patterns. Further, it is connected to a central database and stores data collected from each of the users anonymously. *Somnometer* is available in English and German to target a large and diverse set of users.

<sup>12</sup> Flurry analytic tool: <http://www.flurry.com/>, last accessed August 27, 2014

### 5.2.3 Monitoring Sleep Behavior

A key feature of *Somnometer* monitors a user's night sleep duration, which is automatically based on the user's interaction with the app, aimed at answering RQ1. RQ1 challenges the reliability of monitoring a user's sleep duration using only her interactions with a mobile phone application instead of using any physical or wearable sensors or devices. By recording the active status and monitoring changes, it is possible to estimate how long a user has slept.

In order to monitor sleep patters, we log the following information:

- time of day when users set an alarm ( $t_{alarm\_set}$ )
- scheduled wake up time ( $t_{alarm}$ )
- time of day when users went to bed ( $t_{bed}$ )
- number of times an alarm is snoozed ( $n_{snooze}$ )
- duration of snooze ( $d_{snooze}$ )
- time of day when alarm is deactivated ( $t_{deactivate}$ )

Thus, the sleep duration ( $d_{sleep}$ ) is estimated by the difference between the times the user went to bed and the alarm was deactivated.

$$t_{deactivate} = t_{alarm} + n_{snooze} \times d_{snooze}$$

$$d_{sleep} = t_{deactivate} - t_{bed}$$

$$d_{alarm} = t_{deactivate} - t_{alarm\_set}$$

The deactivation time ( $t_{deactivate}$ ) might differ from the scheduled wake up time ( $t_{alarm}$ ) if the alarm is snoozed. The time of day when users went to bed ( $t_{bed}$ ) is the time the user manually presses the status button and changes the button state from “awake” to “sleeping”.

However, there is a chance that the user just sets an alarm as a reminder, forgets to press the status button when going to bed, or takes a short nap. In this study we are particularly interested in sleep at night, so we try to filter out irrelevant information by defining the following rules:



**Figure 5.4:** HedgeHog is a device worn like an actigraph at the dominant wrist records motion, posture, and light data over a long time-span.

- If the sleeping status is not changed and the alarm duration ( $d_{alarm}$ ) is less than 120 minutes, we presume the alarm is used as a reminder or they had a nap and ignored this dataset.
- If the alarm duration ( $d_{alarm}$ ) is more than 14 hours and if the user forgot to change the sleeping status (manually), the  $t_{bed}$  is thus not available and the calculation of the current sleep duration is ignored. We considered 14 hours as the upper threshold, as this is well above the sleep duration of healthy adults.

To gather information about the user's sleep quality, a dialog box pops up when an alarm is switched off and asks the user to explicitly rate their sleep on a 5-point Likert scale (1=very bad, 5=very good) (see Figure 5.1(e)). This dialog pops up if the calculated sleep duration ( $d_{sleep}$ ) is more than 2 hours. In an effort to provide the user information on her sleep habits, the sleep duration and ratings are visualized using a chart that allows users to toggle between different time frames (week and month) as depicted in Figure 5.2.

## 5.3 Sleeping Data Ground truth

As mentioned above, users' sleep duration is estimated based on their explicit interactions with the *Somnometer* application. To assess how reliable the estimated sleep duration is, we used *HedgeHog*<sup>13</sup>, an open source device that is worn like an actigraph on a user's dominant wrist and records motion data over a long time-span (Figure 5.4). It has sensors

---

<sup>13</sup> HedgeHog: <http://www.ess.tu-darmstadt.de/hedgehog>, last accessed August 27, 2014

such as the accelerometer and light sensors. It also includes internal storage and logs data locally on the device.

*HedgeHog* measures three modalities for night sleep detection:

1. Light intensity, which is typically low during nightly sleep.
2. Amount of motion using a 3D accelerometer sensor, since movements are rare during sleep.
3. Time of day, where nightly sleep occurs usually between 11pm and 9am.

It uses these three pieces of information as inputs for a Hidden Markov Model (HMM) classifier for automatically detecting sleep sessions. The HMM classifier is already trained with the data gathered from 10 users (2 female and 8 male, average age 33 years) over the course of six months. The device has been previously validated to detect nightly sleep and other sleep characteristics in different research projects such as [18, 19]. This actigraph device can measure the sleep duration and quality more precisely. By using these devices, we gathered ground truth data without users having to manually annotate sleep segments themselves or create a diary.

## 5.4 Controlled User Study

We conducted a controlled study in which we applied a mixed-methods approach to develop a fuller picture of user practices through the use of questionnaires, interviews, and automatic logging. In our controlled study we investigated the feasibility of tracking sleep habits based only on users' interaction with the app without the use of wearable sensors. We assembled several *HedgeHog* systems and used them as the ground truth for this study. The controlled study also investigated whether sharing sleeping status can impact connectedness and awareness between friends. It investigated the impact of providing feedback about sleep patterns to users on their subsequent behavior.

### 5.4.1 Procedure

We conducted a within-subject study for six weeks by recruiting eight students (all male, average age 24 years,  $SD=2.5$ ) from different majors at the University of Stuttgart. None of the participants reported having any type of sleeping disorder or had participated in any other sleep study. They were compensated with €20 at the end of the study.



Since our target group was users who use their mobile phone as an alarm clock, we recruited participants who had already adopted this practice prior to using the app. During the study we asked our participants to use *Somnometer* as their main alarm clock and to share their sleep status messages and ratings on Facebook. We deliberately chose participants who were regular Facebook users (spending at least an hour per day browsing Facebook).

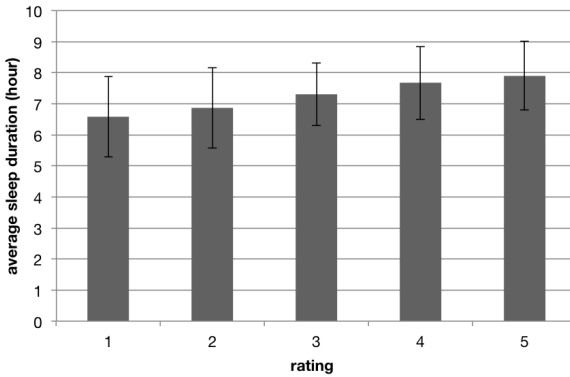
We divided the participants into two groups of four. The first group started the study by sharing status messages and ratings on the Facebook (G1) while the second group did not share any information (G2). After three weeks the participants were informed to switch groups. Before starting the study, all participants were invited to the lab and introduced to the study. They were asked to fill in a demographics questionnaire and received a Hedgehog device. We asked participants to wear the device during the entire duration of the study, particularly when they went to sleep. In order to ensure the app worked on their phone properly, we asked them to download and install the app on their phone and to familiarize themselves with the app before leaving the lab. Researchers were available to answer any and all of the participants' questions.

After six weeks, the participants were invited back to the lab to return the sensor. At this time, they completed a questionnaire about their experience with the app and its usage during the study. They were also interviewed and asked to provide feedback about their experience using the app in order to assess the connectedness. They additionally filled in a SUS (System Usability Scales) questionnaire [23]. Apart from collecting data from the app, we also collected the comments on Facebook posts shared by the app during the study.

## 5.4.2 Results

Based on the data logs, 336 alarms were scheduled during the controlled study and 347 messages were posted on Facebook (Mean=43.37 messages/participant, SD=3.11). However, we encountered with several issues in recording the information, for example, bugs in the app, no network connection, and running out of battery. Furthermore, participants sometimes forgot to rate their sleep. These issues forced us to only consider 217 data out of 336 for further analysis. This data include complete information about the sleep sessions (duration and rate). The average sleep duration was 7.44 hours (SD=1.17) and average rating was 3.45 (SD=1.04).

An ANOVA test revealed that sharing on Facebook had a significant effect on the sleep rating,  $F(1215)=5.487$ ,  $p<.05$ . Interestingly, the participants rated their sleep worse when they shared it on Facebook ( $M=3.31$ ,  $SD=1.06$ ) compared to not using sharing on



**Figure 5.5:** Rating base on the sleep duration. The analysis reveals that longer sleep session get higher ratings. The error bars shows the standard deviation.

Facebook ( $M=3.64$ ,  $SD=.99$ ). The results show no significant difference in their sleep duration between groups.

The Pearson correlation analysis revealed a positive correlation between the sleep duration and its rating ( $r=.23$ ,  $n=217$ ,  $p<.001$ ), depicted in Figure 5.5. Further analysis indicated that the highest rates, i.e., good or very good, were given to sleep durations between 5.20 and 8.06 hours. The questionnaire results revealed that participants found the app very useful and easy to use. The score of 83 out of 100 in the SUS test also reflects the usability of the app. Five out of the eight subjects mentioned that they might use the app further.

### Sleep duration Assessment (RQ1)

Unfortunately, the HedgeHog sensor did not always record all sleep information during the study due to a variety of issues (e.g., battery failures, waterproof problems, etc.). These technical shortcomings reduced the amount of the data available for comparison. Therefore, we considered only the sleep duration datasets calculated by the app on the days that the HedgeHog device successfully measured the sleep duration. This resulted in 20 data-pairs. The statistical analysis, interestingly, revealed that on average the duration obtained from HedgeHog ( $Mean=8.12$ ,  $SD=1.46$ ) was not significantly differed from our app ( $Mean=7.98$ ,  $SD=.87$ ),  $F(119)=.22$ ,  $p=.64$ ,  $r=.44$ . The start and end time of sleep collected by the sensor and the app also significantly correlate, start time:  $r=.57$ ,  $n=20$ ,  $p<.05$ , end time:  $r=.58$ ,  $n=20$ ,  $p<.05$ . No notable individual differences were found in the dataset.

**Self reflection and behavior change (RQ2)**

Based on the qualitative feedback obtained from the questionnaires and interviews, all participants found the analytical chart as the best feature of the app in comparison with other available features such as sharing posts in Facebook, setting an alarm by the countdown timer, or automatically activating the silent mode. This finding is supported by the fact that participants checked the chart on average once per day during the study. They reported that the chart increased their awareness about their own sleep habits and helped them to track their sleep behavior:

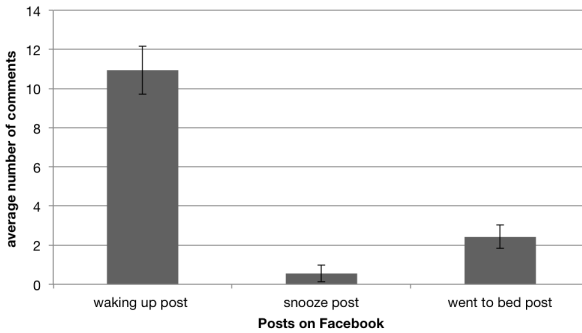
- P3: “With the chart I was able to check if I slept enough during the last days.” (P4 and P7 also gave similar feedback as P3).
- P2: “I tried to keep my sleep duration around 7 h by using the chart.”
- P5: “I [found out] that I didn’t get enough sleep during the week, so promised myself to sleep longer on Sunday.”
- P7: “I never thought such feedback could encourage me to think about my sleep behavior.”

Interestingly, the users reported increased awareness of their sleep behavior, which in turn induced them to change in their sleep behavior; however, the analysis did not convey any significant difference in the users’ sleep patterns.

**Sharing sleep information on Facebook (RQ3)**

We analyzed the Facebook comments received on different posts, i.e., went to bed, snooze, and waking up posts sent by the app. Figure 5.6 depicts average number of comments for the different posts. The results showed the 347 posts received 138 comments in total (Mean = 17.25 comments/message, SD = 14.99). The waking up posts received more comments (Mean = 10.93 comments, SD = 13.31) on Facebook than other posts sent via the app, i.e., snooze (Mean = .56, SD = .74) and went to bed posts (Mean = 2.43, SD = 2.8). In the following we mention some quotes from the dataset:

- P1: “Good Morning”
  - Comment: “Such a long sleep”
  - P1: “why not once in a while!”
- P6: “Good Morning – My sleep was natural”
  - Comment: “mine not:-(”



**Figure 5.6:** Average number of comments received on posts sent on Facebook using Somnometer app.

- P2: “Good Morning”;
  - Comment: “Oh! You are awake, call you now”
- P8: “Good Morning Germany”
  - Comment 1: “Good Morning Japan”
  - Comment 2: “Good Morning Canada”

Further, we grouped the posts and looked at their comments. The analysis describes that, interestingly, the posts that included negative ratings, indicating bad or very bad sleep, received 45% more comments. These comments were mainly concerned with why the user had experienced a bad night sleep and if there was something wrong (P5: “Good morning Facebookers! – My sleep was very bad!”; Comment: “why? probably because you woke up too early on weekend;-)”).

On average participants had 148.5 friends ( $SD=45.9$ ) on Facebook. The users chose to share their posts with 47% of their friends on average ( $SD=68.6$ ). Six out of eight participants had customized the privacy setting of the messages shared on the Facebook using the app feature. When they were asked with whom they shared the sleep data on the Facebook, participants stated that they mainly shared this information with their partner, best friends, family members, and their colleagues. The participants mentioned that they wanted to share the sleep status only with individuals they know well and with whom they are in contact more often as this information might be important for them (P1: “I shared my status with one of my classmates as we work together on [a project]”,

P6: “*I shared the data with my mom to call me when I am awake*”). Meanwhile, they all expressed that sharing this information could invade their privacy.

During the interview, the participants were asked whether sharing the sleep information with friends in social networks impacted their awareness and connectedness. Seven participants experienced the change of awareness among their friends:

- P2: “One day I was absent in a lecture. One of my friends checked my Facebook wall and found out I was still asleep. He sent me SMS and reminded me that I missed the lecture.”
- P5: “My partner told me that before she called me, she checked on Facebook if I am awake.”
- P6: “I saw friends went to bed and I also went to sleep.”
- P8: “My friend asked me why I had a bad sleep last night as she knew it from Facebook.”

However, they mentioned that frequently sharing this information might irritate their friends. We observed this theme in the analysis of Facebook posts, too (P1: “*Good morning fans! – My sleep was very good*”, Comment: “*Do you plan to rate your sleep and share it every single day?:/*”).

### 5.4.3 Discussion

The results of the controlled study suggest that it is feasible to monitor a user’s sleep duration based just on their interactions with an alarm clock app on the mobile phone. Previously, such sleep monitoring had only been possible using specialized tools such as an actigraphy device. For people who already use an alarm clock app daily, the study suggests that monitoring explicit interactions with the app provided enough information to estimate the user’s sleep duration. The statistical analysis showed that the sleep duration acquired from the app is not statistically different than the data obtained from the *HedgeHog*. Thus, it might be possible to successfully monitor users’ sleep duration using just an application on a mobile phone (RQ1). Obviously, for more accurate recording, actigraphy devices are essential.

While asking participants to choose the best feature of the *Somnometer* app, all participants surprisingly ranked the sleep chart the highest, indicating that users are interested in tracking their sleep behavior. Providing a method of visualizing sleep behaviors is

necessary for such applications. Data collected from the interviews and questionnaires revealed that simply providing users with feedback of their personal sleep behavior has the ability to persuade users to think about their sleep behavior and start engaging in healthier behaviors (RQ2). Though, no statistical evidence was revealed in the study. This result aligns well with past work on encouraging exercise [39]. It should be noted that one's sleep behavior can be impacted by many other factors, for example, diet, environment, daily activities, etc., which are not taken into account in this study.

The results further revealed that sharing sleep information with one's social networks impacted awareness among friends (RQ3). Users indicated that they would like to share this information with their social networks. However, users were concerned that sharing sleep information on Facebook had the potential to invade their privacy.

Specifically, our participants expressed that they want to share sleep information with individuals with whom they already have a high degree of social contact as well as with those who they know will find the information useful, e.g., partner, family, close friends, or colleagues. Our participants, however, expressed that they do not want to share their sleep information with the whole community, as they believe that doing so has the potential to annoy, frustrate, or otherwise bother their friends and colleagues. As such, providing a means to easily administrate with whom the sleep information can be shared is crucial. This can be adequately addressed with recent privacy features in the social networks such as Google+'s circles. Analyzing extracted messages shared via the app and their comments from Facebook revealed that posts with negative ratings received more comments. This supports the idea that friends were concerned about each other and curious to discover the reasons for the low sleep ratings. "Good morning" messages also received more comments compared to other messages shared in Facebook via the app. This increase in comments might be explained by the fact that the friends knew the user was awake and thus, they knew that they might be available to react and respond to their comments. Consequently, this behavior led to an increase in interaction and a feeling of connectedness among friends.

It should be mentioned that controlled studies have certain limitations. Similar research uses sample sizes that may be small by general standards. Furthermore, the participants used this system only during the study. Long-term, voluntary usage may reveal other information. The participants in our study were all male; hence, there is a probability that the results are gender biased.

## 5.5 In-the-wild Evaluation

With the first study, we wanted to capture users' natural interaction with the application. This is problematic in a tightly controlled laboratory-style user studies with a small sample size. As such, we decided to develop and release the app on Google Play, the official Google Android market, and let users download and use the app for free over the course of six weeks. We conducted the in-the-wild study in order to capture emergent user behavior in an effort to better understand how users engaged with the app. By using the Android market, we were able to reach many users and rapidly push new updates to users.

### 5.5.1 Procedure

We published the *Somnometer* app on July 7, 2011 and promoted the app by announcing it via mailing lists, forums, and social networks. We completed data collection on August 18, 2011. The app logged all changes in sleep state and any posts were sent to the Facebook via the app in a remote central database. As this was an in-the-wild study, there was no manipulation of the variables or features provided by the app. Every participant had the same version of the app and were instructed how to use the app in the same manner. Doing this gave us the ability to observe users' natural behavior. For example, we let each user decide whether or not they wanted to share information with Facebook via the app.

### 5.5.2 Results

According to the Android market portal, the app was downloaded 725 times during the six weeks study. Based on the Flurry portal, we accrued a total of 3522 sessions of usage (median 2.5 sessions/day). A session was determined to be one use of the application by an end user that typically began when the app was launched and ended when the application was terminated. Furthermore, 55% of users had used the app for only 1 or 2 times over the six-week study period. This left 45% of the participants who used the app more than 3 times.

Based on our database, 173 unique users had set an alarm and at least tried the app, where 166 of them had shared a sleep status message and 165 of them had also shared a sleep rating on Facebook via their use of the *Somnometer* app. In total 10 out of 166 users who shared a sleep status message also customized privacy settings. The users had

on average 258.4 friends ( $SD=242.3$ ) on Facebook. Interestingly, only seven individuals chose to share no information on Facebook. Ten users customized the privacy setting and chose to share with only 18% of their friends on average.

Regarding the qualitative feedback, 120 unique users answered the optional survey (72.5% male, average age 29.3 years), with 62% of whom often or always using the alarm clock feature of their phone. Also 85% of all survey participants used the alarm clock as a wake up alarm and 61% as a reminder. In total 454 alarms were scheduled, 86% of which were set by entering a specific time. The mean duration between setting an alarm and pressing the status button ( $|t_{alarm,et} - t_{bed}|$ ) was 37.6 min. Fifty-five users also checked the sleep chart at least once during the study (Mean=4.8 times,  $SD=1.9$ ).

Since the study occurred in an uncontrolled environment, it was not uncommon for a user to use the app only once or to drop out during the study. To cope with this, we removed data from users who used the app only once or had no night sleep session (a sleep session is the time from when alarm is defined until that alarm is deactivated and the time between the two events is greater than 2 hours ( $d_{alarm} > 2h$ )). This resulted in 268 sleep sessions from 93 users for the analysis. We investigated the sleep duration and rating behaviors between the users who shared data in Facebook vs. those who did not share. A one-way analysis of variance revealed significant differences with regard to their sleep duration ( $F(1266)=3.92, p<.048$ ), but not on their rating behaviors ( $F(1266)=3.35, p>.82$ ). While the average sleep duration of those who shared was longer ( $n=151$ , Mean=7.67 hours vs.  $n=117$ , Mean=7.34 hours), the average rating was lower than those who did not share data in Facebook ( $n=151$ , Mean=3.32 vs.  $n=117$ , Mean=3.46).

Similar to the controlled study, the Pearson correlation analysis indicated a positive correlation between the sleep duration and rating ( $r=.13, n=268, p<.03$ ).

### 5.5.3 Discussion

Almost one hundred participants actively engaged with the application and used the app regularly. Indeed, they demonstrated an interest in tracking their sleep habits and sharing sleep information online with their friends through social networks. The results also revealed that users who shared data with their friends tended to sleep longer. But, interestingly, they rated their sleep lower. This might be a way to redirect friends' attention toward themselves. The correlation between sleep duration and rating was similar to the controlled study. The users recruited from the wild also customized the privacy setting of their posts shared in Facebook. This shows that it is important for users to customize with whom this information should be shared. Unfortunately, due to the



privacy issues we were not able to gather information about comments on the posts sent via the app from the users recruited through the app store.

While studies in-the-wild provide the opportunity to test a system out of a laboratory setting with many users, they are conducted in an uncontrolled environment and, thus, have certain limitations (cf. Section 2.4.3). It was inevitable that some users would download the app without using it, use it infrequently, or opt out of participating in the study [174]. Logging users' interactions and behaviors could also potentially dissuade some users from engaging in the study, as they might not be willing to share this information. Furthermore, it is very hard to get information about participants and their context. Therefore, there are uncertainties whether the participants would be representative for a different age range.

## 5.6 Implication

Based on the controlled study, we determined that it is possible to monitor users' sleep duration using just an application on the mobile phone instead of needing to rely on using wearable actigraphy devices to retrieve contextual information. We also demonstrated that sharing sleep information with social networks impacts feelings of awareness and connectedness among friends. It was important for participants to share their sleep data with people who would find the information valuable and useful. Further, sharing on Facebook had a significant effect on the sleep rating. Specifically, users rated their sleep worse when they shared it on Facebook. One reason could be to redirect friends' attention toward the user themselves.

The qualitative results reveal that providing a means to track and visualize one's sleep habits impacts knowledge of sleep activity. The increase of awareness on the sleep activity can lead to encourage healthier sleep behaviors. However, we did not have any evidence to support this in our quantitative data. Participants' feedback indicates that users would indeed be interested in sharing their sleep charts (P2: *"I'm really interested to compare my sleep chart with friends using the app."* P3: *"It would be great if I could share the chart with my friends."*).

The investigation of the social alarm clock in-the-wild uncovered desires to share sleep patterns with other. However, privacy concerns pertaining to sharing intimate information on social networks. While in the controlled study most of the participants customized the privacy settings, only 10 users in-the-wild did which indicates an interesting privacy paradox. In a hyper-connected world, users' desire to connect with each other might include informing each other of minute and even intimate details of everyday life. Doing this correctly is difficult, especially in the face of users' privacy concerns.

## 5.7 Summary

Users share a large amount of personal information with friends, family members, and colleagues via social networks. Surprisingly, some users choose to share their sleeping patterns, perhaps both for awareness as well as a sense of connection to others. Indeed, sharing basic sleep data, whether a person has gone to bed or waking up, informs others about the availability of a person, but also indicates physical state and reflects a sense of wellness. In this chapter we investigated three research questions: (RQ1) whether it is possible to reliably monitor sleep behavior using simply a mobile phone; (RQ2) how providing users with the ability to track their sleep behavior could empower them to engage in healthier sleep habits; and (RQ3) the impact that sharing sleep information on social networks has on awareness and connectedness among friends.

To address our research questions, we developed a social alarm clock app for Android phones, called *Somnometer*. It helps users to capture and share their sleep patterns. While the sleep rating is obtained from explicit user input, the sleep duration is solely estimated based on monitoring a user's interactions with the app. By observing that many individuals currently utilize their mobile phone as an alarm clock, it revealed behavioral patterns that we were able to leverage when designing the app. We conducted two studies: a controlled study and a study in-the-wild. While the controlled study had its own limitations we simultaneously published *Somnometer* as a free application on the Google Play marketplace and conducted an in-the-wild study to capture natural usage behavior of individuals who had downloaded the app.

The result from a controlled study reveals that it is feasible to monitor a user's sleep duration based just on her explicit interactions with an alarm clock app on the mobile phone. The result of statical analysis between our approach and using a wearable sensor for measuring the sleep duration did not reveal any significant difference. The explicit interactions with an app can indeed be used as a source to obtain such contextual information. The results from both an in-the-wild study and a controlled experiment suggest that providing a way for users to track their sleep behaviors increased user awareness of sleep patterns. The qualitative feedback provided by participants show that the increase of awareness may induce healthier habits. However, we did not observe any change in our quantitative data. We also found that, given the current broadcast nature of existing social networks, users were concerned with sharing their sleep patterns indiscriminately.





# IV

SHARING CONTEXT  
NONVERBALLY



# Outline

As the contexts varied in computer-mediated communication, sharing and exchanging contextual information is essential for enhancement of this type of communication. A common approach to exchange and share information is verbal communication. However, information can be shared nonverbally, too. Indeed, communication of information in nonverbal ways has been the subject of research for more than 150 years. One of the most prominent examples is the invention of the Morse code in the early 1840s. The popularity of this rhythm-based character encoding system can be explained by its ability to be read by humans without any decoding device and its high learnability. In this part of the dissertation we investigate the feasibility of sharing contextual information nonverbally between non-colocated users. We leveraged two nonverbal modalities for imparting information: *melody* and *iconic user interface*. We are particularly interested in the mobile phone as one of the most ubiquitous devices for communication. Users are emotionally attached to their phone and use it in various contexts. We, therefore, leverage the mobile phone as a communication channel for nonverbal information sharing and exchanging.

We explore the melody composition as a means to express and share emotions nonverbally. Humans need to communicate and share their emotions with others. Users have utilized numerous techniques and technologies to maintain emotional connections. We leverage the short message service (SMS) as a form of mass communication to easily and quickly share melodies composed to create awareness of emotional feelings between non-colocated friends or partners. In the further step, we investigated how audio can influence the previewing of received messages. Visual clues or simple audio tones are approaches to inform the receiver about the messages received. Such notifications do not convey any information about the content or intention of a message. We aim at using audio previewing of messages content as means to transmit and communicate information about their content and intention.

On the other hand, the ubiquity of mobile phones and the Internet connectivity inherent on them provide the opportunity to share information between non-colocated users instantly in real-time and connect them together. We explore the feasibility of iconic user interfaces as a nonverbal communication channel for sharing sentiments. In a use case, we investigate how TV viewers can nonverbally share emotional reactions to events shown live on TV. Considering TV as one of main sources of entertainment, watching TV does not necessary have to be a solitary experience. Non-colocated TV viewers can be connected through their phones and are, thus, able to share sentiments about ongoing events. Further, this information can be used to extract highlights based on sentiments and reactions provided and shared nonverbally.

This part of the dissertation is consist of following chapters:

- **Chapter 6 – Melody Composition for Sharing Emotion.** We explore the feasibility of melody composition as onverbal means to express and share emotion between remote users in this chapter. In their role as personal communication devices, mobile phones are the natural choice for sharing and communicating emotions. We use the SMS as one of most popular services on mobile phones for communicating melodies. We present a system that allows users to easily compose melodies and share them as a form of an SMS. We conduct a user study to assess this approach for expressing and sharing emotions nonverbally.
- **Chapter 7 – Sonification Conveys Awareness.** In this chapter we assess the audio previewing (sonification) of text messages received to convey information and create awareness nonverbally. Current notification approaches such as visual cues or audio tones aim at solely informing the receiver about incoming messages without imparting any information about their content. In contrast, we focus on providing notifications to convey information about messages received. We examine audio previewing of messages as a means to communicate the content and reveal awareness nonverbally. We investigate the impact of the approach on users' behavior using mobile phones as well as desktop computers.
- **Chapter 8 – Sharing Sentiments with Iconic Interfaces.** We assess a nonverbal communication channel on the mobile phone as means to share sentiments between non-colocated TV-viewers and connect them together. We present a mobile application that leverages an iconic interface and allows users to express and share their opinions about live TV shows in real-time. The icons and graphical elements used in the interface represent various events relevant to the TV program users watching. The interface further allows users to quickly and with minimum effort share their sentiments at any moment. We conduct a user study in-the-wild with a large number of users to examine nonverbal sentiments sharing and their correlation with moments in TV shows.



# Chapter 6

---

## Melody Composition for Sharing Emotion

Humans are social creatures. They need to communicate and share emotions. Many interactive technologies designed for other purposes have been adapted for use within intimate relationships. Symbols such as flowers, photographs, or love letters have long been used to share emotions between people closely connected to each other. People use numerous techniques and technologies to maintain an emotional connection. Webcams, emails, instant messaging, and blogs are examples of such technologies used as mediators of human interaction for sharing and maintaining emotion. In particular, people who are in distance relationships use these technologies during the physical absence of their partner in order to create a sense of presence-in-absence [101].

In computer-mediated text-based communication, for example, means for explicitly expressing emotions and feelings by abbreviations and symbols have been developed [48]. Hancock et al. [78] reported that emotions are readily communicated in text-based interactions through two strategies: (1) verbal strategies by using changes in disagreement, affect terms, verbosity, etc., or (2) nonverbal strategies such as the use of punctuation. Using emoticons is another approach to express the emotion in the text-based communication. An emoticon is a group of keyboard characters that typically represents a facial expression or suggests an attitude or emotion. For example, the emoticon “:-)” expresses

a happy emotion. Here, text-based or textual information is used to express emotions and communicate nonverbal information.

In contrast, we are interested in how emotions can be shared using non-textual information in non face-to-face communication. In this chapter, we assess how a melody as nonverbal information can be used for expressing and sharing emotions. Users are emotionally attached to their phones [193]. Hence, the usage has a twofold linkage between private and emotional aspects: (1) users engage with their mobile phones and (2) users use them for personal communication. Whereas it seem natural to use this device as a mediator for sharing emotions, currently it only states very basic and simple ways of deploying these devices for sharing emotional feelings. Apart from telephony service, the short messaging service is one of the most popular services available on mobile phones. The short message service (SMS), also referred to as text messaging, has become a form of mass communication since it provides a convenient way of exchanging textual short messages on-the-go. Ahonen reports that SMS is the number one service used on mobile phones as of 2013 [2]. Even though it lacks expressiveness, has confusing syntax, and is error prone [73], sending a text message via a mobile phone is still increasingly being used to forge new romantic relationships [27] and to coordinate with intimate friends [73].

We investigated how SMS as one of most popular mobile phone services can be used for sharing emotions. We use this service as a communication channel and leverage its capabilities to easily, quickly, and cheaply create awareness of the emotional feeling between friends or partners. This is achieved by providing a web-based music composer that allows users to quickly and easily compose a melody. The melody is sent as a short message and played immediately on the receiver's phone. When using asynchronous communication, such as SMS, the intention of the sender may be changed or get lost due to the difference in time between sending and receiving a message. Therefore, we believe that time is a crucial factor while sharing emotions and the emotion sharing should take place synchronously.

While the prior work explores sharing emotion and the remote expression of a person's feelings using (tangible) objects or verbal information, we investigate the impact of music as nonverbal information in the context of emotion sharing on mobile phones. The goal is to assess the self-composed music as a nonverbal communication channel for sharing information and its impact on users. Listening to different types of music can express the mood of a listener or their feelings about another person. To date, mobile music has been studied mainly from the consumption point of view [134, 196]. In the current chapter, we present how composing a melody as a unique piece of art can be used to share emotions.

The contributions of this chapter are twofold:

- We explore the impact of self-composed melody as a crafted piece of art for sharing emotions. Music is an important way to communicate one's state of mind and is often characterized as the language of emotions.
- We present a system that allows users to compose melodies and share their emotions as form of SMS in a synchronous manner.

*This chapter is based on the following publication:*

- A. Sahami Shirazi, F. Alt, A. Schmidt, A.-H. Sarjanoja, L. Hynninen, J. Häkkinen, and P. Holleis. Emotion sharing via self-composed melodies on mobile phones. In *Proceedings of the 2009 International Conference on Human-Computer Interaction with Mobile Devices and Services* (Bonn, Germany), MobileHCI '09, pages 30:1–30:4, New York, NY, USA, 2009. ACM

## 6.1 Related Work

A variety of research projects have explored various approaches for sharing emotions. Several projects have considered methods to represent and share emotion with tangible objects. An early work is the *inTouch* system that creates the illusion of manipulating the same physical object by distant users [20]. Furthermore, picture frames [32], wedding rings [204], or an entire bed [69] have been utilized as emotional communication devices between partners.

The *LumiTouch* system, which is designed as an asymmetric bi-directional channel of communication, enhances the symbolic power of a picture frame by providing a subtle real-time communication link. It consists of a pair of interactive frames. Touching one frame results the other picture frame lighted up. It focuses on communicating emotional content in addition to presence information [32]. A more intimate approach can be found in the Sensing Beds system. These are beds that mediate the needs of two physically distant romantic partners, who are not colocated, by sensing the body position in each bed and using a grid of small heating pads to warm the congruent point in the other partner's bed [32]. The “united-pulse”, composed of two rings, has been designed to share remote intimacy. Each ring can measure the wearer's heartbeat and transmit intimacy by vibrating the partner's ring [204]. Touching the ring also allows distant

romantic partners to share emotions and a small moment of intimacy. A further example is the Lovers' Cup. The project explores the idea of sharing feelings of drinking as a communication channel for a couple in physically different places [37]. The cups are used as tangible communication interfaces that support various interaction techniques such virtual kiss, hands shaking, and toasting.

Sharing emotions and other nonverbal information is an important aspect of interpersonal communication. Researchers have investigated how to enhance such communication, in particular, text-based (messaging) communication. *SenseMS* is designed to augment text messaging with contextual information and human embodiment and aims to provide richer messages by these means [7]. It is reported that the augmentation can result in more pleasant experience for the sender and receiver. *ExMS* is another message system that allows users to concatenate and annotate avatar animations and send them to peers [148]. The user study reveals that the participants use the system to tell micro-stories. Another similar type of application, *Comeks*, enables users to create comic strips as MMS, thus empowering a more expressive communication [169]. The interplay between text and animation allows users to create expressive messages. The use of emoticons also adds emotional expression to text-based communication. Walther and D'Addario investigated the affect of emoticons on message interpretation. They conclude that the emoticons have an impact, but their contribution is outweighed by the textual content [200].

Scent is another medium for communication. The sense of smell has a close link with memory and emotion. Neuroscience research has shown strong links between smell and attention, reactions times, and emotion [122]. Engen and Pfaffman investigated the number of smells subjects could sense [56]. They report that subjects are able to identify 16 smells at one time. Smell has been incorporated in various projects. Brewster et al. assessed smell for tagging photos [21]. They report that the smell can be used to aid photo recall. *Sound Perfume* is a system that augments face-to-face interpersonal communication with auditory and olfactory input [36]. Kaye developed a set of applications such as *Dollars & Scents* or *inStick* that uses scent to convey information [98]. Further, mobile phone manufacturers such Samsung [34] and Motorola [72] have showed the concept and patents pertaining to the scent emitting devices that are controlled by mobile phones. Inspired by related work, we investigate the feasibility of composing a melody and sharing it on mobile phones in order to express and communicate emotions.

## 6.2 Design Rationale

Offering crafted gifts, such as a hand-made birthday card or item, instead of an off-the-shelf one, is a powerful way of expressing emotions and feelings towards a beloved

person. Similarly, self-written (love) songs have the power to deeply touch another person. We expect that applying the concept of craft tradition to emotion sharing via the mobile phone strongly influence the connection between two persons.

Our idea is to encourage users to compose music themselves, instead of using previously composed songs or ringtones. This approach creates an achievement similar to a crafted item. This strongly influenced our design decisions on both the sender and receiver sides. In the following section, we discuss our design rationale for developing our prototype.

### 6.2.1 Sender

We believed that any person could create a unique piece of art, such as a self-composed melody, without having any formal knowledge about the notation of music. To support this, we provide a web-based interface that allows users to compose and send a melody to a mobile phone. The following design principles influenced the development of the music composer. With these guidelines, we were able to create a system that provides a balance between the ease of use and a high level of expressiveness.

#### **Ease of use**

It is not an easy task to write songs based on complex chord patterns and several voices, as it requires certain knowledge. Nevertheless, a basic melody that expresses sadness, happiness, excitement, or longing can be created through trial and error without knowledge of major/minor tonality and notation of notes. The composer interface was designed in such a way that it allowed users to compose melodies from 32 tones in an average time of 30 to 60 seconds, which is comparable to the time required for writing an SMS.

#### **No learning required**

Learning an instrument such as a piano requires a large amount of effort, training, and talent. Our composer interface does not require any learning due to its simple representation of the notes. It provides 32 quarter notes on the y-axis and 8 tones (which are in accordance with a C major diatonic scale) on the x-axis (see Figure 6.1). Breaks between the notes are set simply by not selecting any note.

#### **Smooth creation flow**

Crafted pieces of art such as paintings or hand-written poems are unique due to the flow in their creation process. Similar to a signature they are not a combination of predefined

patterns as we normally see in the digital world, but rather include the current mood of the composer. We tried to simulate this by letting the user compose a melody, not by individually selecting tones (which is nevertheless possible as well), but by smoothly moving the mouse cursor or the pen on a touch screen over the composer interface. Further, we decided not to provide any option to type in the notes using the keyboard. We deliberately choose explicit and nonverbal interaction to compose a melody.

### **No means to store the melody**

Uniqueness is one of the most important properties of a piece of art. This is difficult to realize in the digital world, since everything can be stored and duplicated. We deliberately decided not to provide any means for storing, nor replaying the created melody, nor include previously composed melodies in order to preserve inimitability. Hence, we can match the moment a melody is composed to the emotion of the composer.

### **Full control of the composing process**

Our system provides users full control of the entire composition process. Thus, they can decide when, where, how, and with whom they want to share their emotions. By following these design guidelines, we were able to create a system that supports a balance between ease of use and a high level of expressiveness.

## 6.2.2 Receiver

Mobile phones are designed to be carried wherever the users go and, thus, have the potential to be used as objects that can share emotions. Consequently, we provide an interface to send the composed melodies to a mobile handset. To avoid generating additional costs for the receiver or rely on data connections, we decided to transfer the melodies using the standard SMS channel. In the following section, we address important principles related to emotion sharing such as time constraints, interaction, and control for the receiver.

### **No interaction required**

Traditional forms of digital messages, such as e-cards, emails, as well as SMS, require interaction by the receiver such as opening a certain URL, an email program, or the inbox of his mobile phone. We believe that the time an emotional message is received is a key factor and carefully chosen by the sender (similar to the time when performing a self-written love song). Therefore, the system is designed without requiring interaction

on the receiver side. The incoming melody is detected and played automatically for the receiver.

### **Giving control away**

Allowing a person to send a message at any time requires the receiver to give up control on when to read or listen to a message. However, we consider this to be a sign of trust between friends, which even increases the power of sharing emotions via a mobile phone. Nevertheless, we carefully consider the current configuration of mobile phones to display incoming messages. Respectively, we also consider the current configuration of the phone to play the melody. For example, if the phone is on silent, the melody is not played.

### **Moving from notification to content**

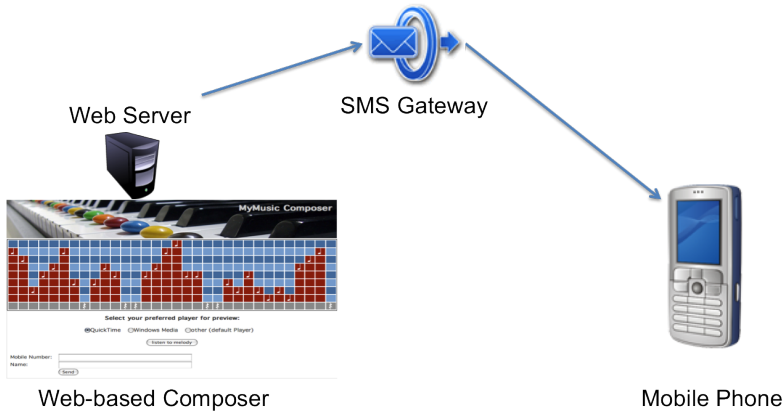
Digital communication is often asynchronous and several mechanisms exist to notify users of new messages. Those mechanisms include visual clues, such as a letter symbol indicating a new email, or audio tones, such as a notification. In our system we use this mechanism to not only indicate that a message is received but also deliver the content of the message. Hence, it is possible to ensure that a person receives a message as close as possible to the sender's intended time.

### **Limited duration**

One important aspect of emotion sharing via the mobile phone is the duration of the message. In order to maintain the sender's intention, the message is immediately delivered. However, the receiver might be busy at the time of receipt. Yet, we believe that the time required for listening to the melody, which is comparable to checking a SMS received, is short enough not to disturb the receiver and long enough to express the emotion of the sender.

### **No means to replay/store**

Similar to the sender side, we do not allow the store or replay the melody on the receiver side. We, thus, avoid that the receiver ignores the message the first time it is played (because it could be replayed later) which increases the awareness of the sender's emotion at that particular moment in time.



**Figure 6.1:** The system architecture consists of a web-based composer, an SMS gateway, and an application for Symbian mobile phones.

## 6.3 Prototype

Based on the design rationale described in Section 6.2 a prototype, called *EmoShare*, is developed. The system consists of two main components (Figure 6.1):

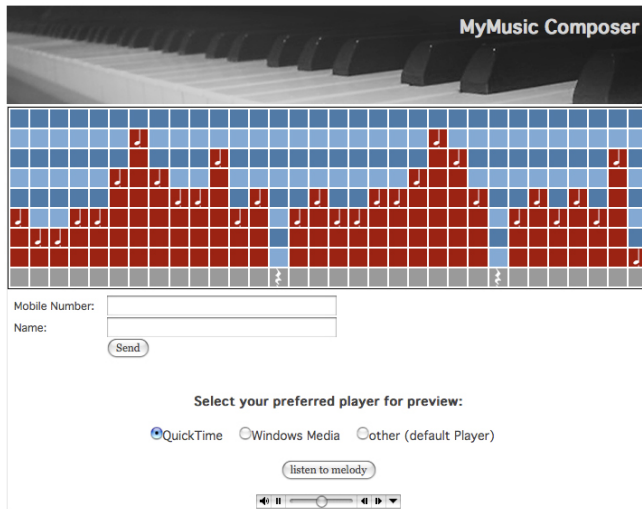
- **Composer:** is a web-based composer hosted on a web server which has the functionality of sending a SMS to any mobile numbers over GSM networks.
- **Music Player Application:** is a JAVA-based application that detects incoming melodies and plays them automatically.

In the following section, we describe each component in detail. The user can compose melodies without installing any specific applications by simply using a web browser and sending the composed melodies as short messages to the recipients. On the receiver side, the application automatically detects the melodies as incoming messages and immediately renders them.

### 6.3.1 Composer

The composer is an AJAX application that allows users to create a melody of 32 quarter-notes. Figure 6.2 depicts the composer's interface. Each note is represented by one





**Figure 6.2:** The Composer Interface: Melody Creator (top) and Preview and sending the melody (bottom)

column in the composer and can be assigned a value between 0 and 8 where 0 represents a crotchet rest, and 1 to 8 is mapped to a C major diatonic scale. To support a flowing composition, the application implements an *onMouseOver* listener, which is activated upon holding down the mouse button. Hence, the melody can be created not only by selecting each single note, but also by smoothly moving the mouse cursor over single notes.

When a melody composition is finished, an XML file is generated. The XML file includes the pitch and length of each note in Midi conform format. To provide users with the possibility to listen to the melody composed, the composer application translates the XML file into a midi-file locally and loads it into any media player such as QuickTime or Windows Media Player upon clicking the “listen to melody” button (see Figure 6.2).

To send the melody to the receiver, the web application establishes a connection to a SMS gateway. It translates the XML file into a string and sends it as an SMS to any mobile number. However, the receiver needs to have the mobile application in order to play the music. Then, a link to the required mobile application and a short description are sent to the receiver in the first SMS. On the receiver side, this string is decoded to generate and play the melody.

### 6.3.2 Melody Player Application

To play the incoming melodies on the mobile phone, a JAVA-based application was developed. The application runs in the background and listens to incoming messages on a specific port. Users are only required to install the application once on their phone and then leave it running in the background. When a message including notes is detected, the notes and the sender's number are extracted. From the extracted notes, a midi melody is generated and directly played. Information about the sender is also shown on the screen. The melody is only played once and there is no possibility to store or replay it.

The Mobile Media API available in JME (Java Micro Edition) is used to generate the melody. Additionally, the player takes the current profile configuration of the mobile phone into account. This means that if the silent mode is activated on the phone, the incoming melody is not played, but rather the sender's number is shown on the screen.

## 6.4 User Study

We conducted a user study to evaluate the prototype and assess the emotion sharing through this approach. We recruited 12 participants (6 male) ranging from 23 to 34 years ( $SD=3.4$ ). Five of them showed musical interests and one of them was a professional composer.

### 6.4.1 Procedure

The study itself consisted of an initial questionnaire, an interview, two tests, and a couple of post-hoc questions about the tests. The initial questionnaire included questions about customizing the ringtone and the SMS notification tone on their mobile phone. During the interview, the participants were asked about their text and multimedia messaging behavior and whether they felt that the current communication methods provided by their phones were sufficient for expressing feelings and emotions.

After the interview, the participants continued with two tasks. In the first task, they composed and sent a musical message to one of their friends using the composer. It was explained that, at this point, the messages would not be sent to anyone in order to create a more relaxed situation for testing. For the second task, we provided them with a mobile phone with the player application installed and asked to imagine a scenario. They were asked to imagine being at home and a melody from the opening titles of

“Looney Tunes<sup>14</sup>” was played for them. After that, they were asked to provide us their thoughts about the musical message and why one of their friends could send this kind of messages to them. We asked whether they would like to answer the message and, in case of positive answer, explain how they would want to replay to such messages.

At the end, few additional questions assessed the participants’ willingness to use this kind of service. The participants were also asked to score ease-of-use and how much they enjoyed using the application on a 5-Point-Likert scale (1=very hard to use, 5=very easy to use). The same scale used to assess how much they enjoyed using the application (1=didn’t enjoy at all, 5=completely enjoyed).

The participants used the composer on the Firefox browser and their performance was recorded with a screen capture program. The experiment and the interview took approximately 30 minutes for each participant.

## 6.4.2 Results

Based on the result of the questionnaire, nine participants expressed their feelings through the phone and only two of them rarely conveyed feelings. Eleven of them shared their feelings by writing text messages and seven by calling. In addition, five mentioned that they use emoticons as feeling indicators. All of the participants felt that they are able to express their feelings with their phones. However, three believed that there is room for improvement. One participant mentioned that sending multimedia messages could be easier than using the application. The initial questionnaire also revealed that eleven participants had already changed the default ring tone of their phone and eight had changed the SMS notification tone.

### *Composing a Melody*

Each participant was able to complete the first composition task in less than seven minutes (including learning the interface, composing, listening to the melody, re-composing, and sending). The shortest one needed approximately two minutes. However, the talkativeness and the participants’ willingness to modify the melody affected the process significantly. The maximum time for composing a melody consisting of only 32 notes and took on average 57 seconds. This demonstrates that the time required to learn how to use the composer interface is quite short.

The average response for the ease of use was 3.75 and for enjoying the use of the application was 3.1. Two of the participants did not like the composer’s layout. One

---

<sup>14</sup> Looney Tunes is a Warner Bros. animated cartoon series

participant considered the composer not to be powerful enough. Another participant was confused because it did not provide any manual. Overall, seven of the users stated that they would not use this kind of service, four indicated would try it, and one, who was a professional composer, said would definitely use it.

### *Receiving a Melody*

All of the participants could describe situations in which this kind of musical messaging could be used. These included amusement, cheering, creativity, sharing memories, and sharing feelings. Overall, two out of twelve participants stated that they definitely liked the idea of receiving music and nine out of twelve were willing to reply to the message using arbitrary ways of communication. Four participants stated that they like to reply the message via a musical message.

As an additional comment, one participant stated that she thought not to be musical enough to use the application and another one thought that she was too lazy to use it. Further concerns were related mainly to the quality of the output and the loss of control. Five of the users had doubts about how the melody would sound on the receiver's phone and five believed that their personalized settings should not be touched.

## 6.5 Implication

According to the user study, on the composer side, users tend to have high subjective requirements for the piece they composed. We reveal that users want messages to be something aesthetically pleasing – something beautiful or cute, which you would dare to send without feeling embarrassed. Creating a musical message, which satisfies the sender is not easily achieved. This should be taken into account when designing such musical user interfaces, and can be incorporated, for example, in selecting the chords, tones and echo.

On the receiver side, the results of the study reveal that most of the participants can imagine a situation in which they would receive a musical message from a friend or a partner. Yet, several concerns were revealed. Predominant concerns are the misinterpretation of the message, social embarrassment, and a feeling of lack of control (e.g., when expecting the phone to stay silent). These aspects illustrate that users need to stay in control over their mobile device. However, we think that this is an opportunity, rather than a flaw of the system, which can be seen as means to indicate trust to friends or partners.

Overall, the findings suggest that the composer can be seen as a tool for three different functions: sharing emotions, creativity, and fun among friends. Composing melodies

is seen as a tool for creating jokes, sending funny things, or even teasing. Here, the technology can support group cohesion. For sharing moments, the music provides different opportunities. It can be used for provoking memories from something or somewhere, which the sender and receiver have experienced together, or describing the current atmosphere.

## 6.6 Summary

People have used various interactive technologies designed for other purposes to express and maintain emotions. In their role as personal communication devices, mobile phones are a natural choice for sharing and communicating emotions. However, their functionalities are currently very limited in power to express affective messages. Emoticons is one strategy for expressing emotions in text-based communication. Above all, textual information is used for expressing emotions and communicating awareness nonverbally.

In this chapter we assessed the feasibility of sharing emotions using nonverbal means, i.e., melodies. We explored melodies as a means to express and share emotions and feelings for synchronous non face-to-face communication. To achieve the goal, we developed a system that allowed users to easily compose melodies and synchronously share them in form of short messages (SMS). The prototype consisted of two parts: a composer and a melody player. Through a web-based composer users could nonverbally interact, compose a melody, and send the melody as a form of SMS to the receiver's mobile phone. The composer application was designed in such a way that users without any knowledge about music could compose a melody through trial and error. On the receiver side, a mobile application monitored incoming messages, detected melodies, and automatically played them.

We conducted a user study with eleven participants to evaluate the goal. The user study revealed that such a form of communication could be used for expressing and sharing emotions nonverbally. Further, it can be used as a means for creativity, fun, and teasing that results in-group cohesion. The composed melodies could be used for sharing moments or describing a situation. Self-composed melodies have a stronger impact than previously composed or downloaded messages, similar to crafted pieces of art offered to a beloved person. The composer was considered to be a tool for describing the current context and provoking memories of situations in which both the sender and receiver had experience together. With a long-term user study, it would be possible to gain insights into the impact of such emotion sharing and create awareness among non-colocated users.



# Chapter 7

---

## Sonification Conveys Awareness

Communication is the main purposes of using mobile phones. The emergence and advances in services and applications on mobile phones allow users to communicate through different communication channels such as text messaging, emails, etc. Visual clues or tones are common notification mechanisms used on the mobile phone to make receivers aware of incoming messages. Synchronous communication tools (e.g., chat clients) use mainly visual clues (e.g., highlighting the application's window). In addition, asynchronous communication tools (e.g., email clients) often make use of audio notifications. However, such notifications neither convey the content, nor the intention of a message. We are particularly interested in how such notifications can be provided in a way that the content of the message is conveyed nonverbally. Sonification is an approach that uses non-speech audios to convey information. It aims to translate relationships or information in data into sounds that exploit the auditory perceptual abilities of human beings such that the data relationships are comprehensible.

The ubiquity of the mobile phone allows users to use their phones in different contexts. The short message service (SMS), as a form of mass communication, provides a convenient way of exchanging information. However, there are situations where users are engaged in other activities or it is difficult or inappropriate to check an incoming message immediately and are made aware of the message arrival by the notification tone. The

notification tone solely notifies the user without revealing further information about the message received. We believe that if the user is made aware of the type of message received, it would lead to a change in the reading behavior. For example, users might want to answer a text message containing a question immediately, whereas, in other cases, they may check messages after finishing their current activity. Prior work have used the rich tactile output as a modality for conveying information [25, 166]. In contrast, we use an abstracted audio preview similar to a notification tone as a means to reveal the content of a message.

In this chapter, we investigate how the sonification of text messages can facilitate conveying message content. To achieve our goal, in the initial step, we explore how providing simple abstract audio previews of SMS messages can influence the reading of the SMS and the writing behavior of users. We conduct a survey and assess the content of the SMS users received. Based on these findings, a prototype is developed that intercepts incoming messages on the phone and plays a tone based on their content. The tone represents some simple and abstract indications on the content of the SMS. Through conducting a controlled user study, we evaluate the feasibility of this approach. In the second step, we assess how a notification can be used to convey more detailed information about a message's content such as its intention, the keywords included, or the precise words. We present an algorithm, which generates a musical representation of a message in such a way that its intention is indicated to the user. At the same time the privacy is preserved, as compared to reading the message out loud.

The contributions of this section are as follows:

- We present how audio previewing of messages can impart information. We further discuss how the use of this mechanism affects users behavior in writing and checking text messages on mobile phones as well as on personal computers.
- We propose an algorithm for the transformation of text messages into euphonic melodies in such a way that the intention of a message can be communicated without reading the message.



*This chapter is based on the following publications:*

- A. Sahami Shirazi, A.-H. Sarjanoja, F. Alt, A. Schmidt, and J. Häkkinen. Understanding the impact of abstracted audio preview of sms. In *Proceedings of the 2010 International Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA), CHI '10, pages 1735–1738, New York, NY, USA, 2010. ACM
- F. Alt, A. Sahami Shirazi, S. Legien, A. Schmidt, and J. Mennenöh. Creating meaningful melodies from text messages. In *Proceedings of the 2010 International Conference on New Interfaces for Musical Expression* (Sydney, Australia), volume 10 of *NIME'10*, pages 63–68, 2010

## 7.1 Related Work

A strand of research has used the sonification approach as means for communication of information in nonverbal ways. In the sonification approach, the goal is to use non-speech audio to convey information. Sonification is a subset of auditory displays. An auditory display uses sounds to communicate information and aims to enable better understanding of the data that underlie the display. Particularly, blind and visually impaired users can benefit from this approach. *MUSART* is a sonification toolkit, which produces musical sound maps to be played in real-time [95]. Walker et al. presented the *Audio Abacus*, an application for transforming numbers into tones [199]. *Babble Online* sonifies browsing activity, trying to communicate information both clearly and in a well-composed and appealing way [79]. The *Sonification Sandbox* allows users to create auditory graphs from several sets of data [198]. Song et al. presented mapping strategies derived from an analysis of various sound attributes, allowing to better represent and access information from complex data sets [183]. Petrucci et al. showed how to use sonification in auditory web browsers to allow visually impaired users to explore spatial information by means of an audio-haptic interface [150].

Audio cues are also used to provide information while interacting with computer [187, 186]. Earcons are auditory icons used in computer interfaces to represent part of an interface and provide information to users [14, 64]. Brewster studied the use of earcons and evaluated whether they provide effective means for communicating information [22].

Several research projects focused on the sonification of synchronous and asynchronous messaging. Specially, in instant messaging communication, research has investigated the

usefulness of providing audio cues for blind users to effectively receive messages. *QnA* estimates the type of an instant message, e.g., whether it is a question or not, and changes the notification mechanism accordingly [10]. The results indicate that modifying the nature of the notification can create a benefit for the user. *Hubbub* is a sound-enhanced mobile instant messenger aiming at increasing background awareness by providing audio clues [94]. Issacs et al. reported that the system helps people feel connected and support opportunistic interactions [94].

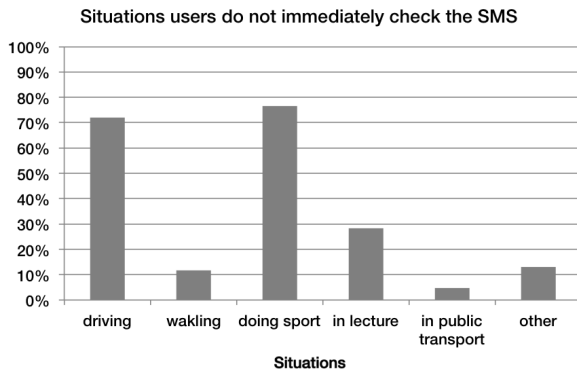
Following similar approaches, we assess an abstract preview of incoming messages' content in mobile phones. We investigate how such sonification can influence the user experience in the mobile context. In contrast to related work, we focus on providing auditory clues by transferring very small chunks of information or the complete content of a message.

## 7.2 Assessment of SMS Usage

Several prior studies have investigated where, when, and for which reasons text messaging is used, e.g., [109] and [211]. Users face several limitations when writing short messages. First, space is scarce and limits the amount of information that can be transmitted. Second, SMS lacks the expressiveness and support for nonverbal communication. These issues lead to the evolution of a distinct language for text messaging [90], characterized by the use of abbreviations, acronyms, and emoticons. Such elements seem to be suitable for defining the types of messages. To gain more insights on common emoticons used in SMS messages we conducted an online survey.

### 7.2.1 Setup

We conducted an online survey in which we assessed the users' behavior with regards to writing and receiving a SMS. The survey was divided into three main parts. First, we assessed general information about the users' SMS behavior (number of SMS, communication partners). Second, we were interested in the users' behavior when receiving an SMS in different situations to understand in which situations they check the messages immediately and in which they do not. The situations were being at home, in public transport, in the office, while driving, and doing sports. Third, we asked them to analyze their last 10 SMS received and provide us with the following data:



**Figure 7.1:** The results reveal that users do not immediately check the SMS while driving, doing sports

- the first word of each SMS
- the number and type of emoticons included in the messages
- the number of question marks

The web survey was an open call announced via mailing lists and social networks, such as Facebook. It did not target any specific group and ran over three weeks in the spring 2009. It was available in four languages namely English, Finnish, German, and Spanish. It took approximately 10 minutes for each participant to complete the survey.

## 7.2.2 Results

In total 347 participants, 46.1% female with the average age of 29.83 years ( $SD=8.2$ ) answered the questionnaire. The participants were from 21 different countries, mainly Germany, Finland, and the United States. They had various backgrounds, e.g., high school or college students, or employees with different academic and vocational backgrounds. The majority of the participants wrote and received on average more than 10 SMS weekly (60.4% senders, 65.6% receivers). Their predominant communication partners were friends, family members, partners, and colleagues.

**Table 7.1:** Top three emoticons that are used in the short messages.

	Emoticons			
	:-) or :)	;-) or ;)	:-( or :(	others
<b>Use percentage</b>	85.03%	62.25%	62.71%	31.41%

**Table 7.2:** The most frequent keywords that the short messages are started with. The keywords are not case-sensitive.

	Keywords			
	hi,hey, hei, or hello	ok	yes	no
<b>Appearances</b>	573	197	135	108

### SMS checking behavior

We were especially interested in situations in which the users checked or did not check the message immediately after they were notified of an incoming message by the notification tone (multiple selection was possible). We discovered that 87.0% of participants check their SMS immediately if being at home, 79.9% in public transport, and still 65.3% in the office. However, more than two thirds of the participants did not check on SMS immediately while driving or doing sports. Figure 7.1 depicts the situations users do not immediately check the SMS.

### SMS analysis

We analyzed the text messages for the use of keywords, emoticons, and punctuation marks to find out whether the type of message could be easily determined. Overall, 90.9% of all participants used emoticons in their SMS. Table 7.1 shows the most popular emoticons. For the question marks, on average 25.4% of the analyzed SMS (approximately 3400 messages) contained at least one question mark. From analyzing the first word of each SMS the most frequently used words were greeting phrases such as *hi*, *hey*, *hei*, or *hello* (573 appearances). Table 7.2 shows additional common keywords.

### Limitations of the survey

The set of SMS assessed was fairly large. However, the survey does not claim to be representative due to the fact that users are openly recruited online, or self-selected. This may have drawn in people who have more expertise with digital technologies than the

average user. Hence, they may not provide a perfect matching sample to the participants of the user study described in the section 7.3.

## 7.3 Audio Previewing of SMS

To evaluate how the abstracted audio preview is experienced and how it changes the text messaging behavior on mobile phone we conducted a four-week field trial with 20 users including seven couples (10 female) with an average age of 28 years ( $SD=3.2$ ).

### 7.3.1 Approach

The survey reported in Section 7.2 reveals that there are situations in which users prefer not to immediately check incoming messages. Further, the survey shows that the scarce space and lack of expressiveness for short messages leads to a widespread use of emoticons and abbreviations. Since emoticons are universal in many languages, we decided to represent them using nonverbal means, i.e., tones. If an incoming message includes certain emoticons or keywords, we map them to specific notification tones. Our hypothesis is that using nonverbal information as content presentation may change the users' behavior in writing and checking SMS. This change, further, results in more use of emoticons and phrases that are mapped to nonverbal information.

Our approach for mapping the message's content to nonverbal information consists of the following steps:

- The content of incoming messages is scanned for certain key strings, emoticons, and punctuations.
- Based on the expected meaning of the spotted key strings we select a specific tone indicating the message's type assumed .
- After the default notification tone the selected tone is played to inform the user on the potential content of the message. We discriminate between the following types of messages:
  - happy messages
  - sad messages
  - answers and responses
  - questions

It should be mention that a similar concept is commonly employed for incoming calls on most mobile phones when it comes to identifying the caller. The phone allows users to assign different ringtones to individual contacts in the address book. This helps users to distinguish the caller by the ringtone without directly checking the phone number.

### 7.3.2 Prototype

To achieve our research goal we developed a prototype, called *EmoDetector*. The application is a standalone Python-based application and works on Symbian S60 mobile phones. It is capable of detecting certain sets of characters from incoming messages and playing a corresponding tone in case of finding a positive match.

After the installation and launch the application, it runs as a background process without having any impact on the other phone's functionalities. The application has a callback feature, which is called and activated whenever a short message arrives. The callback feature analyzes the content of incoming messages. It searches for certain sets of characters. In case of a positive match, it plays a corresponding tone after the normal SMS tone. The tones are played based on the current profile settings of the phone. Therefore the audible notifications are not played if the phone is on silent mode.

Based on the result of the survey, following character sets are considered for matching: six emoticons, the question mark, and the keyword "OK".

- **Emoticon:** “:-) or :)” and “;-) or ;)” (happy message), “:-( or :(” (sad message)
- **Question mark:** “?”
- **Keyword:** OK (not case-sensitive)

It was a deliberate design decision to limit the number of different preview sounds to a small set. The main reason was to avoid the complexity of learning them and minimize effort for the user.

We recruited a professional composer to compose a tone for each character sets. Each tone composed in such a way that it represented the type, characteristics, and emotions included in each character set. The tones are maximum three seconds long. If a message includes more than one character set, the application detects just the first character and plays the related tone.

The application additionally creates a log file and includes a GUI, which shows the characters detected. We did not implement a comprehensive content logging function

since this would have had a major impact on the users' privacy. Furthermore, we did not replace, but rather appended tones with the original SMS notification tone.

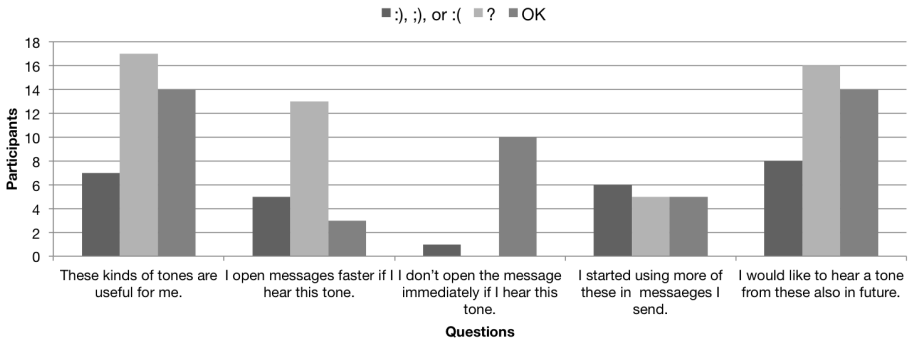
### 7.3.3 Procedure

We invited the participants to the lab and after a short introduction we asked them to install the application on their mobile phone. The participants could either use their own mobile phone if it was compatible with the application or we provided a Nokia 6210 Navigator for the duration of the study. During the four weeks study period, the participants used their own SIM cards and received 20 €.

The procedure of the study consisted of following parts:

1. In the preliminary interviews, we gathered demographics, asked about the participants' current SMS behavior, gave a short briefing about the study, and explained how the application worked.
2. Approximately after one week of usage, the participants were asked to fill in an online web survey, which unveiled the initial impression about the audio preview and if it already had changed the messaging behavior. In addition, the users were asked to complete a System Usability Scale (SUS) questionnaire [23].
3. One week before the final interview the second web survey was conducted. In the survey we also asked participants to provide suggestions on how to enhance the audio preview application.
4. In the final interview, we repeated the questions from the first web survey, to compare the initial impact with the long term use. Additionally, open-ended questions asked for cases where, when, and how the audio preview had changed their SMS behavior.
5. We performed a recognition test. We played the tones to the users and asked them which character set they thought the sounds corresponded in order to evaluate the learnability.

Since the application was running in the background users were not required to interact with the application during the trial. We provided a hotline number and asked the participants to contact us in case of any problem with the application. Meanwhile, we regularly contacted the participants to ask if they had any problems.



**Figure 7.2:** The result of interview indicates how the audio preview have changed the users’ behavior of using SMS.

### 7.3.4 Results

The results of the study indicate that the abstracted audio preview has an impact on how the participants utilize SMS. Already after a week of usage, eleven participants stated that they opened an incoming SMS faster if they heard the question mark tone. Second, eight out of twenty mentioned that they did not need to open a message immediately if they heard the tone for the “ok” keyword. In comparison with the results from the final interview, the results were not statistically significant difference in use after one week.

Figure 7.2 shows the results from the interview. During the interview some of the users mentioned that they usually did not check the incoming messages immediately if their phones were not nearby – unless they heard a tone indicating a question mark. In this case they wanted to check the message immediately. If the participants had an ongoing SMS conversation and they heard a tone indicating *ok* they did not necessarily open the message, but could anticipate the response received. Further, the qualitative feedback revealed that couples, interestingly, tended to use more emoticons and *ok* instead of yes or similar agreement words in their SMS conversations after they started using the application.

Table 7.3 includes an overview of the results from the recognition test. The results show that that the tones for “ok”, “:)", and “?” were the most easily recognized best among the participants. We realized that the degree to which users could recognize the tones correlated with the number of emoticons, keywords, or punctuations received during the study. The correlation coefficient between the number of tones recognized and received



**Table 7.3:** The percentage of users that could recognize the tones.

	ok	:-) or :)	?	:-( or :(	;-) or ;)
<b>Recognition</b>	19%	17%	15%	8%	4%
<b>Occurrences</b>	11.3%	12.2%	22.8%	1.6%	1.7%

is 0.71 ( $r = 0.71$ ). This result indicates that the abstracted audio can be used to preview content, but the learning of the tones depends on how often they are heard.

Further, the SUS test score from the initial survey was 77.12 and from the final interview 83.12. The result from the SUS test indicates that the users are more comfortable with the application after a longer period of use.

### *Limitations*

In the study we did not include a control group to collect comparative data. We assume that SMS behavior does not change significantly in the short-term with experienced mobile phone users. Thus, we relied on the data collected in the preliminary interview. The data collection time was limited to four weeks. Although one can argue that this is not long enough to record the long-term influence of new technology, this time frame seems appropriate as we could observe interesting changes in behavior.

The character set used in the study is also limited. However, it is representative for the most frequent emoticons and keywords used in the SMS communication. The result of the online survey confirms the popularity of the set. Nevertheless, investigating other character sets is also interesting and may reveal other findings.

## 7.4 From a Message to a Melody

The study conducted and described in previous section (Section 7.3) reveals that users' behavior varies based on their situations when it comes to check incoming messages. They also check incoming messages instantly or later if they understand the content type of the messages. These findings suggest that users may eventually even be able to not only guess a message's intention, but to also understand the whole content by learning the musical representation of words frequently used. The success of Morse code, which can be used both for visual and audio encoding of messages, already reveals the feasibility of such approaches. The findings encouraged us to further investigate how



**Table 7.4:** Frequent vowels and constants are mapped to different notes. Other constants are randomly mapped

Mapping Note	Vowel	Constant
tonic	E	N, R
third	I, A	T, S
quint	U, O	-

To enhance the quality of our melody, we decided to not just randomly map characters to the pentatonic scale, but to additionally consider the frequency and the position of a character in the text. In the German language, each syllable contains at least one vowel. Thus, we decided to map vowels to the three tones of the tonic keynote's triad, while also taking into account the average occurrence probability (Table 7.4). Hence *e* as the most frequent vowel (17.40%) is mapped to the tonic keynote, *i* (7.55%) and *a* (6.51%) to the third, and *u* (4.35%) and *o* (2.51%) to the quint. Further, we analyzed the frequency for the consonants as an ending character. Based on the results we mapped the most frequent ending characters *n* (21.0%) and *r* (13.0%) to the tonic keynote, *t* (10.3%) and *s* (9.6%) to the third. The other consonants are mapped randomly.

## 7.4.2 Intention of a message

Users send messages for specific purposes, such as to coordinate, to exchange information (positive, negative), or to express feelings and emotions (happy, sad). To take this into account, we analyze the content of each message for the occurrence of punctuation marks, keywords, and emoticons. We use a simple mechanism to detect a message's intention and then accordingly transform it into a melody. To reduce complexity, we focus on the distinction between major and minor scales only. Table 7.5 gives an overview on the mapping of different intentions, given the phrase "HELLO!". Figure 7.4 shows the melodies of mapping in C major and minor. One option would be to not only transform the music into major and minor scales, but also consider the association of certain keys with specific moods (e.g., flat / sharp keys). However, key-mood associations are invalid for modern (digital) equal temperament keyboards [153].

### *Punctuation*

Punctuation is not only used to indicate the end of a sentence, but also to specify the type of the statement. In Section 7.3, we showed how question marks could be used for

**Table 7.5:** Example for mapping the word “HELLO!” in C major and minor tonality.

Character	Note (C major)	Note (a minor)
H	d	b
E	c	a
L	e	c
O	g	e
!	c/e/g	a/c/e

**Table 7.6:** Intention and according mapping of emoticons, punctuations, and keywords to chords (C= C major, a= a minor, C<sup>7</sup> = C major seventh chord)

	Characters	Intentions	Mapping
<b>Emoticons</b>	:-) :) ;-)	positive	C
	:-( :(	negative	a
<b>Punctuation</b>	?	interrogative	C <sup>7</sup>
<b>(selection)</b>	!	declarative	C
<b>Keywords</b>	when/where	question	C <sup>7</sup>
<b>(selection)</b>	yes, ok	positive	C
	no, sorry	negative	a

creating abstracted audio previews in SMS. We consider both aspects in the following way: whenever a sentence ends with a punctuation mark, we insert a predefined triad into the melody. We use the triad’s type to indicate the type of sentence (Table 7.6). Whereas we use a seventh chord to represent a question, regular triads are used for a full stop or an exclamation point. However, triads are adapted to the intention of the message (major or minor).

### *Greeting and Leave-Taking Phrases*

Text messages often begin with a greeting (hi, hey) and end with a leave-taking phrase (cu, regards). Missing phrases indicate that other messages were previously sent back and forth. Since the beginning often already conveys the message’s intention, we also transform this phrase into a chord.

We deliberately used a simple mechanism to detect the intention behind the message. There is a body of work that looks into text mining and text understanding. However,

this is not the focus of this thesis. By having a simple mechanism, we believe it is possible to increase the learnability of the sonification and its relation to messages. More sophisticated approaches for understanding intentions can be easily included in the algorithm.

### 7.4.3 Algorithm

Based on the approach described, we use the algorithm depicted in Figure 7.5 to create a melodic representation from arbitrary message strings.

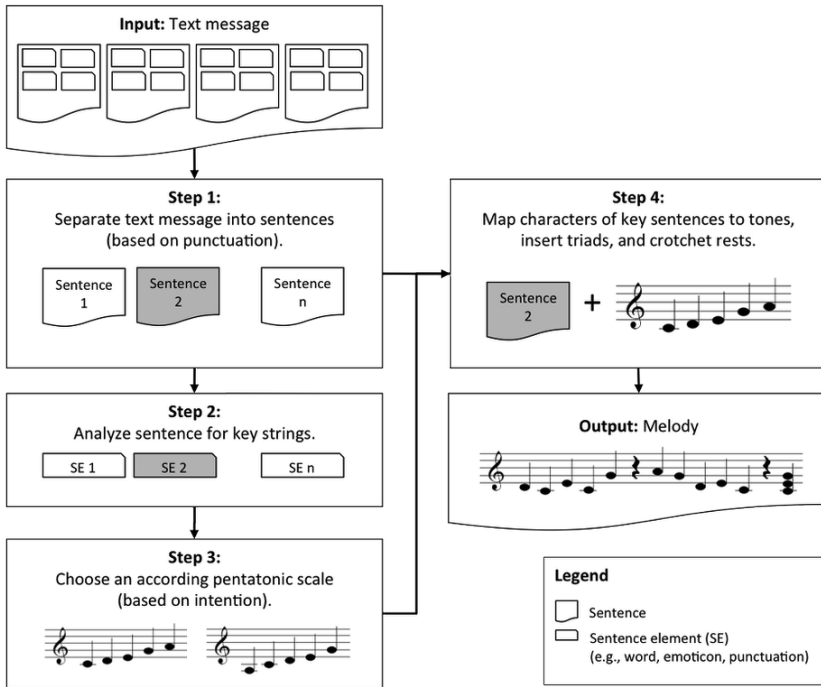
The algorithm steps are as follows:

1. It takes a message string as an input and separates it into sentences by analyzing it for punctuation marks.
2. Each sentence is analyzed for hints (key strings) that reveal its intention. Such hints include emoticons, keywords, and punctuation marks.
3. Based on this analysis we choose a corresponding pentatonic (major or minor) scale.
4. Each single character of each word is mapped to corresponding note.

Besides mapping single characters to keys, we also create triads and tetrads for keyword, emoticons, and punctuation marks. It varies between root position, 1st and 2nd inversion, depending on the intention. Spaces are transformed into crotchet rests. The current version of our script supports text sonification in German language only. However, this approach can be extended to other languages. For a comparable sonification, an analysis of the language as explained above is essential.

## 7.5 Sonification of Message Intention

In an initial online study we tried to prove the feasibility and validity of our approach. The scope of the survey was to reveal whether our approach allowed participants to understand the intentions encoded in the messages. Users may be able to determine a message's intention. This may affect their behavior in such a way that they want to check certain messages immediately, whereas they want to finish their current task first before



**Figure 7.5:** The visualized of the sonification algorithm

checking other messages' content. We expect people with medium/strong musical skills to experience less difficulty understanding our musical representation.

We considered the following hypotheses:

- Users can guess if a message contains positive content, negative content, or a question solely by hearing the melody (H1).
- Users with musical knowledge will learn the melodic description of the intention of the message faster (H2).

We ran an online survey in the summer of 2009 over a period of four weeks. Participants were recruited from music forums, mailing lists, Facebook, and university mailing lists.

### 7.5.1 Apparatus

We used the LimeSurvey tool to realize our online survey. We implemented an AJAX-based web application, which reads a text message, sends it to a PHP-based sonification script, and creates a local MIDI file from the returned XML code. The local MIDI file can then be played back using a media player (e.g., Flash or Quicktime). The web application was integrated into the survey.

### 7.5.2 Procedure

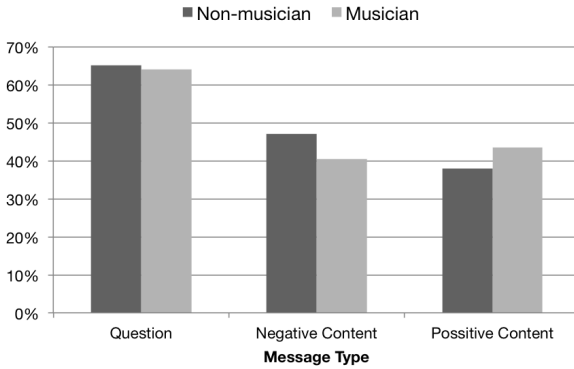
First, we collected demographic information such as gender, age, profession, and musical knowledge. We asked if the participants have played any instrument and had them rate their musical skills on a 5-Point Likert scale (1=Beginner, 5=Professional).

Second, we were interested in the understandability of our mapping and the users' association between intention of the message and the sound of the melody. Therefore, we presented the participants with the sonification of three real text messages encoded with our tool. We used a piano melody for the representation. The melodies were 10-12 seconds in length. The three melodies included the sonification of one message with positive content, one with negative content, and one question (random order). We then asked the users to associate the melody with one of three provided text messages ("no answer" was also an available choice). Hence, we were able to collect the initial impressions participants had of the melody as it related to the intention of the message. Further, we asked if the melody sounded happy, neutral, or sad to them.

Finally, we let people try out the algorithm with their own messages using a web application. We asked them for their personal opinion, privacy concerns, and whether or not they would use the tool.

### 7.5.3 Results

In total, 69 persons completed our online survey (54 male) with an average age of 27.7 years ( $SD=7.32$ ). The participants were mainly students (40) and employees (23). In total 37 participants played a musical instrument. The most popular instruments were piano / keyboard (17), guitar (15), drums (7), and base (6).



**Figure 7.6:** Comparison of correct answers between musicians and non-musicians.

### Interpreting a Message Intention

As depicted in Figure 7.6, the results show that that questions are correctly interpreted by 65.2% of the participants. Messages that contained positive (40.6% correct answers) and negative content (43.5% correct answers) are more difficult to distinguish. However, these results are well above 25% or random choice level.

For those participants who could correctly link the played sounds to messages, we further investigated the degree to which they linked general intention with the melody (e.g., whether they associated a happy sound mapping based on a major scale with positive content and a sad sound mapping based on a minor scale with negative content). The results show that 83.3% considered questions to sound neutral, 84.6% considered the negative messages to sound sad, and 68.2% considered positive messages to sound happy. Hence, this is a strong indication that users who are able to distinguish between negative messages, positive messages, and questions associate the sonification in the way we intended making it very understandable.

### Musicians vs. Non-Musicians

The analysis of responses between musicians and non-musicians did not reveal any significant difference. Figure 7.6 shows that for all melodies both groups produced comparable results. To test the hypothesis that the understandability is not influenced by the experience in playing an instrument, we used a Pearson's  $\chi^2$ -test of independence for each message type. To compare the overall means of correct answers, we used an Analysis of Variances (ANOVA). The ANOVA shows, that the musicians' mean of correct answers (2.54) is lower than the mean of those who did not play an instrument



**Table 7.7:** Comparison of understandability among musicians and non-musicians ( $\chi^2$  Test)

	$\chi^2$	p	% correct answers per group	
			non-musicians	musicians
<b>question</b>	0.004	0.940	0.656	0.649
<b>negative message</b>	0.280	0.597	0.469	0.405
<b>positive message</b>	0.235	0.628	0.375	0.432

(2.81). This difference is not significant ( $F(1,67)=.924$ ,  $df$  1;  $67$   $p=.34$ ). Hence, it is likely that a random effect caused the differences. The significance level (Table 7.7) shows that giving the correct answer never depends on whether or not a person is musician.

### Qualitative feedback

We received numerous interesting hints and suggestions for improvement from the participants. Suggestions included the adaptation of different music genres, different instruments (which could match the receiver's music taste, be mapped to the gender, or be used to distinguish different intentions), adding variations in the tempo of the music, and the inclusion of sequences from popular songs indicating the intention. Several users stated that they would prefer shorter musical representations.

The results, in general, revealed that almost half of the survey's participants could, without any learning, understand the intention of a message. The most understandable message types were the questions, followed by negative and positive messages. The majority of the participants associated a message sonification with the envisioned intention. This is a strong indicator for support of the first hypothesis (H1). We believe that the understandability can be significantly enhanced if people use message sonification over a longer period of time. The results do not reveal any evidence that that musicians perform better than non-musicians. Since both groups perform similarly we reject the second hypothesis (H2).

## 7.6 Sonification of Instant Messages

The online survey gave us a good understanding of the issues and challenges related to message sonification. Yet no evidence has been found that our results are either representative or would be true in the real world. Hence, we implemented *skypeMelody*, a Skype plug-in, which we tested in a real world setting in a study conducted over the course of two weeks.

### 7.6.1 Apparatus

*skypeMelody*, a Java-based Skype Plug-in, is implemented in a similar way as the web application used for the online study. It reads incoming text messages, transforms them into XML-conform MIDI and plays back the melody. For intercepting incoming text messages, users are required to provide a one-time authorization for the connection to Skype. With regard to the users' comments from the online study, we decided to decrease the length of the sonified message. Many participants stated that the message representations were too disruptive when played in full length. However, since we believe users might be able to learn to understand an entire message, we did not simply crop the message in length, but rather filtered out those sentences including no keywords.

We defined in total 24 keywords. The keywords were derived from the online study where we asked the participants to enter each 2 short text messages including a question, a positive content, and a negative content. In total, we analyzed 414 messages.

To gather quantitative data we also included a logging functionality into the Skype plug-in, which allowed us to store certain types of information in an external database. For each message, we stored two time-stamped records each including the hashed user ID and a message ID to later associate both records with each other. We used the Received Event Record for storing the current status of the user (online, away), the message type (type 1: positive, type 2: negative, type 3: question), the number of keywords, the message length, and a list of all enrolled key words. The Read Event Record consisted of the reading time (allowing to calculate a receive-to-read time) and the message id (required for associating both records).

### 7.6.2 Procedure

In the field study, we focused mainly on changes in the user behavior and the understandability as well as learnability of our representation. Users were asked to install

*skypeMelody* and continue their regular Skype behavior. In other words, no specific tasks were given during the study. We used an initial questionnaire to evaluate demographics, text-messaging behavior (number of conversations per week, average length of conversations, communication partners and situations), and musical experience.

To measure how the use of *skypeMelody* influenced their behavior we asked the users to fill in questionnaires after each week. In these questionnaires we were mainly interested in how easily users could distinguish different types of messages and if they checked incoming messages sooner or later than before. Additionally, participants were asked to fill in the system usability scale (SUS) questionnaire [23].

### 7.6.3 Participants

We recruited 14 participants for the study from our lectures, forums, and Facebook. Participants were mainly students with only 2 employed participants, making the sample rather homogeneous, but nonetheless representing a main target group for such an application.

Twelve participants had on average more than 10 text-based Skype conversations per week. Their most important conversation partners were friends (72%), colleagues (61%), partners (40%), and family members (40%). The most important situations in which they used Skype were after work (80%), on weekends (75%), and also during work (60%). The main purpose of the conversations included side conversations (80%), discussion of complex problems (70%), and dating (50%). Our participants used Skype to a large extent for short conversations, e.g., to schedule the time to go out for lunch together (61%).

### 7.6.4 Results

The analysis of log files reveals that each user received on average 20.47 messages per day. Out of the 1085 messages, 658 contained no keywords, 124 were negative, 165 were positive, and 138 were questions. The most common keywords were “yes” (153), “?” (113), “not” (89), “where” (50), “:)” or “:-)” (35) and “;)” or “;-)” (31). The average number of keywords per message sonified was 1.54 (overall mean=0.61).

In total we collected 2533 data records. For consistency reasons, we excluded 66 records where only the read event was registered, and 119 where more than two events occurred per each message (Skype’s message IDs are not unique). We also performed a semantic check of the data. We excluded one message containing all 24 keywords, and

**Table 7.8:** Comparison of receive-to-read time based on message intention. The ANOVA analysis shows that the time to read an incoming message varies based on the intention of the message. This difference is significant ( $F(3,1081)=4.979$ ,  $p < .01$ )

Message Type	question	positive	negative	none	Time (second)
Mean	8.03	5.87	6.95	11.50	
Standard Deviation	20.15	15.19	17.29	21.22	
Number of Messages	138	165	124	658	

23 messages where the recipients' user status was set to "AWAY" as we could not be sure that they received the sonified message. We finally removed 64 outlier messages where the receive-to-read-time delay was beyond a threshold of 120 seconds (assuming read-events after more than 120 seconds were not caused by the sonification). This resulted in 2170 usable records (representing 1085 messages) that could be used for the analysis.

In the following part, we analyzed the results of the study in order to obtain qualitative and quantitative data on (1) changes in the user behavior based on the sonification and (2) the understandability and learnability of our sonification algorithm.

### Messaging Behavior

To assess the effect on the users' message checking behavior, we compared the receive-to-read time for different aspects. First, we compared the receive-to-read time based on the message intentions. The ANOVA analysis shows that the times differed significantly ( $F(3,1081)=4.979$ ,  $p < .01$ ). The results in Table 7.8 show that the users checked non-sonified messages the most slowly. For the sonified messages, questions required the most time until they were checked. Positive messages were checked faster than negative messages.

Further, we analyzed differences between week 1 and 2 of the study. We found out that for all message types the mean time increased. We believe that this is the result of a curiosity effect (people got more used to the sonification in the week 2). However, the increase in time is only significant for positive messages ( $p < .05$ ). Table 7.9 shows the comparison of the message types in both weeks.

**Table 7.9:** Comparison of receive-to-read time between the 1st and 2nd week.

Message Type	Time (second)		F-Value	df	P
	Week 1	Week 2			
<b>question</b>	5.868	9.376	0.99	1, 136	.322
<b>positive</b>	2.391	8.365	6.415	1, 163	.012
<b>negative</b>	7.442	6.597	.072	1, 122	.789
<b>none</b>	10.15	12.44	1.857	1, 656	.173

Subjective user feedback from the questionnaire reveals that surprisingly the perceived receive-to-read time decreased between week 1 and 2. We assume that this effect of “false perception” is a result of the users’ adaption to the system.

### Understandability / Learnability

We used a two-step algorithm for clustering the messages according to the combination of occurred keywords in order to verify, whether our separation of message types was distinct. Further, we use the algorithm for its strong capability to work with discrete values. As it can be seen in the Table 7.10, the results are 4 clusters (columns), which exactly fit the self-chosen classification algorithm (rows). In total, 91.2% of all messages clustered like our algorithm would have predicted. According to [97] the cluster solution can be considered to be “good” (separation accuracy= 0.7). The visualization of the results depicts that only cluster 2 (positive messages) lacks precision beyond 90%. This is explained by the fact that messages containing positive keywords are very likely to also contain other keywords.

The results from the questionnaire show that the understandability of the different message types increased between week 1 and 2 (results based on 5-Point Likert scale, 1=not understandable at all, 5=very understandable). The participants could recognize questions best (mean=3.7, increase=14.6%), followed by positive messages (mean=3.2, increase=28.9%) and negative messages (mean=2.8, increase=20.8%). Although, the results are not significant.

### System Usability Scale (SUS) Test

In the questionnaires in both weeks we asked the participants to fill in the SUS test

Table 7.10: Clustering of Message Types

Message Type	Cluster								Sum
	Cluster-1		Cluster-2		Cluster-3		Cluster-4		
<b>question</b>	83	93.3%	55	22.0%	0	0.0%	0	0.0%	138
<b>positive</b>	4	4.5%	161	64.4%	0	0.0%	0	0.0%	156
<b>negative</b>	2	2.2%	34	13.6%	88	100%	0	0.0%	124
<b>none</b>	0	0.0%	0	0.0%	0	0.0%	658	100%	658
<b>sum</b>	89	100%	250	100%	88	100%	658	100%	1085
<b>#, % corr.</b>	83	93.3%	161	64.4%	88	100%	658	100%	990

[23]. The score for the first week was 72.72 (SD=12.5) and for the second week 82.65 (SD=7.1). Considering the small sample, it is rather unlikely that differences are the result of a random effect ( $p=.108$ ). We evaluated the reliability of the scale using the Cronbach's Alpha measure, representing the inter-item correlation of all answers for each questionnaire (high values represent a consistent opinion of the participants over all questions). The alpha values were 0.544 for week 1 and 0.539 for week 2 (a minimum value of 0.5 is required to justify the combination of all answers to an index). While the results give some good indications, they are not yet empirically significant.

### Privacy

From the online survey, we found that participants considered it to be essential that the sonification matches the intention of the user. Even though 71.4% of the participants liked the overall idea only 19.7% of the participants would use it if the intention was not obvious. Semi-structured interviews conducted after the field study reveals that prevalent, user-centered requirements would be to have aesthetically pleasing sounds and the ability to customize.

From a privacy point of view, we realized that only 19.0% of the participants were afraid of a very strong or strong influence on their privacy, even considering that the content of a message might be understandable to other persons familiar with the encoding (5-Point Likert scale, 1=very strong influence, 5=no influence at all). In contrast, 75.2% of the users felt that reading the message out loud would strongly influence their privacy. Thus, it can be concluded that the sonification of messages still preserves the user's privacy.

## 7.7 Implication

The assessment of SMS usage shows the widespread use of emoticons and punctuation in SMS messages are accepted as a means for expressiveness. Furthermore, it reveals that users check incoming messages depending on their current situations. There are situations where users prefer not or are not able to immediately check incoming messages. The abstract audio previewing of SMS has a significant impact on the users behavior when it comes to checking their messages in situations where users are engaged in other activities. The audio preview of question marks often leads users to check messages immediately. This can be due to the fact that the audio preview informs users that the incoming message includes a question and users want to answer it. In contrast, messages including *ok* are mainly checked after finishing current activities. As reason can be that playing the audio preview of *ok* already reveals the response to an ongoing conversation. Hence, users are aware of incoming message to some degree and do not necessarily check the message immediately.

Additionally, during SMS conversations in which both sides have the application, users use emoticons and popular keywords that are audio previewed more often than any similar keywords words. This is due to the fact that both sides are aware that the other person also uses the application and receives the tone. Therefore, a sender can increase the awareness of the receiver by intentionally using more emoticons and keywords.

The results of the online survey suggest that it is essential that the sonification reflects the intention of the user. However, the privacy should be preserved. Further, there is no significant difference between musicians and non-musicians when it comes to understanding the intention of a message sonified. Nevertheless, it is crucial to use intuitive and easy-to-distinguish musical elements (chords, keys). The field study conducted in the desktop context also reveals that the sonification has a significant influence on users' message checking behavior. There is a significant difference in the receive-to-read time not only between sonified and non-sonified messages but also among the different message types. Users check positive and negative messages quickest. Interestingly, the time delay for checking messages is maximal for messages that include a question. We assume that users tend to finish their task at hand before checking a message containing a question since such messages require a certain level of attention and an effort to reply. On the contrary, messages containing positive or negative news seem to be more interesting to users so that they want to check immediately. Surprisingly, in the mobile context, it is the opposite and users check the questions faster. This indicates that for users the importance and the priority of messages are context based.

Further, the cluster analysis shows that the sonification algorithm proposed produces good results regarding the mapping of the intention of a message to a musical representa-

tion. Even though our results indicate a good understandability, we cannot yet provide empirical evidence on how easy users could learn to understand the precise wording of a text message. The results from the questionnaire are not significant. However, there is a strong indication that users are more comfortable with the sonification when they use the system for a longer period of time.

## 7.8 Summary

In this chapter, we investigated the sonification of text messages as a means to convey information and create awareness about the messages nonverbally. First, we assessed the impact of providing abstracted audio preview of SMS messages on the user's behavior. An online survey with 347 participants revealed the widespread use of emoticons and punctuations in SMS messages. Further, checking incoming messages varies based on users situations. These findings motivated us to assess how the user's behavior is changed if audio previews based on messages' content as nonverbal information are provided. A mobile application was developed which monitored incoming messages and searched for a specific set of characters. In case of a positive match, respective audio preview was played right after the SMS notification on the mobile phone. The audio preview was generated based on most frequent emoticons and keywords used. Indeed, a field study revealed that the audio preview influences user's behavior. Users checked messages that include questions faster. Messages including *ok* were checked after finishing current activities and not necessarily immediately. Furthermore, users more often intentionally used the character sets that are audio previewed if both the sender and the receiver had the mobile application installed and used it. Users particularly exploited this approach for increasing the awareness.

The results of the study in the mobile context encouraged us to explore the same approach for instant messaging in the desktop context. However, we extended the approach in order to convey more information about the content of a message and its intention rather than providing a simple abstract preview. We developed an algorithm that sonifies text messages and transforms them into euphonic melodies. We evaluated the algorithm through an online survey with 69 participants. To assess user's behavior, same as the other case study, we conducted a two-week field study among 14 participants. We developed a plug-in for the *Skype* application. The plug-in read instant messages and sonified them based on their content. The results showed that the sonification had not only a significant influence on user's behavior, but also on different message types. While positive and negative messages were checked fastest, messages including questions needed more time to get checked. This is, surprisingly, opposite on the mobile context. We believe that in



the desktop context such messages require a certain level of attention and more effort to reply. Therefore, users prefer to finish their current ongoing task before checking and replying.

In general, the results depict that tones as nonverbal information can be used for conveying information and creating awareness. Users are able to learn the tones generated based on the contents of a message. They understand the message content based on the provided tones. However, the learnability has a direct relation with the number of times tones played and heard. Further, users' checking/reading behavior is influenced if they understand the content and of incoming messages. The tone can be used to communicate the intention and the content of text messages. Users implicitly exploit the sonification as a means for awareness or even fun. Providing such nonverbal information can, indeed, impact behavior and awareness of users. However, the impacts can be varied between the desktop and mobile context.



# Chapter 8

---

## Sharing Sentiments with Iconic Interfaces

The ubiquity of mobile phones and the Internet connectivity on them provides this possibility to share information among non-colocated users in real-time and connect them together. Considering the television as one of main sources for entertainment, the majority of viewers usually watch TV alone [46]. However, watching TV does not necessarily have to be a solitary experience. It can foster multiple forms of socializing [140]. Researchers have investigated to connect non-colocated TV viewers via telecommunication technologies [43, 116], mainly referred to as Social TV. Typical social TV systems include presence of viewers, voice, video, text, or combinations of these information. The main goal of such systems is to enable viewers to actively share information about TV content in real-time to increase the social experience and connect them together.

Smartphones are indeed used as a second screen for social networking, chatting, and web browsing while watching television. They can serve as standalone platforms for collecting and sharing the user's emotional responses to TV-related experiences [53]. The main research challenge lies in establishing a shared TV watching experience that is meaningful and engaging to users and at the same time does not distract viewers from the actual content. The goal of the work presented in this chapter is to investigate how mobile phones can be used as a communication channel to connect TV viewers together.

The user interface of mobile phones can be used to exchange information that represents emotional reactions to events shown live on TV.

We chose the soccer World Cup 2010 tournament as a shared TV watching experience for this research. The event is most famous soccer tournament worldwide. It has extremely high public attention in many parts of the world and many people have a high emotional involvement to (at least some of) the matches. The matches are broadcast live and synchronized in time with many simultaneous viewers. We focused on exchanging spontaneous emotional feedback between users who are part of a virtual fan block.

We developed a mobile application called *World Cupinion* for expressing reactions. *World Cupinion* is an Android application that allows soccer fans to express their opinions about events and moments in live soccer matches. Through this application users can support their favorite teams and share their opinions with other fans in real-time. As we expect that users' focus of attention is mainly on the match itself and short quick interactions occur when interesting events happen, the design focus is on simplicity and quick usage of the application. Opinions should be expressed with a minimum effort and, in the best case, nonverbally. When the application is not actively used, it mostly serves as an ambient display that conveys the opinions aggregated from the active users. To address these aspects, we designed an iconic user interface. The interface allows users to quickly, and with minimum effort, to share their sentiments. Opinions can be express nonverbally through interactions with icons.

In this chapter, the following research questions are investigated and addressed:

- How mobile phones can be used as a communication channel for sharing sentiments in real-time and connect non-colocated TV viewers together?
- How can TV viewers nonverbally share emotional reactions in real-time using iconic user interfaces?
- How does this communication channel impact the experience of watching TV among non-colocated viewers?

*This chapter is based on the following publications:*

- A. Sahami Shirazi, M. Rohs, R. Schleicher, S. Kratz, A. Müller, and A. Schmidt. Real-time nonverbal opinion sharing through mobile phones during sports events. In *Proceedings of the 2011 International Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada), CHI '11, pages 307–310, New York, NY, USA, 2011. ACM
- R. Schleicher, A. Sahami Shirazi, M. Rohs, S. Kratz, and A. Schmidt. World-cup opinion experiences with an android app for real-time opinion sharing during soccer world cup games. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 3(4):18–35, 2011

## 8.1 Related Work

Various researches have explored the idea of using additional communication channels in parallel with watching TV. *AmigoTV* is an early social TV system that used voice chat communication in combination with broadcast TV [43]. It also provides emoticons and a buddy list with online status. Motorola Labs developed a series of prototypes called “Social TV” system (STV), which allows users to engage in spontaneous communication with their buddies through text or voice chat while watching TV [80, 120]. The system also includes an additional display to convey views of the current TV-watching users. Harboe et al. give a comprehensive overview of social TV systems [80].

Further, various user studies have investigated the communication modalities. Geerts [65] as well as Baillie et al. [12] compared communication via voice with other modalities. Both studies report that most users believe that voice chat is more natural and easier to use than text chat. However, Huang et al. [93] conducted a similar study using the STV system. They concluded that participants tend to prefer text chat and they often communicate about topics unrelated to the TV content. Geerts and DeGrooff reported a set of comprehensive sociability heuristics for social TV systems [66]. It should be mentioned that the social TV systems discussed require the installation of set-top boxes for supporting collaboration. Since set-top boxes are only available in certain locations, users are restricted to particular environments for using such systems.

Another strand of research has studied the content shared through such communication channels. Diakopoulos and Shamma analyzed the sentiments of tweet annotations for a presidential debate to find out their relationship to topics discussed and performance

of the opponents in the event [53]. They report metrics can be used to detect highlights during social media events. Miyamori et al. [125] proposed and examined a method for generating views of TV programs based on viewer's opinions collected from live chats on the Web. Nakamura et al. [133] evaluated affective responses to unstructured video commenting systems. Taking a closer look at what information the audience/watchers of sport events actually wish to share with their friends or fan group, it turns out to be mostly the preliminary evaluation mixed with the personal emotional impact of specific events during the game. The evaluation is rather on the "cold" rational assessment of ongoing maneuvers on the field.

The sudden and strong expressiveness of emotions during sport events allows event viewers to somehow extend their reaction beyond the usual radius of face-to-face communication. However, to express and share of such information through digital communication channels impose a number of requirements. Feedback should be quick and if possible "analogous", i.e., nonverbal to avoid the necessity of lengthy formulation to describe a simple and transient affective rush. Emoticons appear to be an appropriate way to communicate these states [51]. In addition, the provided rating scheme should contain domain-specific labels (e.g., "yellow card" for soccer matches) as well as domain-independent features (e.g., "like-dislike") [146]. Relying on such a limited set of means for expression is also referred to as *lightweight communication* [120].

Prior work has utilized mobile phones to support sports fans. *MySplitTime* allows users to take pictures of bypassing cars at rallies and obtain additional information about the current ranking of the car photographed [57]. *TrottingPal* helps spectators at the trotting track to gather additional information to improve their betting and to coordinate with other visitors who might be dispersed across the area [138]. *World Cup Predictor* is a mobile application that allows users to predict the results of World Cup football matches and awards points for correct guesses [128].

All the mentioned social TV systems require the installation of set-top boxes for supporting collaboration. Since set-top boxes are only available in certain locations, users are restricted to particular environments. To overcome this limitation, we intended a mobile phone application that would give users the opportunity to use it for sharing their opinions in any context in which watching the event is possible, even in bars, the stadium, or at public places – a requirement indispensable to the sports domain.

## 8.2 World Cupinion: a mobile app for sharing opinions

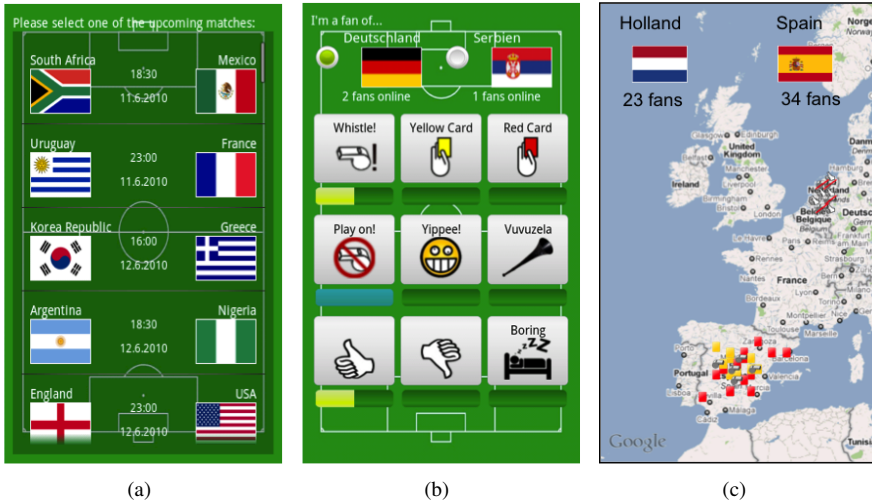
To address the research questions, we developed a mobile application called *Word Cupinion* for Android mobile phones. This application allows users to share their opinions about events that happen during a soccer match in real-time. With this application, we assess whether the mobile phone can be used as a nonverbal communication channel for sharing sentiments and connect users together.

### 8.2.1 Design Rationale

Since the soccer World Cup was chosen and targeted for the share event among users, the *Word Cupinion* app was mainly designed for soccer fans (viewers) to share their opinions while watching a match in real-time. The design of the system took the following rationale into consideration.

- **Simplicity.** Since the user's focus of attention is mainly on the match itself, simplicity of the user interface is crucial. The interface should convey enough information, which the users provided, about the current opinions.
- **Short-term usage.** During a soccer match, situations arise quickly where an interaction might just involve stating one's opinion about the current event; thus, supporting a quick and short burst interaction is necessary.
- **Visualizing.** Providing feedback and visualizing aggregated opinions of the competing teams' fans is essential. The application should convey how the collected opinions evolve even if the user is a "lurker" and not actively interacting with the system.
- **Large number of worldwide users.** The system should be able to handle a large number of worldwide users simultaneously. Hence, localizing and supporting multiple languages is important.

Based on that rationale, an iconic user interface is designed to allow users to share their opinions nonverbally. The iconic user interface contains a set of sentiments, related to events that happen in a soccer match, presented nonverbally using proper icons. Further, the interface lowers the cost of interaction and allows users to quickly, and in short bursts, interact with the application and share their opinion. It further decreases the space



**Figure 8.1: World Cupinion** user interfaces: (a) Initial screen “Match List”, (b) Second screen “Arena”, (c) Map View screen shows the geographic distribution of fan opinions.

between “Readers” and “Leaders” [154]. The feedback visualization conveys the current opinion of users about the match.

## 8.2.2 Iconic User Interface

The application consists of three screens (Figure 8.1). The first screen shows the list of upcoming matches with their starting times and dates in the user’s local time zone (Figure 8.1(a)). The time zone plays an important role here, since we planned to distribute and recruit users worldwide. It should be mentioned that the game selection could have been automatized, except for parallel games during the first phase of the tournament. However, we decided to keep the list to allow users to plan their viewing times in advance of the games. The list can be used as the calendar of the tournament as well.

After selecting a game the user would enter the “Arena” for that game (Figure 8.1(b)). This screen is the main screen of the application. It allows the user to express their opinions during the match and to see the opinions of the fans of their own team and those of the other team aggregated. The screen includes the teams’ name and their flag as well



as a matrix of 3x3 iconic buttons. The buttons are used to share opinions. Two types of icons are used in the interface. The first type represents factual events in a soccer match, for example, whether the referee should give a (yellow/red) card or rather playing an advantage and let the game goes on. The second one is opinions/expressive assessments, i.e., thumb up/down, the vuvuzela sound to express excitement, or applause (“Yippee” button). Below each button there is a horizontal bar that indicates the average opinions of both teams’ fans aggregated. The statistics of the own fans are shown in green and the statistics of the other team in blue. The statistic is calculated based on the last 30 seconds input from users.

The third screen depicts the geographical distribution of both teams fans’ opinions (Figure 8.1(c)). This visualization shows geographical clusters of users having opposing opinions. The map view is based on the standard Google Maps APIs with icon overlays for the feedback that is given at a particular location. Using Google Maps APIs allows interactive panning and zooming of the map. However, we restricted the maximum zoom level due to privacy reasons.

It should be mentioned that when a user enters to the “Arena” screen” the rating buttons are disabled until the user decides which team they wants to support. After choosing a team the buttons are activated and the user can start sharing her opinions. This design decision means that users have to be a fan of a particular team in order to share reactions. Hence, we can collect feedback from fans of each team.

### 8.2.3 System Architecture

A client-server architecture is implemented for the application. The mobile application sends two types of requests to the server: input requests and update requests. The input requests are used to send user opinions to the server. This request is sent as soon as an opinion button is pressed. The server logs all inputs in a SQLite database and maintains statistics of the user opinions received in the last 30 seconds.

The update requests are used to poll the state of the mobile application’s user interface. In response to update requests the 30-second statistics are sent to the mobile clients. After a successful update request, the statistics of the “Arena” screen is updated. Further, the map view sends another request type, to which the server generates a response containing the user inputs for the last 5 minutes.

We initially used UDP datagrams for communication, as our protocol does not require an active connection. The UDP datagram also imposes a lower load on the server, which is beneficial if there are many simultaneous server requests. However, it soon appears that

certain network firewalls and also mobile network providers may block UDP packets that have non-standard destination ports. To remedy this, our mobile application has a fallback mechanism that automatically switches to HTTP requests if UDP communication is unsuccessful. User input events are always sent via HTTP to ensure that they do not get lost. Supporting HTTP requests gave us this opportunity to also implement the same application web-based. This allowed users, without Android phones, to use the app without installing any additional app. Users only need a web browser to use the app.

A further important issue of mobile phone application is energy consumption [142, 123]. Over the 90 minutes of a game (plus the 15 minutes break and an optional 30 minutes extension), the application continuously communicates with the server via the mobile phone network or WiFi. There is a trade off between the update rate of the interface and energy consumption. In pilot tests we found that one update every three seconds is sufficient. A significant contribution to energy consumption comes from continuously using the device as an ambient display for the opinion state. Even if the user is not interacting with the device the community opinion is updated and shown.

## 8.3 User Study in-the-Wild

To evaluate the *World Cupinion* app on a larger scale, we conducted a 4-week user study in the wild and recruited a large number of users. It was crucial to conduct the study this way to be able to observe natural user behavior, which is problematic in tightly controlled experiments. We used the mobile and web-based applications for the user study.

### 8.3.1 Procedure

We distributed the app from June 4th, 2010 (one week before the start of the World Cup) for 4 weeks in Google Play, the official Google marketplace for Android phones. The application was available for free. We used several channels to advertise our app. In addition to mailing lists and social networks, we added press releases by the Deutsche Telekom Laboratories and the Technical University of Berlin. The ability to rapidly push new releases of the application to the Android market allowed us to publish weekly updates containing bug fixes or new features during the actual soccer World Cup. After the last match in the tournament an update containing a questionnaire about the application was released. It consisted of 22 questions ranging from simple demographics (age, sex) to open suggestions for improvements. For all evaluative questions, a 5-point Likert scale was offered.

### 8.3.2 Results

Based on the Google play portal at the end of the World Cup, we had 1645 downloads and 448 “active” installations (=29% of all downloads). The number of active installations denotes the number of users that still had the app installed on their devices at that point. Based on our database, 71% of inputs were from the Android client and 29% from the Web-based client. In total 21205 inputs from 925 unique users (71% Android users) were collected during the 64 matches of the tournament (an average of 331 inputs/match). The results presented in the following are based on data from two sources: the logs of user activities during the matches and from the in-application questionnaire provided with the last update.

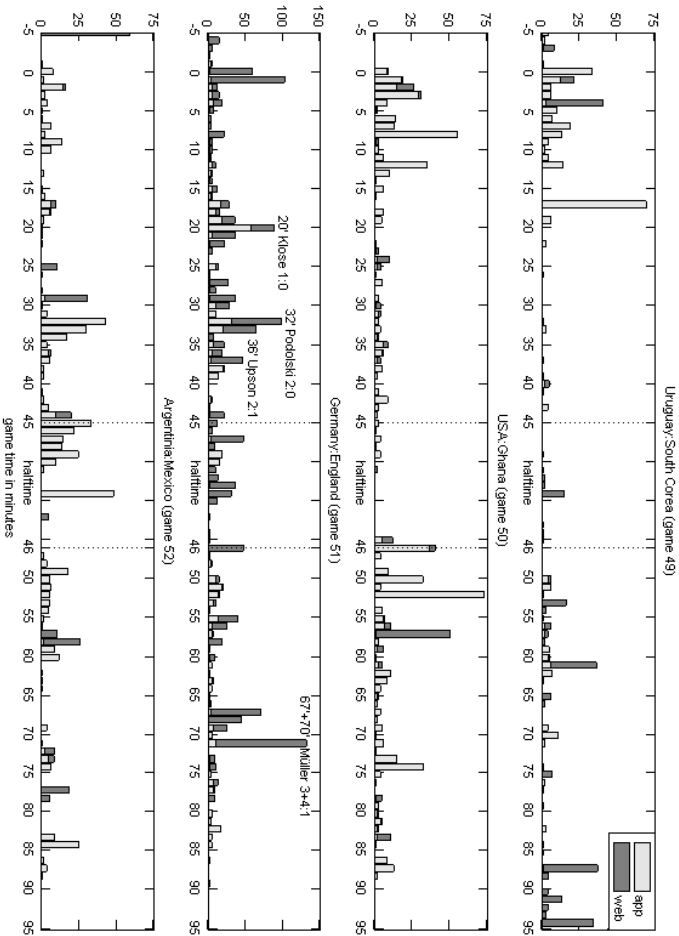
#### *Usage Statistics*

On average 28.6 users (SD=19.1) were active during the matches, with a maximum of 94 and a minimum of 8 users for a single game. The number of active users per match was highly dependent on the nationality of the teams playing. A general decrease after the first couple of matches was observed. From the match 49 the round of sixteen started, which led to a temporary increase in usage. The most prominent match of the World Cup, the final (game 64) had surprisingly low number participants.

The average participation lasted 681 seconds (SD=1316.2) during which 17.6 actions (SD=33.6), i.e., button presses were performed. Table 8.1 shows the average number of inputs during a game, and average session length divided by interface type, i.e., the mobile phone client or web interface.

A multivariate analysis of variance (MANOVA) reveals that there is a significant effect of input interface on these parameters. Users tend to give less ratings during longer sessions ( $F(2,820)=6.584$ ,  $p=.001$ ). However, the ANOVA comparison indicates that this difference is only significant for the number of inputs ( $F(1,821)=9.063$ ,  $p=.003$ ), but not for session length ( $F(1,821)=.714$ ,  $p=.398$ ). While the means in number of inputs clearly vary for both interface types, the medians are identical. This indicates that in the web interface there were some users with very high number of inputs as also shown by the higher maximum (591 total inputs for the web interface vs. 372 for the app). This might be due to the excessive clicking of the mouse.

To provide a better impression of input quality, we plotted the inputs for the first 4 games of the round 16 (matches 49-52) in all of which more than 50 users participated. Figure 8.2 shows the click distribution for these four games. There is no clear difference in input patterns between the mobile phone and the web client. Furthermore, the graphs indicate that ratings are not random and accidental, but rather are related to events in



**Figure 8.2:** Number of inputs during the first four matches of the round of 16, distinguished by interface type, i.e., the Android mobile phone app or the web interface. Bars represent absolute numbers. Please notice the for the match (England vs. Germany), the moment of the goals are also labeled [174].

**Table 8.1:** Usage statistics by interface type, i.e., Android mobile phone app vs. web interface.

	Android app		Web Interface	
	no. of inputs	session length (second)	no. of inputs	session length (second)
<b>Mean</b>	15.58	703.79	23.61	615.13
<b>Median</b>	9.00	60.87	9.00	44.86
<b>Std. Deviation</b>	24.56	1320.22	51.11	1305.39
<b>Min</b>	1.00	.38	2.00	1.44
<b>Max</b>	372.00	6997.30	591.00	6611.56

the game. This is exemplified by match 51 (Germany vs. England). The analysis shows that there is a relation between the collected sentiments and events in the chosen ground truth, i.e., Y! Sports ticker <sup>15</sup>. It is clearly visible that the inputs correspond to important moments of a match such as scored goals and goal kicks. Hence, it can be concluded that users understood the app as intended to communicate moments of high relevance to the other participants. Further, generating a summary of important moments of a match based on the collected sentiments appears feasible.

### *Icon Usage*

In addition to the amount of activity over time, we were also interested in the usage of the iconic buttons. Table 8.2 shows the relative frequency of button clicks per game for the match 49 till the match 52 and across all matches. For the matches 49-52 the differences in relative usage frequency were assessed statistically using the  $\chi^2$  test. The results revealed a clear effect of button meaning on usage frequency ( $\chi^2(24)=407$ ,  $p<.001$ ). Table 8.2 also indicates to what extent the usage frequency of single buttons varies from game to game and whether this difference is statistically significant.

While there were slight differences between single games, the vuvuzela was by far the most frequently used button across all games with the highest number in the match 51 (Germany vs. England). The second most frequently used button serves to annotate a typical soccer controversy, namely that the referee should whistle and stop the match or

<sup>15</sup> Y! Sport ticker: <http://g.sports.yahoo.com/soccer/world-cup>, last accessed August 27, 2014

**Table 8.2:** Relative frequency of button clicks per game in percent (column-wise), for the first 4 games of the round of 16 and across all games. Values are rounded to whole numbers. Percentages with the same subscript letter are not statistically different from each other in the corresponding  $\chi^2$  test when comparing row-wise across games, i.e., 27<sup>a</sup>% Vuvuzela clicks in game 49 is not statistically different from 29<sup>a</sup>% in game 50, but both are significantly lower than 36<sup>b</sup>% in game 51. 13<sup>a,b</sup>% indicates that 13% is neither different from other numbers in the same row with the subscript ‘a’ nor from ones with the subscript ‘b’.

Icon	% usage					
	match number				average	
	49	50	51	52	matches 49-52	all matches
<b>Vuvuzela</b>	27 <sup>a</sup>	29 <sup>a</sup>	36 <sup>b</sup>	20 <sup>c</sup>	28	23
<b>Whistle!</b>	18 <sup>a</sup>	19 <sup>a</sup>	6 <sup>b</sup>	17 <sup>a</sup>	15	19
<b>Play on!</b>	7 <sup>a</sup>	5 <sup>a</sup>	3 <sup>b</sup>	6 <sup>a</sup>	5	6
<b>Thumbs up</b>	9 <sup>a</sup>	9 <sup>a</sup>	15 <sup>b</sup>	13 <sup>a,b</sup>	12	15
<b>Thumbs down</b>	6 <sup>a</sup>	8 <sup>a,b</sup>	8 <sup>a</sup>	12 <sup>b</sup>	9	10
<b>Red Card</b>	11 <sup>a</sup>	10 <sup>a</sup>	2 <sup>b</sup>	8 <sup>a</sup>	8	5
<b>Yellow Card</b>	5 <sup>a</sup>	6 <sup>a</sup>	3 <sup>b</sup>	5 <sup>a</sup>	5	6
<b>Yippie!</b>	12 <sup>a</sup>	9 <sup>a</sup>	21 <sup>b</sup>	11 <sup>a</sup>	13	9
<b>Boring</b>	5 <sup>a</sup>	5 <sup>a</sup>	7 <sup>a</sup>	6 <sup>a</sup>	6	6

play on advantage in a specific situation. “Yippie!” and thumbs up and down were also used regularly with around 9-15% across all games. The remaining buttons are rather specific to particular moments in the game and, thus, were used less often. Figure 8.3 shows the button usage for match 51. It depicts that fans of opposing teams evaluate the match differently. While fans of the winning team, Germany, used the vuvuzela and the Yippie! button significantly more often, the English fans expressed their disappointment with a clearly higher use of the thumbs down, the whistle, and the yellow card.



**Figure 8.3:** Relative frequency of button clicks during match 51 (England vs Germany) split by team. Stars (\*) indicate a statistically significant difference between both fan groups ( $p < .05$ ) [174].

### *Qualitative Feedback*

The questionnaire was introduced with the last update as a link to a web page, which people could access from within the app on their mobile phone.

In total 46 users (mean age=20.1, SD=62.39, 6 female users) replied to the questionnaire. Based on the feedback from participants, 37% followed the World Cup matches frequently and 50% watched occasionally. Further, 55% considered themselves as knowledgeable fans and 30% as experts (knowledgeable of players' details). Thirty-three participants stated that they normally watched the matches at home and 30 watched with family or buddies.

The participants were asked about their behavior on using the *World Cupinion* application. In total, 18% of the participants stated that they used the app for most matches, 40% used it regularly, and 42% occasionally or just once. Also, those who considered themselves to be knowledgeable or expert fans used the app more frequently. Sixty percent of those who watched the matches in a group (with family, buddies, or in a crowd) still used the app regularly or for most of the matches. Additionally, participants rated the app in

general. The average rating was 3.6 out of 5 (SD=7.46, Median=4). Those who used the app more regularly rated the app higher.

We also asked participants to rate if the level of fun and connectedness to other fans changed while using the app. Eleven percent (11%) mentioned that the fun aspect did not change at all. Thirty percent (30%) believed that the fun aspect increased sometimes and 59% reported to have more fun most of the time or (almost) always. None of the responses indicated that the app reduced the fun of watching. Also, 7 out of 46 participants (15%) did not feel connected to other fans at all. Thirty-two percent felt a (very) strong connection and 53% average/little bit connection. Those who had more fun felt more connected to the other fans (Spearman's  $\rho = .636$ ,  $p < .001$ ).

## 8.4 Implication

The results of the user study reveal that it is possible to connect TV viewers through their mobile phones. This communication channel can be used to exchange sentiments nonverbally and in real-time. The iconic user interface allows user to share opinions that are related to the events. It further enables instant reactions to the live events and minimizes user distraction when operating the mobile phone; whereas text input requires more time and cognitive resources.

The use of four expressive icons “Vuzuzela”, “Yippie!” and “thumbs up/down” comprise more than half of all clicks across all games (57%, see table 8.2. This is an indicator that the app is used as intended, or to share experiences/opinions. Fans of opposing teams clearly vary in term of how they evaluate a specific game. This result is not surprising, but did serve as a “sanity check” and was the first step in the direction of identifying group-specific rating patterns. For more a detailed analysis, e.g., evaluative meaning of ratings in reaction to activity of the opposing fan group or depending on momentary location, a larger data set is essential.

One of the challenges was establishing a shared TV watching experience that was meaningful and engaging to users. The soccer World Cup 2010 provided a good setting for creating such an experience. The event is quite popular in many parts of the world and extends over four weeks. On the other hand, as the app was released to the public, large number of users might use the app that led to high traffic and performance issues. The technical realization should be capable of dealing with a large number of devices and provide instant feedback. Users expect that the application would work at any moment. Therefore, permanent maintenance and monitoring are crucial and should already be considered during such studies.



### *Limitations*

In general, studies in-the-wild are essentially conducted in an uncontrolled environment. There is no direct contact between instructors and participants. Hence, it is not possible to directly describe the procedure of the study. Further, there is no control over participants to ensure that the application is used in the intended way or used at all. On the other hand, as participants are recruited via application marketplaces, the background (gender, language, profession, nationality etc.) of the subjects is completely out of researchers' control. Users may be much larger in numbers, but still demographically biased.

Even though large number of users downloaded the apps, the fact that only a subset of those downloads were still "active" at the end of the study already demonstrates out that a high dropout rate has to be taken into account. The number of users use the application is decreasing over time even if the events became more and more thrilling, as in our case. This appears to be another example for the novelty effect that has been reported before for Social TV [93] and mobile phone applications alike [40].

## 8.5 Summary

In this chapter we assess how iconic user interfaces can be used for expressing and sharing sentiments nonverbally and connect non-colocated users together. Through a use case, we investigated whether real-time opinion sharing about TV shows through a nonverbal (non-textual) iconic user interface on mobile phones is feasible and reasonable. Even with the rise of the World Wide Web, TV has remained the most pervasive entertainment medium and is nowadays often used together with other media, which allows for active participation. With advances and pervasiveness of telecommunication the idea of connecting non-colocated TV viewers via telecommunication technologies, referred to as Social TV, has received considerable attention. Such systems typically include set-top boxes for supporting collaboration.

Further, we assessed how such a nonverbal communication channel could impact the experience of watching TV among non-colocated viewers. We considered the soccer World Cup 2010 as the TV show for establishing a shared setting among users. This event is one of most popular sporting events in many parts of the world and extends over four weeks. It allowed us to recruit a large number of users for sharing opinions in large user communities in real-time.

To achieve our research goal, we designed and developed a mobile app including an iconic interface. The iconic interface includes icons relevant to a soccer match. We released the app in the app stores in order to recruit participants. The results reveal that

the iconic interface allows users to share their opinion about the event. The sentiments collected, indeed, correlated with what happened in the event. This information, for example, can be used to annotate the event and extract important and interesting moments. Further, such a communication channel connects non-colocated users and provides them with the possibility to exchange their opinions nonverbally. Sharing opinions thereby increases the TV viewers' sense of connectedness and enjoyment. Remarkably, even users who watched the matches in groups still used the app to virtually connect to non-colocated fans. Anecdotal evidences showed that the implicit action of an ambient vuvuzela sound – amplified sound when the majority of fans had pressed the Vuvuzela button – resulted in an “aha” reaction in viewers and promoted the conscious experience of connectedness between viewers.

The results suggest it is possible to use iconic interfaces to acquire nonverbal information. Such interfaces can be used for exchanging and sharing sentiments nonverbally in real-time. Sharing this information can result in increase of connectedness between non-colocated users. The type of application proposed and evaluated here can be useful for other types of events, e.g., election debates, quiz shows, contests, etc. Providing real-time feedback directly on the TV while watching events can encourage users to contribute even more.





V

CONCLUSION



# Chapter 9

---

## Guideline for Research in the Wild

We discuss in Section 2.4 different research methodologies for assessment of hypotheses and addressing research questions. While studies in laboratories with a controlled environment is one way of conducting evaluations, researchers have tried to increase the external validity of findings by carrying out studies in more realistic contexts. Early research indeed shows the importance to conduct in-situ experiments when analyzing mobile usage behavior [144]. The emergence of application stores has provided the opportunity to conduct studies in-the-wild with a large number of users. We also used this approach to answer our research questions in Chapter 5 and 8. However, there has been little investigation regarding a guideline on designing and conducting such studies.

Prior work has highlighted certain challenges of large scale studies conducted in the wild based on their experiences [85, 129]. They mainly focus on either on specific aspects or are limited to the respective authors' experience. What still missing is a comprehensive overview of the lessons learned from all the different studies conducted in-the-wild and how research can address the identified challenges. Based on the experiences we gained during carrying out user studies reported in [164, 165] and this dissertation as well as the analysis of related work, we provide best practices on how to conduct studies through applications stores. We identify challenges and limitations of such studies. The

guideline can help other researchers who would like to use this methodology to address their research questions.

*This chapter is based on the following publication:*

- N. Henze, A. Sahami Shirazi, A. Schmidt, M. Pielot, and F. Michahelles. Empirical research through ubiquitous data collection. *Computer*, 46(6):74–76, 2013

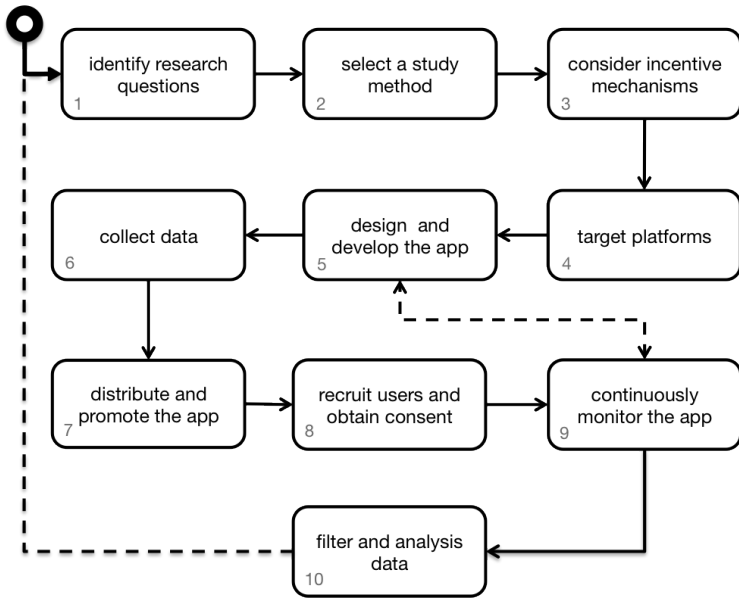
## 9.1 Research in the Wild: 10 Steps Practices

The emerging area of research in the wild enables researchers to investigate research questions through studies conducted in the relevant context with a large number of diverse participants in a variety of situations. Applications (apps) are used as apparatuses and mobile application marketplaces are prime for the recruitment of participants. In the following section, we provide practices identified for conducting large scale studies using applications stores based on the analysis of prior studies and studies conducted by ourselves. We describe these practices in ten steps 9.1. We, first, describe the diversity of research questions that have been investigated and the methods that have been used to conduct these types of studies. We discuss potential incentive mechanisms and target platforms. We present important aspects that need to be considered when developing the apparatus and recording the data. We review approaches to distribute the app and recruit participants. Finally, we look at important aspects regarding continuously monitoring the app and the data analysis. In the following section, we describe each of the aspects in detail.

### *1. Research Questions*

The main goal of conducting a study is to answer one or more research questions. Research in-the-wild has been used to study a truly broad span of research questions, from investigating general user's behaviors to specific questions. Henze et al., for example, used Fitts' Law to model and examine basic human motion performance [86]. Böhmer et al. investigated the time of the day apps on mobile phones are used [15]. Sahami Shirazi et al. assessed notifications received on mobile phones [165]. Girardello et al. assessed how to automatically estimate apps' quality by observing a large number of users [67]. Oliver examined how users consume energy on their mobile devices [142]. The findings of these examples are not necessarily useful to refine the study of a specific





**Figure 9.1:** Steps to conduct research through app stores, from identifying research questions to filtering and analyzing the data.

app, but generally increase our knowledge about mobile user behavior and ways to improve the interaction.

The method has been also used to investigate very specific questions. In Chapter 8, for example, we investigated how mobile phones can be used as a communication channel for exchanging non-verbal information that represents emotional reactions to events shown live on TV. Pielot et al. used a widely disseminated app to study the effect of tactile feedback for mobile navigation systems [152]. Cramer et al. used the method to study Spotisquare, a novel mash-up of the location-based service Foursquare and music streaming service Spotify [45]. The answers are mainly important for specific application domains.

As for any study, it is essential to clearly identify the research questions beforehand. Overall, the conducted research shows that research in-the-wild is not limited to a specific application domain. It has been argued that it is not suited for research questions that require collecting subjective feedback. The method is, however, limited to questions that do not require providing participants with additional hardware beyond a standard smartphone.

## 2. *Study Methods*

We reviewed various research types in Section 2.4. In correlational research, researchers typically observe a phenomenon without interfering with it. In experimental research, researchers manipulate one or more aspects to study its effect on variables of interest. While correlational research can help to identify and describe phenomena, experiments are used to explore the underlying causes.

Most of the in-the-wild studies carried out and reported in prior work are correlational studies. Examples include previous studies on smartphone app usage [15, 67]. While these studies let us learn about the contextual factors that correlate with the use of apps, they cannot reliably isolate the cause. To test specific hypotheses, one approach is to validate suspected cause and effect relationships by complementing observational with experimental research [86, 87]. For example, Henze et al. first studied how people interacted with smartphones (observational), developed improved interaction techniques, and ultimately showed that they increase performance compared to the state-of-the-art (experimental). However, since research in-the-wild apparatuses are used for certain purposes, some experimental manipulations may not be feasible. For example, when announcing a novel tactile interface as a key feature [152], users might not accept that this interface is not available to them, because they are assigned to the control group. Hence, some experiments will have to allow users to switch between conditions, which turns them into quasi-experiments. While this may be necessary in order to not lose participants, this setup makes it more difficult to rule out confounding variables, i.e., participants who are more curious tend to use tactile feedback more often, but are simultaneously also better navigators.

Thus, if possible, studies in-the-wild should try to use pure observational or experimental approaches. However, at times, it may be necessary to sacrifice internal validity in order to attract and satisfy participants.

## 3. *Incentive Mechanisms*

When participants volunteer to take part in a study conducted in a lab, they are usually motivated. However, for studies in-the-wild users first need to get motivated to participate in the studies. Therefore, incentive mechanisms must be considered when designing the study. Incentives encourage users to start and keep participating in the study, which is crucial for a successful experiment. Various incentive mechanisms can be considered. For example, In the study conducted in Chapter 8 we used fun and access to information as incentives for their users. Henze et al. [87], McMillan et al. [118], and Kranz et al. [104] considered fun through playing games as an incentive mechanism for encouraging users. Providing utilities as well as microblogging are other mechanisms to be considered.

We used these two mechanism together to encourage our users taking part in the study described in Chapter 5. The app allowed users to share micro information with their social networks.

The economic incentive is another mechanism. Frei reported a comparison between paid and free crowdsourcing [62]. The results show that in most application domains the paid crowdsourcing can produce better completion rate and processing time than free crowdsourcing. While the quantity of participants can be increased, the quality is not necessarily also increased. Other mechanism such as gamification of non-game apps [114, 113], and social psychological incentives, e.g., historical reminders of past behavior or ranking of contributions can also be considered [33]. The concept of gamification is to integrate game mechanics and game thinking into non-game applications. The mechanics used to gamify a system include point systems, badges, rewards, levels, etc.

In comparison to economic incentives, gamification is a non-monetary incentive that requires light cost when operating. It attracts fewer malicious users which can lead to an increase in quality of results. Anint also suggested ways of taking advantages of the positive social facilitation and avoiding the negative social loafing [9].

### *4. Target Platforms*

To target and recruit many participants for a study, the apparatus used must be easily accessible. Hence, developers need to decide which platform they want to target. Various platforms such as the Web, iOS, Android, and Windows Phone are available. The platform selected can impact the number of users available. Web-based apparatuses can be easily accessed through the Web. Cramer et al., for example, used a web-based apparatus for their work [45]. Users only require a web browser to use the system. They can use smartphones, PCs, or any other devices with a web browser. While the number of potential users is virtually unlimited, it does not have the advantages of the distribution through application stores.

The availability of software development kits for all smartphone platforms allows developers to implement applications and deploy them on smartphones and tablets. iOS, Android, Windows Phone, and BlackBerry operating systems are the most common smartphone platforms. While the iOS platform and BlackBerry OS are used on a small set of device models, Windows Phone and, in particular, Android run on many different devices. Such a variety enables researchers to target users with very different devices. Addressing a variety of devices can enable researchers to test hypotheses with different device characteristics, such as screen sizes (e.g., [86]). Supporting this variety, however, requires considering all variants during development and data analysis, both of which can be a burden for the developer.

The platforms allow access to different information and enable to develop different services. Android, for example, allows background services while this is not possible on iOS devices. Apparatuses used in [15, 67, 87] are developed for the Android platform and McMillan et al. [118] used the iOS platform. The apparatus we described in Chapter 8 supports both Android and the Web to target more users. Oliver reported that there is no single platform to do everything reliably [141]. The choice of platform therefore depends on the requirements that result from the research questions that have to be addressed.

### *5. App Design & Development*

For developing the apparatus, it is necessary to turn the research question into an app that has the potential to be successful and attracts users. Therefore, it is often required to find a compromise between research and attractiveness for users. For example, Henze et al. reported that during the design of their game they had to find a balance between providing players with a game that was worth playing and a test application that collected meaningful data [86]. Pielot et al. also sacrificed validity to make their navigation system more attractive. They reported that they decided against conducting an experiment that randomly assigns users to conditions and allows them to select the conditions in order not to confuse or annoy the users [152].

Based on a case study Ferreira et al., unsurprisingly, reported that stability, reliability, usability, and performance of an app are crucial for user acceptance [60]. This is in contrast to research prototypes that are used in controlled lab and field studies where prototypes often only need to work for a short time in a controlled environment. Zhai et al., for example, already had prototypical implementation of their app's core features available when they started developing an app for the app store [212]. Yet, they reported on intensive testing and iterative releases. Platform characteristics can significantly increase the required development effort. For example, Han et al. claimed that the fragmentation of the Android ecosystem causes portability and compatibility issues within the entire Android platform, which increases developer workload, delays application deployment, and ultimately disappoints users [77]. Thus, extensive testing using a range of different devices might be required. Various aspects should be considered during implementing the app. Apps with noticeable energy consumption might, for example, not be accepted by users. A noticeable or perceived decrease in battery life can encourage users to uninstall the app. This is particularly relevant for background services that permanently collect data.

On a user's device, the developed app does not necessarily run on its own: other applications can affect the behavior of the app - it is not isolated. Oliver, for example, expressed concerns that other software installed on a device can affect the data recording [142]. Oliver suggested that apps that can potentially be affected by other processes should

record which other apps are installed on the device. This approach might, however, induce privacy issues and raise ethical concerns. Oliver further reported that time synchronization can also be a challenge. The clock on a smartphone can be updated with the timestamp broadcast by the cellular network as well as by plugging the device to a computer. A reliable reference clock becomes crucial if the activities in an app have to be synchronized with external events, e.g., the app presented in Chapter 8.

In developing the apparatus, researchers have to ensure that they do not trade too much validity for increasing the attractiveness of their app. As the app will be released in the wild, it is not surprising that extensive testing using different devices and software configurations is required and essential. In particular, developers must expect that other applications and external factors can influence the behavior of their apparatus.

### 6. *Data Collection*

One of the core challenges of conducting studies via application stores is that there is no easy direct contact between the researcher and the participant. Thus, common approaches of collecting data, such as think-aloud or observation, cannot be applied. Data has to be collected in other ways.

Previous research has explored various ways of data collection. First, researchers can monitor the device status. For example, Böhmer et al. [15] used a background service to collect data for when users open and close applications. Second, the interaction with the device itself can be logged. One example is the logging of touch events [86]. Third, studies have made use of the mobile phone's sensors to collect information about usage and context, as, for example, logging the user location and travel speed [152]. When using sensors, it is advisable to consider their additional battery usage. Users can be very sensitive to increase battery drain by an app [152]. These three approaches allow researchers to learn about the "what", but oftentimes not the "why". Hence, researchers have investigated different ways to additionally elicit qualitative feedback from the user. For example, Church and Cherubini [38] explored the use of Experience Sampling. McMillan et al. [118] used Facebook to contact participants and invite them to Skype interviews.

Typically, researchers may want to collect data in the least obtrusive way possible. Hence, it makes sense to automatically collect data through the device where possible. When using this approach, researchers must be careful with the use of additional resources, bandwidth, and battery in particular. Furthermore, at times it may be necessary to clarify the meaning of the data. In these cases, interviews, experience sampling, diary studies, or local small-scale replications of the experiments might be needed [128].

### 7. *Distribution and Promotion*

A successful experiment requires a set of participants. It is crucial to recruit enough participants. Thus, the apparatus should be distributed and accessible to many users. Based on the platform chosen, different distribution channels are available. The emergence of application marketplaces provides the chance to make an application accessible to millions of users. Researchers use available marketplaces to distribute their apps among users and, thus, recruit potential participants for their user studies. Releasing and distributing apps through marketplaces have various advantages. It is possible to easily maintain the apps. A new version of an app can be easily released and distributed among users who have already installed the app. Furthermore, marketplaces have a developer console, which provides different information about the status of released apps, e.g., the number of downloads, the number of installs, the errors reported by users, etc. This allows developers to identify possible bugs of an app, fix them, and release a new version. It is also possible to specify the countries in which an app should be released. For the Android platform, it is even possible to release an app for a specific set of Android mobile phones based on their features.

Chrome Web Store <sup>16</sup>, an online marketplace for the Google Chrome browser, is a possible channel to distribute web-based apps support the Chrome web browser. Firefox Marketplace <sup>17</sup> is also the marketplace for the FireFox browser.

For the Android platform various marketplaces are available. The largest Android market is Google Play <sup>18</sup>, the Google's official Android market. This channel is used in many projects (e.g., [15, 87, 152, 165]). Since apps are published in this market without review, the publishing and distribution process is very fast. There are other Android marketplaces such as the Amazon Appstore <sup>19</sup>, the official Amazon Android store, available for distributing Android applications. Further, Henze et al. assessed installations of several thousands games and apps on the Google Play market and suggest that the best time to release a game is Sunday evening GMT [84]. Möller et al. raised the awareness for a potential slow to the update propagation on the Android platform [126].

McMillan et al. provided a comparison of distribution channels for large-scale deployments of iOS apps [119]. Apple Store is the official store for the iOS platform. Cydia <sup>20</sup> is another channel for distributing iOS apps on Apple devices that are jailbroken. However,

---

<sup>16</sup> Chrome Web Store: <http://chrome.google.com/webstore/>, last accessed August 27, 2014

<sup>17</sup> Firefox Marketplace: <http://marketplace.firefox.com/>, last accessed August 27, 2014

<sup>18</sup> Google Play: <http://play.google.com/store/>, last accessed August 27, 2014

<sup>19</sup> Amazon Appstore: <http://amazon.com/mobile-apps/>, last accessed August 27, 2014

<sup>20</sup> Cydia Appstore: <http://cydia.saurik.com/>, last accessed August 27, 2014

McMillan et al. reported that there are problem with user density within the jailbreak community. On the other hand, apps uploaded and released in the Apple Store are subjected to a lengthy review process [124].

Promoting the app is the next step after releasing the app in order to recruit enough users to install and use the app. Several approaches can be used to increase the number of users. Localization of the app in popular languages can target different users all over the world. One free and easy approach to advertise the app would be through social networks such as Facebook, Twitter, and Google+. Also announcing the app via possible mailing lists could help to introduce the app among the peers. Blog posts and reviews can recruit users, too. Reviews and blog posts, especially by popular bloggers or reviewers, could aid in promotion of the app. Sahami Shirazi et al. described that by featuring their app in well-known technology blogs helped them to reach a growing number of users [165].

Another channel is using advertising platforms. Many advertising platforms are available that allow researchers to promote apps. AdMob<sup>21</sup> is an advertising platform from Google that offers an advertising solution for mobile phone platforms and mobile websites. It allows developers to promote their apps with advertisements. iAd<sup>22</sup> is Apple's mobile advertising platform for its mobile devices. It should be considered that it is essential to have enough participants for a successful experiment. It is recommended that potential possibilities be used to recruit numerous participants.

## 8. *Recruitment and Consent*

Attracting a large number of users does not necessarily result in a large number of participants. Henze et al. analyzed five of the apps they used to conduct studies and looked at the percentage of users that contributed actual data points [85]. For one of the apps, they used a rather complicated approach. A menu item is added after a user used the app for a certain time. Only after selecting the menu item, agreeing to be part of the study, and filling out a form with demographic information, a user was turned into a participant. As a result, only for 0.46% of the installations of the app was turned into a participant. This is dramatically small compared to one of their games that did not inform users at all. For this game, data was collected from 83.68% of the installations and, thus, the user considered a participant.

Chalmers et al. identified ethical challenges surrounding large scale user trials [31]. They highlight that researchers need to ensure that users provide informed consent before becoming participants. Pielot et al. compared four approaches to inform participants

---

<sup>21</sup> AdMob: <http://google.com/ads/admob/>, last accessed August 27, 2014

<sup>22</sup> iAd: <http://developer.apple.com/iad/>, last accessed August 27, 2014

that they will participate in a study when playing the game [151]. Not surprisingly, they showed that just using a modal dialog with a single okay button results in the highest conversion rate. Morrison et al., however, investigated if participants were really aware what is being recorded, even if this is clearly stated in the app [127]. They record the participants' location and asked them about their opinion. Morrison et al. found that showing a map with the participants' position increased their concerns, which might imply that participants were not fully aware of the potential implications of recording their position.

Most researchers use a pragmatic approach when conducting a study. The apps and games developed by Böhmer et al. [15] and by Henze et al. [86, 87] mentioned the propose of the user study in the apps' description in the application store. Several other work such as [118, 15] informed the users when the app was started for the first time. Further, they enabled users to refuse to participate in the study and still use the app or game. Providing additional information in a section of their apps or on an external web site is another approached used, e.g., [152]. No consensus has been reached on how users should be informed or asked. It seems commonly agreed, however, that researchers have to at least try their best to inform the users. Multiple approaches should be combined to increase the probability that users can make an informed decision. Users need to have the possibility to not become a participant either by opting out or by not using the app.

### *9. Continuously Monitor the App*

In contrast to studies that are conducted over a short period of time, such as experiments in a usability lab, a research in-the-wild study typical last months or even years. It therefore requires the researchers to monitor the app over a long period of time. In addition, it can become necessary or desirable to refine the app. Previous researchers engaged in these types of studies have frequently updated their apparatus [15, 60, 85, 87] and even provide suggestions about the best time to do so [84]. Zhai et al. reported that they made seven releases of their notepad application in a few months [212] and Kranz et al. reported that they released 21 updates for one of their apps [104].

It is essential to continuously monitor the app to ensure that it runs sufficiently and collects meaningful data. A sudden increase in number of installations can significantly increase the amount of data that is collected. As the number of installations is difficult to predict, additional resources can suddenly be required to handle a large amount of data. During an ongoing in-the-wild study, in particular, it is crucial to monitor the app to ensure it works sufficiently. In the case that updates are required or intended, it is important that data from different versions are not mixed up. In addition, not all users will update. For the Android platform, Möller et al. reported that not all users update their apps and it could take weeks until the majority adopt a new version [126]. Limited



resources might become apparent only after an app is released for quite a while, but is likely to require immediate attention.

### *10. Data Analysis and Filtering*

The data collected usually needs to be filtered, as not all data may reflect the kind of usage that should be studied [45, 85]. While a traditional user study makes sure that all participants contribute roughly the same amount of data, this cannot be guaranteed in research conducted in-the-wild. It is important to recognize such effects to avoid drawing invalid conclusions.

Most applications users just start the application once, while a few power users contribute the bulk of the data [85, 129, 152]. To avoid biases, researchers need to decide whether to exclude users that contribute too much or too little information to the data set. Morrison et al. [129] highlighted that one solution is to “make a virtue of a large number of single use participants”. Another solution to this challenge is to visualize and carefully inspect the data for undesired use. If the data itself is not conclusive, the in-the-wild study should be accompanied by a small-scale, local study, as suggested by Morrison et al. [128].

## 9.2 Limitations & Challenges

Similar to studies conducted in laboratory and controlled settings, studies carried out in the wild have their own challenges and limitations. An unsupervised and uncontrolled setting results in the issues discussed in the following part.

### *Uncertain Number of Users & Demographic Information*

In studies conducted in-the-wild, after the app is distributed, it needs to be promoted to recruit participants. However, there is still a risk that only a few users download and use the app. This can depend on many factors, mainly the incentive mechanism used, promotion, and distribution of the app. For example, in five case studies presented by Henze et al. [85] one of the apps was completely unsuccessful and with another app, only little data could be collected.

As there is no direct contact with participants in studies conducted in the wild, collecting demographic information is critical. One possibility is to ask users about their demographic data through a form as we presented in the Chapter 5. However, it can be possible that users do not provide real, accurate information [118]. Another approach, which is more reliable, is authenticating users using social websites. With this approach it is

possible to access information from the social website. Furthermore, it allows researchers to make contact with users. With this approach, it is possible to contact users [118].

### *Unforeseen Usage & Noisy Data*

While it can be ensured that participants follow the study procedure in lab settings, experimenters do not have any control over participants during a study in the wild. Hence, the unsupervised use of apps can lead to unforeseen usage patterns. Users may get interrupted while using the app [152]. Thus, it is essential to detect unpredicted usage patterns and minimize their effects in order to keep internal validity. Though it is possible to collect a fairly large amount of data during studies conducted in the wild, not all data will be useful. The uncontrolled setup can lead to users not completely contributing in the experiment or dropping out during the study. Thus, the data collected is noisy and needs to be filtered. For example, only 13% of the data collected during the study described in [152] was used for further analysis.

### *Qualitative Feedback*

It is common to collect qualitative feedback during user studies in the lab. Cooper et al. discuss that understanding the user cannot be achieved by digging through the piles of numbers that come from quantitative study [41]. So collecting qualitative feedback during studies in the wild is valuable, but, at the same time, not straightforward. Several channels are available to collect qualitative feedback from users. Collecting comments from marketplaces is one possibility. The Google Play market, for example, allows users to rate apps and leave comments. In [85] comments received on several apps used for studies are assessed. The authors stated that the comments mainly report technical problems and did not provide any feedback relevant to the research questions. However, technical comments can be used to identify bugs [60]. On the other hand, Zhai et al. reported that they indeed collected comments useful to their research questions[212].

Another approach is embedding a feedback form or a questionnaire in the app [85]. We embedded a questionnaire in the *Word Cupinion* app and collected qualitative feedback (cf. Chapter 8). McMillan et al. suggested rewarding users for providing feedback using in-game badges or bonuses [118]. Further, they stated that it is feasible to get indirect contact with participants, e.g., using Facebook and providing vouchers for participation in phone interviews.

# Chapter 10

---

## Conclusion

Communicating and sharing emotions is one of main purposes to fulfill humans' social needs. The communication includes not only verbal information, but also nonverbal information. Indeed, nonverbal information makes up the majority of communication. The advances in technologies emerge new mediums for communications. Users are able to communicate together independent of their locations. Such computer-mediated communication is mainly non face-to-face in different contexts. As such, this type of communication lacks the nonverbal component of communication, which can lead to confusion and incoherence. Further, communicators in both sides do not perceive contexts in which communication is taken place. Obtaining and sharing contextual information between non-colocated can enhance the communication. In this dissertation, we investigated how to obtain context awareness using certain sources. We further explored how contextual information can be shared and awareness can be conveyed nonverbally. Table 10.1 includes the research questions answered in this thesis.

### 10.1 Summary of Contributions

With the ubiquity of computing technologies, users use devices such as smartphones in different contexts. Sensors embedded in such devices, as well as available sensing technologies, allow researchers to retrieve information about the context in which the technologies are used. Further, the user himself can be a resource for retrieve contextual

information. In Part III of this dissertation we utilized commercial brain user interfaces (BCI) to acquire brain signals and obtain emotional information of users. On the other hand, we presented how it is possible to obtain awareness about a certain activity through users explicit interactions with a mobile phone instead of using any sensor attached to the user.

Further, the availability of pervasive Internet connectivity provides the opportunity to share context and awareness information in real-time between users and connect them together. In Part IV of the dissertation we discussed how context and awareness information can be shared using nonverbal modalities. We presented that tones can be used as a nonverbal modality to express and share context and awareness. Indeed, the intention of text messages can be conveyed through sonification. Further, we showed how iconic user interfaces as another nonverbal communication channel can be used to share sentiments nonverbally and connect non-colocated users together.

Using application stores and conducting several user studies in the wild, further allowed us to identify challenges and limitations of this research methodology. We provided a guideline outlining how to conduct such studies based on the experiences we gained while carrying out our studies as well as by reviewing prior research that used similar methodology.

Table 10.1: Overview of Contributions to Research Questions.

---

**No. Research Question**

---

**I. Exploit Awareness**

---

**R1 Is it possible to implicitly determine a set of users activities, i.e., reading and relaxing, by using only brain signals?**

We conducted a user study and collected brain signals for certain activities. We used a commercial brain-computer interface to acquire brain signals. We investigated the classification of reading and relaxing activities from other common daily tasks based on only gestures extracted from brain signals. The results reveal that such sensors can be used to obtain context information about certain users activities (Chapter 3).

---

*Continued on next page ...*

Table 10.1 – . . . Continued from previous page

No.	Research Question
R2	<p><b>How does information acquired from brain signals correlate with emotional states of the user by using videos as stimuli?</b></p> <p>Through a case study we assessed how emotional information obtained from a commercial brain-computer correlates with users emotional state. We conducted a user study and asked users to explicitly indicate when they get excited while watching a movie. In parallel, we recorded the emotional information acquired from the BCI device. The results reveal that this information indeed correlates to each other. Based on the results, we proposed an algorithm that extracts moments in which users get excited. The results suggest that BCI devices provide reliable emotional information about users. The emotional information can be shared in computer-mediated communication to enhance the communication (Chapter 4).</p>
R3	<p><b>How is it possible to monitor user's sleep activity based on only explicit interaction with a mobile application without using any sensor?</b></p> <p>Common approaches for monitoring user sleep behavior use actigraphy devices. In contrast, we explored monitoring such an activity using solely user explicit interactions with a mobile application. We also provided users feedback about their sleeping activities. The results revealed that we could, indeed, be able to monitor users' sleep behavior. Further, providing feedback increase users awareness about their behavior. This may induce users to ward healthier behavior (Chapter 5).</p>
<b>II. Sharing Context nonverbally</b>	
R4	<p><b>Is it feasible for users to express and share their emotion through the composition of a melody as a nonverbal mean for communication?</b></p> <p>A melody composer was developed allowing users to compose melodies and use it as nonverbal means to communicate and share their emotion with others. The result of a controlled study unveiled that this form of communication can be used to express and share emotions, similar to crafted pieces of arts. The composer can be also used for describing the current context and provoking memories (Chapter 6).</p>
R5	<p><b>How can audio previewing of a text message convey and share its intention nonverbally?</b></p> <p>We explored conveying information about text messages and creating awareness on their content by providing audio previews. In the first step, we assessed audio previews that were based on certain frequent keywords and emoticons. The results showed that the audio previewing approach could be used to convey information nonverbally. Further, it impacts users behavior in checking short messages (SMS) as well as users intentionally to use more characters, which are audio previewed to increase awareness. The results of the study in mobile context encouraged us to explore the same approach for instant messaging in the desktop context. We further extended the audio previewing approach in such a way that intentions and contents of messages are conveyed as well. Indeed, the results revealed influences on user's behavior. In contrast to the mobile context, messages including questions needed more time to be checked. Thus, the importance and priority of messages are completely dependent on the context (Chapter 7) .</p>

Continued on next page . . .

Table 10.1 – ... Continued from previous page

---

**No. Research Question**


---

**R6 How can iconic interfaces be used to share sentiments nonverbally in real-time?**

We explored the feasibility of using iconic interfaces for exchanging sentiments between non-colocated TV viewers in real-time and connecting them together. We chose the World Cup 2010 tournament as a shared event among the viewers. We developed mobile applications that included an iconic interface for sharing opinions about moments during a soccer match. We conducted a study in the wild with a large number of users. The results unveiled that iconic interfaces can be used to nonverbally and instantly share sentiments. Sharing this information increase awareness and connectedness among users. The sentiments further correlate with the key moments in an event (Chapter 8).

### 10.1.1 Exploiting Awareness

Sharing context between non-colocated users in computer-mediated communication can lead to an increase in the communication quality. The initial step, however, is to exploit context information about users. Context information can be related to users' surrounding environments. This information can be obtained by embedding sensors and sensing the environment. Prior work has explored various sensors and sensing technologies to retrieve such context information. We focused on obtaining context information using physiological information. In this thesis, we explored humans' brains as well as their explicit interactions with a device/application as two origins for exploiting context information about users.

**Brain signals can be used as a source for implicitly obtaining context.** We investigated using the brain, as the central organ of the human nervous and intelligent system, to exploit context and awareness. Through a controlled user study and using a commercial off-the-shelf brain-computer interface (BCI) we collected data from certain daily activities. We presented that users' activities, i.e., reading and relaxing can be determined from other daily activities using brain signals. While most of prior work suggested a user-dependent classification, we showed that classifying activities and obtaining mental awareness independent of users is feasible. Hence, it is possible to develop context-aware systems without any calibration. Further, the results reveal that mental awareness can be retrieved implicitly without adding any additional (mental) cost to users.

**Commercial BCIs can be used to acquire emotional awareness.** We further assessed how the emotional information commercial BCI interfaces provide correlates with users emotional states. In a case study we carried out an experiment in which emotional

information (i.e., excitement) obtained from a commercial BCI and explicitly by users were assessed. The results unveiled that information obtained from the BCI can be used to detect moments which users believe they got excited. Based on the findings an algorithm was developed that can be used to annotate videos and generate a summary of highlights. The results, in general, suggest that commercial BCI devices provide valuable information to obtain contextual information and particularly emotional information about users. This provides the opportunity to develop systems that leverage mental and emotional awareness. Such systems can be used outside of laboratory environments and in the daily life.

**Explicit interaction can be leveraged to implicitly obtain context.** We also explored using users' explicit interactions instead of using any sensor to retrieve context information. We developed a mobile application that monitors users' sleep behaviors based on their interaction with the application. We presented that users' explicit interactions is a rich source to determine state of users and exploit contextual information about users' sleep behavior. This approach is unobtrusive and does not require wearing any physical sensor. The results of the study further unveil that providing users feedback about their sleep behavior can encourage them towards healthier behavior. The results suggest that interactive applications can be designed in such a way that users' context is derived without using any physical sensor. Providing users with feedback about their activities can also lead to the increase of awareness. In particular, providing users feedback about behaviors, which are related to their health, can increase their awareness and encourage them toward changing behaviors.

### 10.1.2 Sharing Context Nonverbally

In computer-mediated communication it is essential to share context information among non-colocated users to enhance communication. While various approaches are available to share such information, we were particularly interested in nonverbal means. We investigated two nonverbal modalities to share information: *rhythmic tones* and *iconic user interfaces*. We used mobile phones as the communication channel for exchanging and sharing information due to the ubiquity of mobile phones and pervasive Internet connectivity available on them.

**Rhythmic tones can be used to express and share emotions.** We explored melody composition as a nonverbal means to express and share emotional information. The results of the user study show that rhythm-based tones as nonverbal means can be used to convey information. Self-composed melodies, similar to crafted pieces of arts, can be used to express and share emotions between non-colocated users. They indeed have a stronger impact than previously composed melodies. Further, such form of

communication can be used as means for creativity, fun, and teasing which results in group cohesion. It can be used for describing the current context as well as provoking memories.

**Message intention can be conveyed by mapping texts to tones.** We further used rhythmic tones modality for sharing communicating information. We assessed how the audio previewing of text messages can transfer and share information about the messages. We, first, explored this approach in the mobile context. We sonified short messages based on certain keywords. The results revealed that sonification can effectively communicate the content of the messages. Indeed, it impacts user behavior by checking their messages more or less quickly and increasing their use of certain keywords. Users purposely use keywords which are sonified more frequently to increase awareness on the receiver side. We investigated whether this approach can be also applied on instant messaging on the desktop context. We extended the sonification approach in a way that the message's contents and intention were also conveyed. The results unveil that it is possible to convey more information about contents and intention of text messages using sonification. The impact is almost similar to what we found in the mobile context. However, users behavior on prioritizing and checking/replying is context dependent. Messages including questions are checked faster in the mobile context, whereas users prefer to finish the ongoing task before checking/replying such messages in the desktop context.

**Iconic user interfaces are proper means to share sentiments nonverbally.** We explored the use of iconic user interfaces on mobile phones as a nonverbal means to share and exchange context information. We used this approach for nonverbally sharing sentiments between non-colocated TV viewers in real-time and connecting them together. The results of the conducted study unveil that it is feasible to use an iconic user interface on mobile phones and share opinions during an event in real-time. The sentiments users provided and shared indeed correlated with key moments and could be used to extract highlights based on this information. Sharing and visualizing such information between non-colocated users increased their awareness and connected them together. It further decreases distance between active users and lurkers, who only observe sentiments without actively share information. In contrast to text-based approaches, the iconic interface is a proper way to express opinions nonverbally and quickly with minimal efforts. This approach is not limited to only sharing sentiments. It could also be used for sharing other information such as emotions.



### 10.1.3 Practices for Research in-the-wild

The emergence of application stores has provided researchers with this opportunity to move user studies from controlled laboratory environments in to the wild. Such studies are carried out in uncontrolled environments with thousands of users. They have higher external validity in comparison to lab studies. We also used this methodology to address some of the research questions in this dissertation (Chapter 5 and 8) as well as other research questions [164, 165]. The experience we gained while conducting these studies allowed us to identify challenges and limitations that researchers are likely to face. Additionally, we also reviewed studies conducted in prior work and determine other challenges and shortcomings. Based on the findings, we proposed a guideline for conducting studies through application stores. The guideline includes practices that are helpful in designing such studies. It can help researchers who want to address their research questions using these methodologies.

**Ten steps towards conducting a user study in the wild.** The practices consist of ten steps and describe aspects that should be taken into account for setting up and carrying out studies via application marketplaces. In the initial step, we discuss the diversity of research questions that have been investigated and the methods that have been used to conduct the studies. We consider potential incentive mechanisms and target platforms when it comes to developing an apparatus. We discuss important aspects which need to be considered during the development of the apparatus and data recording. We review approaches for distributing applications and recruit participants. Finally, important aspects regarding continuously monitoring the app and the data analysis are described.

**Limitations are inevitable and multifold.** The limitations and challenges of studies in the wild are multifold. It is very hard to ensure the number of participants and the intended use of the apparatus. Further, collecting demographic information is very hard. Participants may not provide real information or find it violates their privacy. The uncontrolled setting can lead to unforeseen usage. Users may not also completely contribute in the study and drop out. The data collected is very noisy and should be filtered before using for the analysis and evaluation.

## 10.2 Future Work

Numerous sources and approaches have been explored to obtain context information. This thesis provides an overview on using certain information and modalities to exploit and share contextual information. A number of open research questions have been identified while conducting the research presented here. For example, we assessed

rhythmic tones as a nonverbal means to convey information through controlled studies. In these studies, participants used the approach for a short period of time. The findings have high internal validity. However, the sonification approach can be assessed through in-the-wild approach to observe the use in a more realistic context. The study in the wild allows researchers to expose findings with higher external validity. While each study has its own advantages and disadvantages, both sets of findings will be very valuable to constructing a better understanding of using sonification to share and communicate contextual information. In the following section, we point out other possible research directions for future work.

### 10.2.1 Physiological Information

In this thesis, we explored brain signals for obtaining context information. This information gives us insights into mental activities. We investigated the classification of two common activities, i.e., reading and relaxing. However, other common physical activities such as eating, talking, walking, seating, etc. can be considered for investigation. These activities unveil other context information about users. Monitoring physical activities, for example, can reveal the number of calories a user burns daily.

On the other hand, other physiological information can be explored for retrieving context information. Heartbeat rate, body temperature, and amount of oxygen in blood can provide valuable information about humans' body state. This information can be used to develop context-aware systems that monitor users and implicitly obtain information. For example, by using heartbeat information, it is possible to determine when the user is relax or under stress. It can also be used to identify certain physical activities based on physiological information. Commercial sensing technologies available for monitoring physiological states allow researchers to implicitly collect information from the human's body outside of a clinical environment. Such devices can be smoothly integrated into clothes and, thus, be unobtrusive for users. Using this emotional information, similar to calming technologies, systems can be developed which encourage users to be more relaxed and/or reduce information overload and ultimately, increase users attention.

### 10.2.2 Sharing Mental & Emotional Awareness

We explored sharing sentiments explicitly using iconic user interfaces on mobile phones. The results unveiled that sharing this information increase awareness and connectedness among non-collected users. Following a similar direction, sharing mental and emotional information between non-colocated users could also be explored. Sharing such

information may let users to obtain additional information. It is interesting to find out whether sharing such information implicitly and explicitly also impacts users awareness and connectedness. Being aware of the fact that a friend is frustrated, for example, may encourage others to support them. One main challenge is how this information should be visualized and shared between users. Using emoticons or avatars could be one feasible approach to visualize this information. Audio and haptic feedback could be another modalities for communicating this information. Nevertheless, privacy is a crucial aspect that should be considered and further investigated.

### 10.2.3 Beyond Text Sonification

We investigated sonification of instant messages and specific common keywords to communicate information. However, the approach has the potential to be used to convey other information, such as emails, content of images, or emotions. Meta information can be extracted and sonified in a way that it conveys information. The main advantage of sonification is that a decoding medium is not required. Through the sense of hearing, users can receive and decode tones. Particularly, sonification can be beneficial on mobile contexts. Users most of the time have their phone with them. In situations in which users are dedicated to other activities (like driving), sonification is a useful approach to convey information and create awareness while still preserving privacy.

### 10.2.4 Research in the Wild for other Domains

Having thousands of users downloading and using prototypes in their natural environments provides an opportunity to observe users in realistic contexts. Information collected through such an approach allows researchers to gain more insights into users interaction behavior. This knowledge can be of great interest to designers and developers. The current in-the-wild-research methodology is mainly conducted through application stores for mobile devices. However, this method is not only limited to this domain. Applications stores for tablets, PCs, and TVs are already emerged. Even stores for cars, appliances, and public displays are currently emerging. It is expected that the guideline and practices suggested can be transferred to other domains. Nevertheless, it is worthwhile to investigate and highlight other potential challenges in other domains. Findings from both lab and in-the-wild studies allow us to gain insights into users and model their behavior. Such models are precious in designing interactive systems and increasing usability.

## 10.3 Concluding Remarks

This dissertation addresses several research questions in the exciting research field of (mobile) human-computer interaction. With advances in technologies, it is expected that other sources could emerge and be used to obtain context awareness. This can result in new context-aware systems and services becoming available in the marketplace. Furthermore, more sophisticated nonverbal means of communication could be used as a common, or even default, form of communication. Indeed, nonverbal means of communication could convey information once these means are widespread on pervasive devices, such as mobile devices and personal computers.

Furthermore, it is expected that in-the-wild research methodology enables findings not only in HCI, but also in other major research directions in computer science and other disciplines. It is hoped that these research questions, findings, and the guideline discussed in this thesis be helpful and constructive for designers and developers ■





# VI

## BIBLIOGRAPHY





# Bibliography

- [1] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles. Towards a better understanding of context and context-awareness. In *Handheld and ubiquitous computing*, pages 304–307. Springer, 1999.
- [2] T. T. Ahonen. *Almanac 2012: The Mobile Telecoms Industry Annual Review for 2012*. TomiAhonen Consulting, 2013.
- [3] E. Almirall and J. Wareham. Living labs and open innovation: roles and applicability. *The Electronic Journal for Virtual Organizations and Networks*, 10(3):21–46, 2008.
- [4] F. Alt, A. Sahami Shirazi, A. Kaiser, K. Pfeuffer, E. Gurkan, A. Schmidt, P. Holleis, and M. Wagner. Exploring ambient visualizations of context information. In *Proceedings of the 2010 International Conference on Pervasive Computing and Communications Workshops* (Mannheim, Germany), PERCOM '10, pages 788–791. IEEE, 2010.
- [5] F. Alt, A. Sahami Shirazi, S. Legien, A. Schmidt, and J. Mennenöh. Creating meaningful melodies from text messages. In *Proceedings of the 2010 International Conference on New Interfaces for Musical Expression* (Sydney, Australia), volume 10 of *NIME'10*, pages 63–68, 2010.
- [6] F. Alt, A. Sahami Shirazi, A. Schmidt, and J. Mennenöh. Increasing the user's attention on the web: Using implicit interaction based on gaze behavior to tailor content. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction* (Copenhagen, Denmark), NordiCHI'12, New York, NY, USA, 2012. ACM.
- [7] A. K. Amin, B. T. A. Kersten, O. A. Kulyk, P. H. Pelgrim, C. M. Wang, and P. Markopoulos. Sensems: a user-centered approach to enrich the messaging experience for teens by non-verbal means. In *Proceedings of the 2005 International*

- Conference on Human Computer Interaction with Mobile Devices & services* (Salzburg, Austria), MobileHCI '05, pages 161–166, New York, NY, USA, 2005. ACM.
- [8] K. Ansari-Asl, G. Chanel, T. Pun, et al. A channel selection method for eeg classification in emotion assessment based on synchronization likelihood. In *In Proceeding of European Signal Processing Conference* (Poznan, Poland), Eusipco 2007, 2007.
- [9] J. Antin. Designing social psychological incentives for online collective action. In *In Proceeding of Conference on Directions and Implications of Advanced Computing* (Berkeley, CA, USA), DIAC, June 2008.
- [10] D. Avrahami and S. E. Hudson. Qna: augmenting an instant messaging client to balance user responsiveness and performance. In *Proceedings of the 2004 International Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA), CSCW '04, pages 515–518, New York, NY, USA, 2004. ACM.
- [11] N. T. Ayas, D. P. White, J. E. Manson, M. J. Stampfer, F. E. Speizer, A. Malhotra, and F. B. Hu. A prospective study of sleep duration and coronary heart disease in women. *Archives of Internal Medicine*, 163(2):205, 2003.
- [12] L. Baillie, P. Frohlich, and R. Schatz. Exploring social tv. In *Proceedings of 2007 International Conference on Information Technology Interfaces* (Cavtat, Croatia), ITI '07, pages 215–220. IEEE, 2007.
- [13] J.-P. Banquet. Spectral analysis of the eeg in meditation. *Electroencephalography and clinical neurophysiology*, 35(2):143–151, 1973.
- [14] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg. Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, 4(1):11–44, 1989.
- [15] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, and G. Bauer. Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden), MobileHCI '11, pages 47–56, New York, NY, USA, 2011. ACM.
- [16] M. Bonnet and D. Arand. Insomnia, metabolic rate and sleep restoration. *Journal of internal medicine*, 254(1):23–31, 2003.
- [17] M. H. Bonnet, D. L. Arand, et al. We are chronically sleep deprived. *SLEEP, NEW YORK*, 18:908–911, 1995.

- [18] M. Borazio and K. Van Laerhoven. Predicting sleeping behaviors in long-term studies with wrist-worn sensor data. In *Proceedings of the Second international conference on Ambient Intelligence* (Amsterdam, The Netherlands), AmI'11, pages 151–156, Berlin, Heidelberg, 2011. Springer-Verlag.
- [19] M. Borazio and K. Van Laerhoven. Combining wearable and environmental sensing into an unobtrusive tool for long-term sleep studies. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (Miami, FL, USA), IHI '12, pages 71–80, New York, NY, USA, 2012. ACM.
- [20] S. Brave, H. Ishii, and A. Dahley. Tangible interfaces for remote collaboration and communication. In *Proceedings of the 1998 International Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA), CSCW '98, pages 169–178, New York, NY, USA, 1998. ACM.
- [21] S. Brewster, D. McGookin, and C. Miller. Olfoto: designing a smell-based interaction. In *Proceedings of the 2006 International Conference on Human Factors in Computing Systems* (Montreal, Quebec, Canada), CHI '06, pages 653–662, New York, NY, USA, 2006. ACM.
- [22] S. A. Brewster, P. C. Wright, and A. D. N. Edwards. An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of the 1993 International Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands), CHI '93, pages 222–227, New York, NY, USA, 1993. ACM.
- [23] J. Brooke. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189:194, 1996.
- [24] J. B. Brown. The use of focus groups in clinical research. *Doing Qualitative Research*, 2:109–124, 1999.
- [25] L. M. Brown and T. Kaaresoja. Feel who's talking: using tactons for mobile phone alerts. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (Montreal, Quebec, Canada), CHI EA '06, pages 604–609, New York, NY, USA, 2006. ACM.
- [26] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster. Robust recognition of reading activity in transit using wearable electrooculography. In *Proceedings of the 6th International Conference on Pervasive Computing* (Sydney, Australia), Pervasive '08, pages 19–37, Berlin, Heidelberg, 2008. Springer-Verlag.
- [27] R. Byrne and B. Findlay. Preference for sms versus telephone calls in initiating romantic relationships. *Australian Journal of Emerging Technologies and Society*, 2(1):48–61, 2004.

- [28] A. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, and R. D. Raizada. Neurophone: brain-mobile phone interface using a wireless eeg headset. In *Proceedings of the 2010 ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds* (New Delhi, India), MobiHeld '10, pages 3–8, New York, NY, USA, 2010. ACM.
- [29] D. Canary, M. Cody, and V. Manusov. *Interpersonal Communication: A Goals Based Approach*. Bedford/St. Martin's, 2008.
- [30] S. K. Card, T. P. Moran, and A. Newell. The model human processor: An engineering model of human performance. In K. R. Boff and J. P. Thomas, editors, *Handbook of Perception and Human Performance*, chapter 45, pages 1–35. John Wiley & Sons, New York, New York, USA, 1986.
- [31] M. Chalmers, D. McMillan, A. Morrison, H. Cramer, M. Rost, and W. Mackay. Ethics, logs and videotape: Ethics in large scale user trials and user generated content. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada), CHI EA '11, pages 2421–2424, New York, NY, USA, 2011. ACM.
- [32] A. Chang, B. Resner, B. Koerner, X. Wang, and H. Ishii. Lumitouch: an emotional communication device. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (Seattle, Washington, USA), CHI EA '01, pages 313–314, New York, NY, USA, 2001. ACM.
- [33] C. Cheshire and J. Antin. The social psychological effects of feedback on the production of internet information pools. *Journal of Computer-Mediated Communication*, 13(3):705–727, 2008.
- [34] W.-j. Cho, S.-w. Lee, J.-h. Won, and S.-o. Choi. Mobile phone having perfume spraying apparatus, Nov. 24 2009. US Patent 7,622,084.
- [35] E. K. Choe, S. Consolvo, N. F. Watson, and J. A. Kientz. Opportunities for computing technologies to support healthy sleep behaviors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada), CHI '11, pages 3053–3062, New York, NY, USA, 2011. ACM.
- [36] Y. Choi, A. D. Cheok, X. Roman, T. A. Nguyen, K. Sugimoto, and V. Halupka. Sound perfume: designing a wearable sound and fragrance media for face-to-face interpersonal interaction. In *Proceedings of the 2011 International Conference on Advances in Computer Entertainment Technology* (Lisbon, Portugal), ACE '11, pages 4:1–4:8, New York, NY, USA, 2011. ACM.

- [37] H. Chung, C.-H. J. Lee, and T. Selker. Lover's cups: drinking interfaces as new communication channels. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (Montreal, Quebec, Canada), CHI EA '06, pages 375–380, New York, NY, USA, 2006. ACM.
- [38] K. Church and M. Cherubini. Evaluating mobile user experience in-the-wild: Prototypes, playgrounds and contextual experience sampling. In *Proceedings of the Workshop on Research in the Large: Using App Stores, Markets and other Wide Distribution Channels in Ubiquitous Computing Research, Copenhagen, Denmark*, pages 29–32, 2010.
- [39] S. Consolvo, K. Everitt, I. Smith, and J. A. Landay. Design requirements for technologies that encourage physical activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montreal, Quebec, Canada), CHI '06, pages 457–466, New York, NY, USA, 2006. ACM.
- [40] S. Consolvo, P. Klasnja, D. W. McDonald, D. Avrahami, J. Froehlich, L. LeGrand, R. Libby, K. Mosher, and J. A. Landay. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 2008 International Conference on Ubiquitous Computing* (Seoul, Korea), UbiComp '08, pages 54–63, New York, NY, USA, 2008. ACM.
- [41] A. Cooper, R. Reimann, and D. Cronin. *About face 3: the essentials of interaction design*. Wiley, 2012.
- [42] D. Cooper and P. Schindler. *Business Research Methods*. McGraw-hill/Irwin. McGraw-Hill Education, 2010.
- [43] T. Coppens, L. Trappeniers, and M. Godon. Amigotv: towards a social tv experience. In *Proceedings of the 2004 European Conference on Interactive Television Enhancing the experience* (Brighton, U.K.), volume 36, 2004.
- [44] C. Cornelius and M. Boos. Enhancing mutual understanding in synchronous computer-mediated communication by training trade-offs in judgmental tasks. *Communication Research*, 30(2):147–177, 2003.
- [45] H. Cramer, M. Rost, and L. E. Holmquist. Services as materials: Using mashups for research. In *Proceedings of the International Workshop on Research in the Large* (Beijing, China), LARGE '11, pages 9–12, New York, NY, USA, 2011. ACM.
- [46] D. Crispell. Tv soloists-statistics on number of television sets owned in households from sri consulting media futures program. *American Demographics*, May 1997.

- [47] K. Crowley, A. Sliney, I. Pitt, and D. Murphy. Evaluating a brain-computer interface to categorise human emotional response. In *Proceedings of the 2010 IEEE International Conference on Advanced Learning Technologies* (Sousse, Tunisia), ICAIT '10, pages 276–278, Washington, DC, USA, 2010. IEEE Computer Society.
- [48] D. Crystal. Language and the internet. *IEEE Transactions on Professional Communication*, 45(2):142–144, 2002.
- [49] C. Darwin and S. California. *The Expression of the Emotions in Man and Animals*. Oxford University Press, USA, 1998.
- [50] J. del R Millan, J. Mouriño, M. Franzé, F. Cincotti, M. Varsta, J. Heikkonen, and F. Babiloni. A local neural classifier for the recognition of eeg patterns associated to mental tasks. *IEEE Transactions on Neural Networks*, 13(3):678–686, 2002.
- [51] D. Derks, A. E. Bos, and J. von Grumbkow. Emoticons in computer-mediated communication: Social motives and social context. *CyberPsychology & Behavior*, 11(1):99–101, February 2008.
- [52] A. K. Dey, G. D. Abowd, and D. Salber. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Hum.-Comput. Interact.*, 16(2):97–166, Dec. 2001.
- [53] N. A. Diakopoulos and D. A. Shamma. Characterizing debate performance via aggregated twitter sentiment. In *Proceedings of the 2010 International Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA), CHI '10, pages 1195–1198, New York, NY, USA, 2010. ACM.
- [54] C. Dodge. The bed: a medium for intimate communication. In *CHI 1997 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, GA, USA), CHI EA '97, pages 371–372, New York, NY, USA, 1997. ACM.
- [55] N. K. Duke and P. D. Pearson. Effective practices for developing reading comprehension. *Journal of education*, 189(1):107, 2008.
- [56] T. Engen and C. Pfaffman. Absolute judgment of odor quality. *Journal of Experimental Psychology*, 59:214–219, 1960.
- [57] M. Esbjörnsson, B. Brown, O. Juhlin, D. Normark, M. Östergren, and E. Laurier. Watching the cars go round and round: designing for active spectating. In *Proceedings of the 2006 International Conference on Human Factors in Computing Systems* (Montreal, Quebec, Canada), CHI '06, pages 1221–1224, New York, NY, USA, 2006. ACM.

- [58] C. A. Everson, B. M. Bergmann, A. Rechtschaffen, et al. Sleep deprivation in the rat: Iii. total sleep deprivation. *Sleep*, 12(1):13–21, 1989.
- [59] R. Faubel, E. Lopez-Garcia, P. Guallar-Castillón, T. Balboa-Castillo, J. L. Gutiérrez-Fisac, J. R. Banegas, and F. Rodríguez-Artalejo. Sleep duration and health-related quality of life among older adults: a population-based cohort in Spain. *Sleep*, 32(8):1059, 2009.
- [60] D. Ferreira, V. Kostakos, and A. K. Dey. Lessons learned from large-scale user studies: Using android market as a source of data. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 4(3):28–43, 2012.
- [61] B. J. Fogg. Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002(December), 2002.
- [62] B. Frei. Paid crowdsourcing: Current state & progress toward mainstream business use. *Produced by Smartsheet. com*, 2009.
- [63] G. Garg, V. Singh, J. Gupta, A. Mittal, and S. Chandra. Computer assisted automatic sleep scoring system using relative wavelet energy based neuro fuzzy model. *WSEAS Transactions on Biology and Biomedicine*, 8, January 2011.
- [64] W. W. Gaver. Synthesizing auditory icons. In *Proceedings of the 1993 International Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands), CHI '93, pages 228–235, New York, NY, USA, 1993. ACM.
- [65] D. Geerts. Comparing voice chat and text chat in a communication tool for interactive television. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles* (Oslo, Norway), NordiCHI '06, pages 461–464, New York, NY, USA, 2006. ACM.
- [66] D. Geerts and D. De Grooff. Supporting the social uses of television: sociability heuristics for social tv. In *Proceedings of the 2009 International Conference on Human Factors in Computing Systems* (Boston, MA, USA), CHI '09, pages 595–604, New York, NY, USA, 2009. ACM.
- [67] A. Girardello and F. Michahelles. Appaware: which mobile applications are hot? In *Proceedings of the International Conference on Human computer interaction with mobile devices and services* (Lisbon, Portugal), MobileHCI '10, pages 431–434, New York, NY, USA, 2010. ACM.
- [68] E. Goodman and M. Misilim. The sensing beds. (Seattle, WA, USA), In *Proceedings of UbiComp Workshop*, 2003.

- [69] E. Goodman and M. Misilim. The sensing beds. In *UbiComp 2003 Workshop*, 2003.
- [70] D. J. Gottlieb, N. M. Punjabi, A. B. Newman, H. E. Resnick, S. Redline, C. M. Baldwin, and F. J. Nieto. Association of sleep time with diabetes mellitus and impaired glucose tolerance. *Archives of Internal Medicine*, 165(8):863, 2005.
- [71] B. Graimann, B. Allison, and G. Pfurtscheller. *Brain-computer Interfaces: Revolutionizing Human-computer Interaction*. The Frontiers collection. Springer Berlin Heidelberg, 2010.
- [72] P. M. Greco, S. D. Hunt, and J. W. Seuck. Communication device having a scent release feature and method thereof, Apr. 3 2007. US Patent 7,200,363.
- [73] R. Grinter and M. Eldridge. y do tngrs luv 2 txt msg? In W. Prinz, M. Jarke, Y. Rogers, K. Schmidt, and V. Wulf, editors, *Proceedings of the 2001 European Conference on Computer-Supported Cooperative Work* (Bonn, Germany), ECSCW '01, pages 219–238, Dordrecht, The Netherlands, 2002. Springer Netherlands.
- [74] E. Haapalainen, S. Kim, J. F. Forlizzi, and A. K. Dey. Psycho-physiological measures for assessing cognitive load. In *Proceedings of the International Conference on Ubiquitous computing* (Copenhagen, Denmark), UbiComp '10, pages 301–310, New York, NY, USA, 2010. ACM.
- [75] E. T. Hall. *The silent language*, volume 3. Doubleday New York, 1959.
- [76] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, Nov. 2009.
- [77] D. Han, C. Zhang, X. Fan, A. Hindle, K. Wong, and E. Stroulia. Understanding android fragmentation with topic analysis of vendor-specific bugs. In *Proceedings of the 2012 19th Working Conference on Reverse Engineering WCRE '12*, pages 83–92, Washington, DC, USA, 2012. IEEE Computer Society.
- [78] J. T. Hancock, C. Landrigan, and C. Silver. Expressing emotion in text-based communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA), CHI '07, pages 929–932, New York, NY, USA, 2007. ACM.
- [79] M. H. Hansen and B. Rubin. Babble online: applying statistics and design to sonify the internet. In *Proceedings of the 2001 International Conference on Auditory Display* (Espoo, Finland), ICAD '01, pages 10–15, 2001.



- [80] G. Harboe, N. Massey, C. Metcalf, D. Wheatley, and G. Romano. The uses of social television. *Comput. Entertain.*, 6(1):8:1–8:15, May 2008.
- [81] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Human mental workload*, 1(3):139–183, 1988.
- [82] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the International Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA), CHI '10, pages 203–212, New York, NY, USA, 2010. ACM.
- [83] F. Hemmert, S. Hamann, and R. Wettach. The digital hourglass. In *Proceedings of the 2009 International Conference on Tangible and Embedded Interaction* (Cambridge, United Kingdom), TEI '09, pages 19–20, New York, NY, USA, 2009. ACM.
- [84] N. Henze and S. Boll. Release your app on sunday eve: Finding the best time to deploy apps. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden), MobileHCI '11, pages 581–586, New York, NY, USA, 2011. ACM.
- [85] N. Henze, M. Pielot, B. Poppinga, T. Schinke, and S. Boll. My app is an experiment: Experience from user studies in mobile app stores. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 3(4):71–91, 2011.
- [86] N. Henze, E. Rukzio, and S. Boll. 100,000,000 taps: Analysis and improvement of touch performance in the large. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden), MobileHCI '11, pages 133–142, New York, NY, USA, 2011. ACM.
- [87] N. Henze, E. Rukzio, and S. Boll. Observational and experimental investigation of typing behaviour using virtual keyboards for mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA), CHI '12, pages 2659–2668, New York, NY, USA, 2012. ACM.
- [88] N. Henze, A. Sahami Shirazi, A. Schmidt, M. Pielot, and F. Michahelles. Empirical research through ubiquitous data collection. *Computer*, 46(6):74–76, 2013.
- [89] B. Hjorth. Eeg analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29(3):306–310, 1970.
- [90] J. Höflich and J. Gebhardt. Changing cultures of written communication: Letter — e-mail — sms. In R. Harper, L. Palen, and A. Taylor, editors, *The Inside Text*,

- volume 4 of *The Kluwer International Series on Computer Supported Cooperative Work*, pages 9–31. Springer Netherlands, Dordrecht, The Netherlands, 2005.
- [91] K. Hogan and R. Stubbs. *Can't Get Through: 8 Barriers to Communication*. Pelican Publishing Company, Incorporated, 2003.
- [92] S. M. Hosni, M. E. Gadallah, S. F. Bahgat, and M. S. AbdelWahab. Classification of eeg signals using different feature extraction techniques for mental-task bci. In *Proceedings of the International Conference on Computer Engineering & Systems* (Cairo, Egypt), ICCES'07, pages 220–226. IEEE, IEEE, 2007.
- [93] E. M. Huang, G. Harboe, J. Tullio, A. Novak, N. Massey, C. J. Metcalf, and G. Romano. Of social television comes home: a field study of communication choices and practices in tv-based text and voice chat. In *Proceedings of the 2009 International Conference on Human Factors in Computing Systems* (Boston, MA, USA), CHI '09, pages 585–594, New York, NY, USA, 2009. ACM.
- [94] E. Isaacs, A. Walendowski, and D. Ranganthan. Hubbub: a sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. In *Proceedings of the 2002 International Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA), CHI '02, pages 179–186, New York, NY, USA, 2002. ACM.
- [95] A. J. Joseph and S. K. Lodha. Musart: Musical audio transfer function real-time toolkit. In *Proceedings of the 2002 International Conference on Auditory Display* (Kyoto, Japan), ICAD '02, 2002.
- [96] S. Kanoh, K. Miyamoto, and T. Yoshinobu. A p300-based bci system for controlling computer cursor movement. In *Proceedingd of the 2011 IEEE Engineering in Medicine and Biology Society* (Boston, MA), EMBC '11, pages 6405–6408. IEEE, 2011.
- [97] L. Kaufman and P. J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley-Interscience, 1990.
- [98] J. J. Kaye. Making scents: aromatic output for hci. *interactions*, 11(1):48–61, Jan. 2004.
- [99] Z. A. Keirn and J. I. Aunon. A new mode of communication between man and his surroundings. *IEEE Transactions on Biomedical Engineering*, 37(12):1209–1214, 1990.
- [100] S. Kim, J. A. Kientz, S. N. Patel, and G. D. Abowd. Are you sleeping?: sharing portrayed sleeping status within a social network. In *Proceedings of the 2008*

- ACM Conference on Computer Supported Cooperative Work* (San Diego, CA, USA), CSCW '08, pages 619–628, New York, NY, USA, 2008. ACM.
- [101] S. King and J. Forlizzi. Slow messaging: intimate communication for couples living at a distance. In *Proceedings of the 2007 International Conference on Designing pleasurable products and interfaces* (Helsinki, Finland), DPPI '07, pages 451–454, New York, NY, USA, 2007. ACM.
- [102] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the International Conference on Human Factors in Computing Systems* (Florence, Italy), CHI '08, pages 453–456, New York, NY, USA, 2008. ACM.
- [103] M. Knapp and J. Hall. *Nonverbal Communication in Human Interaction*. Wadsworth, Cengage Learning, 7 edition, 2010.
- [104] M. Kranz, L. Murmann, and F. Michahelles. Research in the large: Challenges for large-scale mobile application research—a case study about nfc adoption using gamification via an app store. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 5(1):45–61, 2013.
- [105] B. M. Landry, J. S. Pierce, and C. L. Isbell, Jr. Supporting routine decision-making with a next-generation alarm clock. *Personal Ubiquitous Computing*, 8(3-4):154–160, July 2004.
- [106] V. Lavrenko, S. Feng, and R. Manmatha. Statistical models for automatic video annotation and retrieval. In *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Montreal, Quebec, Canada), volume 3 of *ICASSP '04*, pages 1044–1047. IEEE, 2004.
- [107] J. Lazar, J. H. Feng, and H. Hochheiser. *Research Methods in Human-Computer Interaction*. Wiley Publishing, 2010.
- [108] J. C. Lee and D. S. Tan. Using a low-cost electroencephalograph for task classification in hci research. In *Proceedings of the ACM symposium on User interface software and technology* (Montreux, Switzerland), UIST '06, pages 81–90, New York, NY, USA, 2006. ACM.
- [109] L. Leung. Unwillingness-to-communicate and college students' motives in sms mobile messaging. *Telematics and Informatics*, 24(2):115–129, May 2007.
- [110] Y. Li, J. Zhang, Y. Su, W. Chen, Y. Qi, J. Zhang, and X. Zheng. P300 based bci messenger. In *Proceedings of International Conference on Complex Medical Engineering* (Tempe, Arizona, USA), ICME '09, pages 1–5, 2009.

- [111] W.-H. Liao and C.-M. Yang. Video-based activity and movement pattern analysis in overnight sleep studies. In *Proceedings of the 2008 International Conference on Pattern Recognition* (Tampa, FL, USA), ICPR '08, pages 1–4. IEEE, 2008.
- [112] R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 1932.
- [113] Y. Liu, T. Alexandrova, and T. Nakajima. Gamifying intelligent environments. In *Proceedings of the 2011 International ACM Workshop on Ubiquitous Meta User Interfaces* (Scottsdale, Arizona, USA), Ubi-MUI '11, pages 7–12, New York, NY, USA, 2011. ACM.
- [114] Y. Liu, V. Lehdonvirta, M. Kleppe, T. Alexandrova, H. Kimura, and T. Nakajima. A crowdsourcing based mobile image translation and knowledge sharing service. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia* (Limassol, Cyprus), MUM '10, pages 6:1–6:9, New York, NY, USA, 2010. ACM.
- [115] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, B. Arnaldi, et al. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of Neural Engineering*, 4(2):1–13, June 2007.
- [116] K. Luyten, K. Thys, S. Huypens, and K. Coninx. Telebuddies on the move: social stitching to enhance the networked gaming experience. In *Proceedings of the 2006 ACM SIGCOMM workshop on Network and System Support for Games* (Singapore), NetGames '06, New York, NY, USA, 2006. ACM.
- [117] P. Maquet. The role of sleep in learning and memory. *Science*, 294(5544):1048–1052, 2001.
- [118] D. McMillan, A. Morrison, O. Brown, M. Hall, and M. Chalmers. Further into the wild: Running worldwide trials of mobile systems. In *Proceedings of the 2010 Conference on Pervasive Computing*, Pervasive '10, pages 210–227. Springer, Springer Berlin/Heidelberg, Berlin, Heidelberg, 2010.
- [119] D. McMillan, A. Morrison, and M. Chalmers. A comparison of distribution channels for large-scale deployments of ios applications. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 3(4):1–17, 2011.
- [120] C. Metcalf, G. Harboe, J. Tullio, N. Massey, G. Romano, E. M. Huang, and F. Bentley. Examining presence and lightweight messaging in a social television experience. *ACM Trans. Multimedia Comput. Commun. Appl.*, 4(4):27:1–27:16, Nov. 2008.

- [121] A. N. Mhóráin and S. Agamanolis. Aura: an intimate remote awareness system based on sleep patterns. (Portland, OR, USA), In *Proceeding of the CHI Workshop on Awareness Systems*, 2005.
- [122] G. A. Michael, L. Jacquot, J.-L. Millot, and G. Brand. Ambient odors modulate visual attentional capture. *Neuroscience Letters*, 352(3):221 – 225, 2003.
- [123] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In *Proceedings of the 2008 Conference on Embedded network sensor systems* (Raleigh, NC, USA), SenSys '08, pages 337–350, New York, NY, USA, 2008. ACM.
- [124] E. Miluzzo, N. D. Lane, H. Lu, and A. T. Campbell. Research in the app store era: Experiences from the cenceme app deployment on the iphone. (Copenhagen, Denmark), In *Proceedings of the 2010 Workshop on Research in the Large*, 2010.
- [125] H. Miyamori, S. Nakamura, and K. Tanaka. Generation of views of tv content using tv viewers' perspectives expressed in live chats on the web. In *Proceedings of the 2005 International Conference on Multimedia* (Hilton, Singapore), MULTIMEDIA '05, pages 853–861, New York, NY, USA, 2005. ACM.
- [126] A. Möller, S. Diewald, L. Roalter, F. Michahelles, and M. Kranz. Update behavior in app markets and security implications: A case study in google play. In *Proceedings of the 3rd International Workshop on Research in the Large. Held in Conjunction with Mobile HCI, B. Poppinga, Ed.(Sep 2012)*, pages 3–6, 2012.
- [127] A. Morrison, O. Brown, D. McMillan, and M. Chalmers. Informed consent and users' attitudes to logging in large scale trials. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada), CHI EA '11, pages 1501–1506, New York, NY, USA, 2011. ACM.
- [128] A. Morrison, D. McMillan, S. Reeves, S. Sherwood, and M. Chalmers. A hybrid mass participation approach to mobile software trials. In *Proceedings of the SIGCHI Internation Conference on Human Factors in Computing Systems* (Austin, Texas, USA), CHI '12, pages 1311–1320, New York, NY, USA, 2012. ACM.
- [129] A. Morrison, S. Reeves, D. McMillan, and M. Chalmers. Experiences of mass participation in ubicomp research. (Copenhagen, Denmark), In *Proceedings of the 2010 Workshop on Research in the Large*, 2010.
- [130] J. Mostow and J. Beck. When the rubber meets the road: Lessons from the in-school adventures of an automated reading tutor that listens. *Scale-Up in Education*, 2:183–200, 2007.

- [131] J. Mostow, K.-M. Chang, and J. Nelson. Toward exploiting eeg input in a reading tutor. In *Proceedings of the 15th international conference on Artificial intelligence in education* (Auckland, New Zealand), AIED '11, pages 230–237, Berlin, Heidelberg, 2011. Springer-Verlag.
- [132] K. Nagao, S. Ohira, and M. Yoneoka. Annotation-based multimedia summarization and translation. In *Proceedings of the 2002 International Conference on Computational linguistics* (Taipei, Taiwan), volume 1 of *COLING '02*, pages 1–7, Stroudsburg, PA, USA, 2002. Association for Computational Linguistics.
- [133] S. Nakamura, M. Shimizu, and K. Tanaka. Can social annotation support users in evaluating the trustworthiness of video clips? In *Proceedings of the 2008 Workshop on Information Credibility on the Web* (Napa Valley, California, USA), WICOW '08, pages 59–62, New York, NY, USA, 2008. ACM.
- [134] E. Nettamo, M. Nirhamo, and J. Häkkinen. A cross-cultural study of mobile music: retrieval, management and consumption. In *Proceedings of the 2006 Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments* (Sydney, Australia), OZCHI '06, pages 87–94, New York, NY, USA, 2006. ACM.
- [135] NeuroSky. Brain wave signal (eeg) of neurosky,inc. White paper, December 2009.
- [136] Nielsen. *Smartphones Account for Half of all Mobile Phones, Dominate New Phone Purchases in the US*, accessed August 27, 2014. [http://blog.nielsen.com/nielsenwire/online\\_mobile/smartphones-account-for-half-of-all-mobile-phones-dominate-new-phone-purchases-in-the-us/](http://blog.nielsen.com/nielsenwire/online_mobile/smartphones-account-for-half-of-all-mobile-phones-dominate-new-phone-purchases-in-the-us/).
- [137] V.-P. Niitamo, S. Kulkki, M. Eriksson, and K. A. Hribernik. State-of-the-art and good practice in the field of living labs. In *Proceedings of the International Conference on Concurrent Enterprising* (Milan, Italy), ICE '06, pages 26–28, 2006.
- [138] A. Nilsson. Using it to make place in space: Evaluating mobile technology support for sport spectators. *ECIS 2004 Proceedings*, pages 1398–1409, 2004.
- [139] D. A. Norman. *The Design of Everyday Things*. Basic Books, New York, reprint paperback edition, 2002.
- [140] L. Oehlberg, N. Ducheneaut, J. D. Thornton, R. J. Moore, and E. Nickell. Social tv: Designing for distributed, sociable television viewing. In *Proceedings of the 2006 European Interactive TV Conference* (Athens, Greece), volume 2006 of *EuroITV '06*, pages 25–26, 2006.

- [141] E. Oliver. A survey of platforms for mobile networks research. *SIGMOBILE Mob. Comput. Commun. Rev.*, 12(4):56–63, Feb. 2009.
- [142] E. Oliver. The challenges in large-scale smartphone user studies. In *Proceedings of the 2010 International Workshop on Hot Topics in Planet-scale Measurement* (San Francisco, California, USA), HotPlanet '10, pages 5:1–5:5, New York, NY, USA, 2010. ACM.
- [143] M. Othman, A. Wahab, and R. Khosrowabadi. Mfcc for robust emotion detection using eeg. In *International Conference on Communications* (Kuala Lumpur, Malaysia), MICC, pages 98–101. IEEE, 2009.
- [144] A. Oulasvirta, S. Tamminen, V. Roto, and J. Kuorelahti. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci. In *Proceedings of the International Conference on Human Factors in Computing Systems* (Portland, Oregon, USA), CHI '05, pages 919–928, New York, NY, USA, 2005. ACM.
- [145] K. F. Ozenc, J. P. Brommer, B.-k. Jeong, N. Shih, K. Au, and J. Zimmerman. Reverse alarm clock: a research through design example of designing for the self. In *Proceedings of the 2007 conference on Designing pleasurable products and Interfaces* (Helsinki, Finland), DPPI '07, pages 392–406, New York, NY, USA, 2007. ACM.
- [146] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135, Jan. 2008.
- [147] M. B. Perkusich, T. S. Rached, and A. Perkusich. Thinkcontacts: Use your mind to dial your phone. In *Proceedings of the 2011 International Conference on Consumer Electronics* (Las Vegas, NV, USA), ICCE '11, pages 105–106. IEEE, 2011.
- [148] P. Persson. Exms: an animated and avatar-based messaging system for expressive peer communication. In *Proceedings of the 2003 International Conference on Supporting Group Work* (Sanibel Island, Florida, USA), GROUP '03, pages 31–39, New York, NY, USA, 2003. ACM.
- [149] M. K. Petersen, C. Stahlhut, A. Stopczynski, J. E. Larsen, and L. K. Hansen. Smartphones get emotional: mind reading images and reconstructing the neural sources. In *Affective Computing and Intelligent Interaction*, ACII '11, pages 578–587. Springer, 2011.

- [150] L. S. Petrucci, E. Harth, P. Roth, A. Assimacopoulos, and T. Pun. Websound: a generic web sonification tool, and its application to an auditory web browser for blind and visually impaired users. In *Proceedings of the 2000 International Conference on Auditory Display* (Atlanta, Georgia, USA), ICAD '00, 2000.
- [151] M. Pielot, N. Henze, and S. Boll. Experiments in app stores-how to ask users for their consent. In *Proceeding of the CHI Workshop on Ethics, Logs & Videotape*, 2011.
- [152] M. Pielot, B. Poppinga, W. Heuten, and S. Boll. Pocketnavigator: studying tactile navigation systems in-situ. In *Proceedings of the International Conference on Human Factors in Computing Systems* (Austin, Texas, USA), CHI '12, pages 3131–3140, New York, NY, USA, 2012. ACM.
- [153] J. Powell and N. Dibben. Key-mood association: A self perpetuating myth. *Musicae Scientiae*, 9(2):289, 2005.
- [154] J. Preece and B. Shneiderman. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 1(1):13–32, 2009.
- [155] M. Rhoads. Face-to-face and computer-mediated communication: What does theory tell us and what have we learned so far? *Journal of Planning Literature*, 25(2):111–122, 2010.
- [156] C. Robson. *Real world research: A resource for social scientists and practitioner-researchers*, volume 2. Blackwell Oxford, 2002.
- [157] R. Rosenthal and R. Rosnow. *Essentials of Behavioral Research: Methods and Data Analysis*. McGraw-Hill higher education. McGraw-Hill Companies, Incorporated, 2008.
- [158] J. Ross, L. Irani, M. S. Silberman, A. Zaldivar, and B. Tomlinson. Who are the crowdworkers?: shifting demographics in mechanical turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA), CHI EA '10, pages 2863–2872, New York, NY, USA, 2010. ACM.
- [159] N. S. Ryan, J. Pascoe, and D. R. Morse. Enhanced reality fieldwork: the context-aware archaeological assistantthe context-aware archaeological assistant. In *Computer applications in archaeology* (Birmingham, UK), CAA. Tempus Reparatum, 1997.
- [160] A. Sahami Shirazi, F. Alt, A. Schmidt, A.-H. Sarjanoja, L. Hynninen, J. Häkkinä, and P. Holleis. Emotion sharing via self-composed melodies on mobile phones. In



- Proceedings of the 2009 International Conference on Human-Computer Interaction with Mobile Devices and Services* (Bonn, Germany), MobileHCI '09, pages 30:1–30:4, New York, NY, USA, 2009. ACM.
- [161] A. Sahami Shirazi, J. Clawson, Y. Hassanpour, M. J. Tourian, A. Schmidt, E. H. Chi, M. Borazio, and K. Van Laerhoven. Already up? using mobile phones to track & share sleep behavior. *International Journal of Human-Computer Studies (IJHCS)*, 71(9):878 – 888, 2013.
- [162] A. Sahami Shirazi, M. Funk, F. Pfeleiderer, H. Glück, and A. Schmidt. Mediabrain: Annotating videos based on brain-computer interaction. In *Mensch & Computer 2012: interaktiv informiert—allgegenwärtig und allumfassend!?*, pages 263–272. Oldenbourg Wissenschaftsverlag, 2012.
- [163] A. Sahami Shirazi, M. Hassib, N. Henze, K. Kunze, and A. Schmidt. What’s up in your mind? mental task awareness using single electrode brain computer interfaces. In *5th Augmented Human International Conference* (Kobe, Japan), AH'14, page 4. ACM, March 2014.
- [164] A. Sahami Shirazi, N. Henze, T. Dingler, K. Kunze, and A. Schmidt. Upright or sideways?: analysis of smartphone postures in the wild. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services* (Munich, Germany), MobileHCI '13, pages 362–371, New York, NY, USA, 2013. ACM.
- [165] A. Sahami Shirazi, N. Henze, T. Dingler, M. Pielot, D. Weber, and A. Schmidt. Large-scale assessment of mobile notifications. In *Proceedings of the 2014 ACM annual conference on Human factors in computing systems (to appear)* (Toronto, ON, Canada), CHI '14, pages –, New York, NY, USA, 2014. ACM.
- [166] A. Sahami Shirazi, P. Holleis, A. Schmidt, and J. Häkkinä. Rich tactile output on mobile devices. In *Proceedings of the European Conference on Ambient Intelligence* (Nuremberg, Germany), Aml '08, pages 210–221, Berlin, Heidelberg, 2008. Springer-Verlag.
- [167] A. Sahami Shirazi, M. Rohs, R. Schleicher, S. Kratz, A. Müller, and A. Schmidt. Real-time nonverbal opinion sharing through mobile phones during sports events. In *Proceedings of the 2011 International Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada), CHI '11, pages 307–310, New York, NY, USA, 2011. ACM.
- [168] A. Sahami Shirazi, A.-H. Sarjanoja, F. Alt, A. Schmidt, and J. Häkkinä. Understanding the impact of abstracted audio preview of sms. In *Proceedings of the*

- 2010 *International Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA), CHI '10, pages 1735–1738, New York, NY, USA, 2010. ACM.
- [169] A. Salovaara. Appropriation of a mms-based comic creator: from system functionalities to resources for action. In *Proceedings of the 2007 International Conference on Human Factors in Computing Systems* (San Jose, California, USA), CHI '07, pages 1117–1126, New York, NY, USA, 2007. ACM.
- [170] L. A. Samovar, R. E. Porter, E. R. McDaniel, and C. S. Roy. *Communication Between Cultures*. Cengage Learning, 8th edition, 2012.
- [171] M. Saylor. *The Mobile Wave: How Mobile Intelligence Will Change Everything*. Vanguard Press, 2012.
- [172] B. Schilit, N. Adams, and R. Want. Context-aware computing applications. In *Proceedings of the Workshop on Mobile Computing Systems and Applications WMCSA '94*, pages 85–90, Washington, DC, USA, 1994. IEEE Computer Society.
- [173] W. N. Schilit. *A system architecture for context-aware mobile computing*. PhD thesis, Columbia University, 1995.
- [174] R. Schleicher, A. Sahami Shirazi, M. Rohs, S. Kratz, and A. Schmidt. Worldcupin-ion experiences with an android app for real-time opinion sharing during soccer world cup games. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 3(4):18–35, 2011.
- [175] A. Schmidt. Implicit human computer interaction through context. *Personal Technologies*, 4(2-3):191–199, 2000.
- [176] A. Schmidt. Network alarm clock (the 3ad international design competition). *Personal Ubiquitous Computing*, 10(2-3):191–192, Jan. 2006.
- [177] A. Schmidt. *Context-Aware Computing: Context-Awareness, Context-Aware User Interfaces, and Implicit Interaction*. The Interaction Design Foundation, Aarhus, Denmark, 2013.
- [178] A. Schmidt, M. Beigl, and H.-W. Gellersen. There is more to context than location. *Computers & Graphics*, 23(6):893–901, 1999.
- [179] A. Schmidt, A. Sahami Shirazi, and K. van Laerhoven. Are you in bed with technology? *Pervasive Computing, IEEE*, 11(4):4–7, oct.–dec. 2012.
- [180] C. Smith, H. Hancock, J. Blake-Mortimer, and K. Eckert. A randomised comparative trial of yoga and relaxation to reduce stress and anxiety. *Complementary therapies in medicine*, 15(2):77, 2007.

- [181] R. Snow, B. O'Connor, D. Jurafsky, and A. Y. Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Honolulu, Hawaii), EMNLP '08, pages 254–263, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics.
- [182] R. Sobolewski, R. B. Reilly, S. Finnigan, P. Dockree, K. O'Sullivan, and I. H. Robertson. Monitoring of cognitive processes in older persons. In *Proceedings of the International Conference on Neural Engineering* (Antalya, Turkey), NER '09, pages 132–135. IEEE, 2009.
- [183] H. J. Song and K. Beilharz. Aesthetic and auditory enhancements for multi-stream information sonification. In *Proceedings of the 2008 international conference on Digital Interactive Media in Entertainment and Arts* (Athens, Greece), DIMEA '08, pages 224–231, New York, NY, USA, 2008. ACM.
- [184] K. Stanovich and A. Cunningham. What reading does for the mind. *Journal of Direct Instructions*, 1(2):137–149, 2001.
- [185] L. Stewart, K. von Kriegstein, J. D. Warren, and T. D. Griffiths. Music and the brain: disorders of musical listening. *Brain*, 129(10):2533–2553, 2006.
- [186] D. A. Sumikawa. Guidelines for the integration of audio cues into computer user interfaces. Technical Report UCRL-53656, Lawrence Livermore National Lab., CA (USA), 1985.
- [187] K. Sumikawa, M. Blattner, K. Joy, and R. Greenberg. Guidelines for the syntactic design of audio cues in computer interfaces. Technical Report UCRL-93464, Lawrence Livermore National Lab., CA (USA), 1985.
- [188] D. D. S. Y.-P. Tan and S. R. K. P. J. Ramadge. Automated analysis and annotation of basketball video. *Storage and Retrieval for Image and Video Databases V*, page 176, 1997.
- [189] A. Temko, G. Boylan, W. Marnane, and G. Lightbody. Speech recognition features for eeg signal description in detection of neonatal seizures. In *International Conference of Engineering in Medicine and Biology Society* (Buenos Aires, Argentina), EMBC, pages 3281–3284. IEEE, 2010.
- [190] P. T. Terenzini, L. Springer, E. T. Pascarella, and A. Nora. Influences affecting the development of students' critical thinking skills. *Research in higher education*, 36(1):23–39, 1995.

- [191] J. van Erp, F. Lotte, and M. Tangermann. Brain-computer interfaces: Beyond medical applications. *Computer*, 45(4):26–34, Apr. 2012.
- [192] K. Van Laerhoven, M. Borazio, D. Kilian, and B. Schiele. Sustained logging and discrimination of sleep postures with low-level, wrist-worn sensors. In *Proceedings of the 2008 12th IEEE International Symposium on Wearable Computers* (Karlsruhe, Germany), ISWC '08, pages 69–76, Washington, DC, USA, 2008. IEEE Computer Society.
- [193] L. Ventä, M. Isomursu, A. Ahtinen, and S. Ramiah. My phone is a part of my soul - how people bond with their mobile phones. In *Proceedings of the 2008 International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies UBIComm '08*, pages 311–317, Washington, DC, USA, 2008. IEEE Computer Society.
- [194] J. Vidal. Real-time detection of brain events in eeg. *Proceedings of the IEEE*, 65(5):633–641, 1977.
- [195] J. J. Vidal. Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2(1):157–180, 1973. PMID: 4583653.
- [196] A. Voida, R. E. Grinter, N. Ducheneaut, W. K. Edwards, and M. W. Newman. Listening in: practices surrounding itunes music sharing. In *Proceedings of the 2005 International Conference on Human Factors in Computing Systems* (Portland, Oregon, USA), CHI '05, pages 191–200, New York, NY, USA, 2005. ACM.
- [197] U. Wagner, S. Gais, H. Haider, R. Verleger, and J. Born. Sleep inspires insight. *Nature*, 427(6972):352–355, 2004.
- [198] B. N. Walker and J. T. Cothran. Sonification sandbox: A graphical toolkit for auditory graphs. In *Proceedings of the 2003 International Conference on Auditory Display* (Buston, MA, USA), volume 3 of *ICAD '03*, 2003.
- [199] B. N. Walker, J. Lindsay, and J. Godfrey. The audio abacus: Representing a wide range of values with accuracy and precision. In *Proceedings of the 2004 International Conference on Auditory Display (ICAD2004)*, Sydney, Australia (Sydney, Australia), ICAD '04, 2004.
- [200] J. B. Walther and K. P. D'Addario. The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review*, 19(3):324–347, 2001.

- [201] J. B. Walther, T. Loh, and L. Granka. Let me count the ways the interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of language and social psychology*, 24(1):36–65, 2005.
- [202] M. Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, and Y. Song. Unified video annotation via multigraph learning. *IEEE Trans. Cir. and Sys. for Video Technol.*, 19(5):733–746, May 2009.
- [203] P. Ward, N. J. Hodges, A. M. Williams, and J. L. Starkes. 11 deliberate practice and expert performance. *Skill acquisition in sport: Research, theory and practice*, page 231, 2004.
- [204] J. Werner, R. Wettach, and E. Hornecker. United-pulse: feeling your partner’s pulse. In *Proceedings of the 2008 International Conference on Human Computer Interaction with Mobile Devices and Services (Amsterdam, The Netherlands), MobileHCI ’08*, pages 535–538, New York, NY, USA, 2008. ACM.
- [205] J. Wood. *Communication Mosaics: An Introduction to the Field of Communication: An Introduction to the Field of Communication*. Wadsworth, 2010.
- [206] D. Yamamoto, T. Masuda, S. Ohira, and K. Nagao. Video scene annotation based on web social activities. *IEEE MultiMedia*, 15(3):22–32, July 2008.
- [207] Y. Yasui. A brainwave signal measurement and data processing technique for daily life applications. *Journal of physiological anthropology*, 28(3):145–150, 2009.
- [208] Z. Ye, Y. Li, A. Fathi, Y. Han, A. Rozga, G. D. Abowd, and J. M. Rehg. Detecting eye contact using wearable eye-tracking glasses. In *Proceedings of the International Conference on Ubiquitous Computing (Pittsburgh, Pennsylvania), UbiComp ’12*, pages 699–704, New York, NY, USA, 2012. ACM.
- [209] M.-S. Yoh, J. Kwon, and S. Kim. Neurowander: a bci game in the form of interactive fairy tale. In *Proceedings of the 2010 ACM International Conference adjunct papers on Ubiquitous computing - Adjunct (Copenhagen, Denmark), UbiComp ’10 Adjunct*, pages 389–390, New York, NY, USA, 2010. ACM.
- [210] P. C. Zee and F. W. Turek. Sleep and health: everywhere and in both directions. *Archives of Internal Medicine*, 166(16):1686, 2006.
- [211] P. Zerfos, X. Meng, S. H. Wong, V. Samanta, and S. Lu. A study of the short message service of a nationwide cellular network. In *Proceedings of the 2006 International Conference on Internet Measurement (Rio de Janeiro, Brazil), IMC ’06*, pages 263–268, New York, NY, USA, 2006. ACM.

- [212] S. Zhai, P. O. Kristensson, P. Gong, M. Greiner, S. A. Peng, L. M. Liu, and A. Dunnigan. Shapewriter on the iphone: from the laboratory to the real world. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA), CHI EA '09, pages 2667–2670, New York, NY, USA, 2009. ACM.



