From Context to Content: Designing Sensor Support for Reflective Learning

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für Elektrotechnik und Informationstechnik
des Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

M.Sc. Lars Müller
geb. in Crivitz

Hauptreferent:	Prof. Dr. Wilhelm Stork
Korreferentin:	Prof. Dr. Monica Divitini
Tag der mündlichen Prüfung:	21. November 2014

Abstract

Reflection can help professionals to learn more effectively during their on-the-job training. This thesis examines how wearable sensor systems can support reflective learning by monitoring work experiences. A design space is defined that guides designers to build systems that can provide content for reflection by selecting the relevant context, the appropriate capturing method, and visualizations that trigger reflection and can lead to new insights on work practices. Wearable sensors and applications have been developed and evaluated to capture the affective and social context in workplace settings. Healthcare professionals have been selected as target group because reflective practice is an important part of their hands-on training.

An ethnographic study explored the use of wearable ECG sensors in a stroke unit. As a reaction to the dynamic requirements at the workplace, a rapid-prototyping-framework was created to facilitate the development of mobile applications that process psychophysiological data.

In addition, a system based on wearable low-power proximity sensors was developed to capture social contacts in care homes. The system enables a quantitative analysis of care practices and demonstrated in four studies that it supports reflective learning in care homes. Carers reflected on their work practices and gained new insights during the analysis of the sensor data.

This thesis explores a new application domain for wearable sensor systems. It is a first step towards the generation of learning content from sensor data.

Zusammenfassung

Durch die Reflexion von Erfahrungen kann das arbeitsbegleitende Lernen effektiver werden. Diese Arbeit untersucht, wie tragbare Sensoren dieses Lernen unterstützen können, indem sie den Arbeitsalltag erfassen und die resultierenden Daten visualisiert werden. Zu diesem Zweck wurde ein Entwurfsraum definiert, der den Entwicklungsprozess in die Auswahl der Daten, der Erfassungsmethode und der geeigneten Visualisierung strukturiert. Darauf aufbauend wurden tragbare Sensoren und Anwendungen entwickelt und evaluiert, die den affektiven und sozialen Kontext im Arbeitsumfeld erfassen. Das Gesundheitswesen dient als Anwendungsfall, da in diesem Bereich die Reflexion ein elementarer Teil der praktischen Ausbildung ist.

Eine ethnographische Studie evaluierte den Einsatz von tragbaren EKG Systemen zur Erregungserkennung in einer Schlaganfalleinheit. Um agiler auf die Anforderungen in der Praxis reagieren zu können, wurde ein Rapid-Prototyping-Framework für mobile Anwendungen entwickelt, die physiologische Daten verarbeiten.

Darüber hinaus wurde ein System zur Messung der sozialen Kontakte in der Altenpflege entwickelt, das auf tragbaren Näherungssensoren aufbaut. Damit wird eine quantitative Erfassung der Pflege möglich. Vier Studien zeigten, dass das System das reflexive Lernen unterstützt. Die teilnehmenden Pfleger reflektierten über ihren Arbeitsalltag und gewannen neue Einsichten durch die Analyse der aufgezeichneten Daten.

Diese Arbeit ist ein erster Schritt, um Sensordaten als Lerninhalte zu verwenden. Sie eröffnet damit ein neues Anwendungsfeld für tragbare Sensorsysteme.

Acknowledgments

I dedicate this thesis to my parents and siblings who supported and encouraged me during the last four years.

I thank Wilhelm Stork, who encouraged me to work on an interdisciplinary topic and Monica Divitini for being a source of inspiration and learning. Their advice and the discussions have been invaluable to shape this thesis.

I am grateful for the exceptional team that worked with me at FZI, especially Verónica Rivera-Pelayo who contributed on so many levels and infused rigor and an incredible drive to succeed. Andreas Schmidt and Christine Kunzmann provided guidance and encouraged me to pursue my own ideas. My managers, Stephan Heuer and Christophe Kunze offered me the freedom to realize these ideas in the Embedded Systems and Sensors Engineering (ESS) division of FZI. Special thanks to my colleagues Sebastian Chiriac, Nico Adler, Kristina Schaaff, Christoph Zimmermann, and Daniel Orradre, who patiently answered my often confusing questions and motivated me when I was hunting another bug. I would also like to thank all the other members of the Embedded Systems and Sensors Engineering (ESS) division of FZI for their insightful comments and questions. It was a pleasure to work with you and learn from you.

I thank the many students from which this thesis benefited so much, in particular Mark Sonnentag, Pascal Weber, Lukas Platzek, Lukas Jornitz, Stephan Baumeister, André Syndikus, Andreas Übelhör and Daniéle Pevoubou for their important contributions to this work.

Last but not least, I would like to thank all members of the MIRROR project and especially Kevin Pudney, Malcolm Rose, Dominik Walter and Lisa Reinmann. We made an excellent team and I enjoyed working with you. Finally, I would like to acknowledge the European Commission for funding the MIRROR project.

Contents

1 Introduction

Learning is mainly associated with formal education at schools or universities, but learning is a lifelong activity, which is largely informal and work-specific. The studies by Eraut [1] demonstrated that *"most workplace learning occurs on the job rather than off the job."* Hence, work and learning are intertwined; experiences at work are the source of new insights. However, these insights are often not attributed to learning because the gained knowledge is tacit or assumed to be part of a person's existing capability. Furthermore, the learning process is not confined to a specific place. Learning is not facilitated by a teacher and there is no syllabus for instruction. The content that provides new knowledge comes not from a book or other formal learning content but from practical experiences. The learner drives the experiential learning process by creating new experiences and aims to bring them into continuity with her existing knowledge [2].

1.1 Motivation

In hospitals and care homes, on-the-job training and reflective practice are important activities for improving the quality of care and support personal and organizational competence development [3]. Nurses, physicians, and care staff are working in a challenging environment in which quick decisions are often vital. In care homes, reflective practice helps to identify one's own shortcomings and tailor custom solutions for individual patients and residents. Reflecting on work can facilitate learning from experience by drawing additional insights from a situation. Reflective learning refers to *"those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations"* [4]. The individual carefully re-evaluates experiences to come to new outcomes. The outcomes of reflection range from new knowledge to the decision to react differently in future similar situations. While a number of reflective learning theories [5, 4, 6] describe reflection and the underlying mechanisms, the technical support is so far very limited.

From the beginning, diary writing has been a key activity in reflective practice. Recently, handwritten diaries have been replaced by mobile applications [7] and automatic image capturing systems like the Sense-Cam [8] or Google Glass [9]. Mobile diary applications require users to key in data manually and can only record what is already known to the user. These applications act either in the form of a memory aid or help to identify long term trends using visualizations. Nevertheless, many users still prefer writing to typing or flinch at the thought of manually tracking data. Image-capturing systems produce an overwhelming amount of data that is difficult to aggregate. Moreover, these systems cannot be used in many work environments without violating the privacy of third parties that do not agree with their images being used or cannot be asked beforehand.

A growing number of sensors are finding their way into our daily lives to record activity and allow applications to adapt to the current context [10]. Current smartphones include sensors to measure the acceleration, the magnetic field, and many more. Moreover, the microphone [11] and the radio interfaces [12] can be used as sensors as well. In addition, wearable sensors count steps [13] or measure the heart rate [14]. Until now, the usage of these sensors was focused on well-being and fitness by encouraging reflection in our private lives. If sensors could support reflection at work in a similar manner, professionals would be able to learn more effectively on the job.

Ubiquitous computing has created a wide variety of sensors that can be applied in the work process [15, 16]. These sensors often aim at enabling the system to assist according to the user's state rather than empowering the user with additional knowledge. However, sensors can record data at a much more flexible level of detail than diaries and can provide completely new perspectives on one's own work practices. Sensors do not need user interaction and, therefore, can record data in the background without interfering with the primary work processes.

Sensor based approaches have been used in persuasive technology (see Section 3.3.2), but a persuasive system automatically interprets sensor data to guide behavior. The user does not have to understand the data, but only to believe that the systems analysis is correct. Hence, the outcomes are fixed and do not account for the variety of insights that are required at the particular workplace. In addition, the predetermined outcomes may not be relevant to an professional because the relevancy of outcomes for an individual depends on the personality and existing work practices. For instance, problems might be unknown or fuzzy until the right data are inspected. In conclusion, persuasive technology has shown how sensors can change behavior in a predetermined manner. They provide useful tools and methods to design technology that facilitates behavior change but selecting the right sensors to support reflective learning and introducing them into the work process is an open challenge.

1.2 Research Questions

This thesis examines how wearable sensors can support professionals to reflect on their work by providing additional content. In each work domain, professionals face unique requirements and challenges. Reflecting on the right experience can help workers to overcome these challenges. This thesis aims at developing an approach to design wearable, sensor-based reflection support in a manner that can be generalized to other domains. Towards this end, the following three research questions (RQ) are addressed:

1. RQ1: How can we design reflection support using captured data?
2. RQ2: Can sensors capture relevant data for reflection in workplace settings?
3. RQ3: What is the impact of the captured data on reflection and learning in a particular workplace setting?

In general, sensor-based reflection support has to be tailored to the particular workplace to capture relevant data for professionals. Nevertheless, the desired design approach should aim at sensor systems that have the potential to be used in other work environments with minimal changes. Sensor systems should be developed and tested in the workplace to evaluate the design approach, gain insights about the acceptance of the systems in practice, and finally measure the desired impact on learning.

1.3 Research Approach

Reflection and the possible support depend on the particular experience that is reflected upon. Therefore, we followed an user centered design [17] to build sensor based computer supported reflective learning (CSRL) applications. We used an iterative approach based on a strong collaboration with end-users. The development and evaluation of new capturing approaches helped to understand the requirements when designing sensor-based reflection support, as addressed by RQ1. Multiple CSRL applications are required to generalize findings beyond the individual, the particular challenges found in one use case and one technology.

The healthcare domain was selected as the target context because (a) reflective practice is seen as promising in this field [3], and (b) it is one of many non- or little computerized work environments. Healthcare professionals increasingly use computers for their documentation, but they are not looking at a computer screen during most of the time. Two design studies adapted wearable sensors and developed applications according to the needs of a stroke unit (see Chapter 5) and care staff in care homes (see Chapter 6).

The developed systems have been evaluated in the respective target context. Sensor technology and feasibility studies can and must be explored in lab or research environments, but the real impact of a developed system must be studied in the target context. Rogers et al. [18] emphasize that most insights for context-sensitive systems are possible only if they are evaluated in the respective target context. These in situ studies must account for the unpredictable factors that are present only in the field by using broad evaluation approaches. Deliberate broad questionnaires, observations, and semi-structured interviews can capture those effects. In situ studies are prone to result in noisy data because of undesired effects induced by uncontrolled variables. Nevertheless, these results provide the most realistic perspective on the actual impact of a system regarding its original design goal, which is, as stated in RQ3, the support of reflective learning at the workplace.

1.4 Contribution

The core contributions of the thesis are: (a) a design space to structure the conception and implementation of sensor based reflection support, (b) two design studies that build and evaluate CSRL applications to capture affective and social context at work, and (c) insights on the impact of data captured with these applications on the reflective learning process. All studies within this thesis have been conducted during the normal operation in the workplace setting. Therefore, the results provide a realistic impression of the acceptance and impact of the developed application and sensor concepts.

The thesis is inspired by the design methods for persuasive technology [19, 20], which also aims to influence behavior by data capturing. However, the design methods for persuasive technology are of limited use for reflective learning because they assume a predetermined target behavior, whereas reflective learning aims at identifying problem specific alternatives. A designer of sensor supported CSRL applications has to develop a technology that accounts for workplace-specific requirements and provides insights that can be only estimated at design time. The developed design space structures the resulting complexity and provides guidance on options and pitfalls.

Existing approaches that facilitate reflection with data capturing use camera images [8] or self-reporting applications [7], which are not applicable in most workplace-settings. In contrast, the conducted design studies use wearable sensors to facilitate reflective learning. The design studies adapt and evaluate sensor technology from affective computing and ubiquitous computing to capture data on the affective and social context in workplace settings. The design studies resulted in a new wearable sensor system for low-computerized environments that improves learning and reflection at work.

Until now, sensors have been used in technology enhanced learning (TEL) either to select content [21], to adapt content [22], or to train specific skills such as emotion regulation [23]. The impact of sensor data on reflection and learning has not been studied before.

1.5 Structure of the Thesis

Chapter 2 introduces the theory behind reflective learning, provides an introduction into ubiquitous computing, and explains the relevant aspects of affective computing. Building on these fundamentals, Chapter 3 presents the related work in form of existing systems, relevant sensor technologies, and design methods. Figure 1.1 depicts the structure of the following main chapters 4-6 and their relation to the research questions defined in Section 1.2. The first research question is addressed in Chapter 4, which analyzes the requirements, describes the developed design space, and how the design space was used to design the prototypes that are described in the following chapters. The design study in Chapter 5 explores options to capture affective context at work. It presents an ethnographically inspired study with wearable heart rate sensors in a stroke unit and a mobile framework to rapidly prototype such applications. Chapter 6 sheds light on the role of social contacts for reflection. A new proximity sensor is iteratively developed and improved in multiple care home studies. The final studies evaluate the impact on learning. In Chapter 7, we discuss the insights, and the conclusion in Chapter 8 summarizes the results and outlines how future work can built on these results.

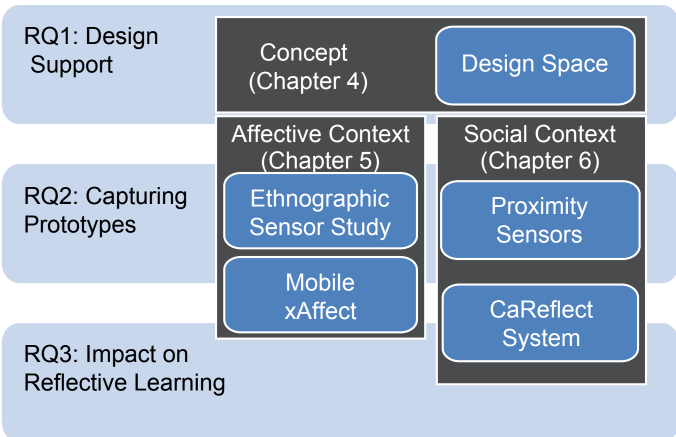


Figure 1.1: Structure of the thesis in relation to research questions

2 Background and Theory

This thesis can be placed within the overlapping context of three fields of research: reflective learning, ubiquitous computing, and affective computing. This chapter provides the background and theory from these fields as related to this thesis. The following sections build a general understanding of the tools and research approaches used in the remainder of the thesis. Reflective learning provides the theoretical underpinning for this thesis. The reflective learning theories that are briefly presented below have developed over several decades. While focusing on the pedagogical aspects the potential role of technology has been mainly ignored. Conversely, ubiquitous computing was driven by the technological vision of Mark Weiser [24]. Affective computing as defined by Picard [25] strives to improve the human computer interaction by including affective aspects. Ubiquitous and affective computing have developed the sensors and algorithms that are the basis of our design studies.

There is a significant overlap between affective and ubiquitous computing in the sense that ubiquitous computing has provided technology that can be used within affective computing. Conversely, there are so far few connections between reflective learning theories and the two technology driven research fields. This thesis aims to close this gap. The theoretical considerations from reflective learning theory can be put to use by building on the technical advances in the domain of ubiquitous computing and affective computing.

2.1 Reflective Learning

When referring to learning in this thesis, I refer to a process that Kolb [5] defines as follows:

> *"Learning is the process whereby knowledge is created through transformation of experience."*

Reflective learning or learning by reflection refers to a set of learning theories that have evolved over several decades resulting in a variety of definitions [2, 5, 4, 6]. Hence, it is difficult to define a shared understanding of reflection. In the following sections, the most important approaches are briefly summarized. A detailed description and discussion of reflective learning theories can be found in [26]. The presented theories focus on the concept of learning from experience and refer to a cyclic or iterative approach to such learning.

2.1.1 Experiential Learning

The origins of reflective learning refer back to the idea of experiential education presented by Dewey [2]. In 1938, Dewey summarized his thoughts on the connection of experience and education. He argued that learning should be tailored to the existing knowledge and experience of a learner by taking the individual experiences into account to build on them. According to Dewey, a continuity of experiences that build on one another leads to the most effective learning. Existing knowledge is confronted with new experiences that in the best case leads to learning. Dewey writes

> *"Continuity and interaction in their active union with each other provide the measure to educative significance and value of the experience."*

Dewey's work was mainly focused on formal learning scenarios and aimed to advise the educator. He argued that teachers should use a less authoritarian style and encourage reflective thinking. His work was a starting point to introduce reflection into formal training and helped to conceptualize experience as an important entity in education.

Kolb [27] was inspired by the work of Dewey and defined experiential learning as a cyclic process, as depicted in Figure 2.1. Therefore, he emphasized the learning process instead of the outcomes. According to the defined four-stage cycle, concrete experiences are turned into observations and reflections that serve as the basis for the following formation of abstract concepts and generalizations. The cycle is closed by testing these implications in new situations which leads to new experiences.

Kolb's model [5] targets not only formal learning, but also refers to organization development as well as training. Drawing from Dewey, Kolb

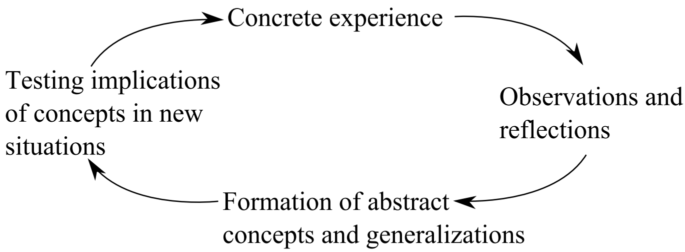


Figure 2.1: The experiential learning model, according to Kolb [27]

defines reflection as a process that spans education, work and personal development. This broader scope made it attractive for a number of domains that deal with knowledge that is difficult to formalize. For instance project management can be only partially learned from books. Hence, business schools and the industry employ business simulation games to create this experience artificially [28]. Similar models are used in case-based learning in hospitals [29].

Professional training programs in a wide range of disciplines strive to teach employees to reflect in action (while doing something) and on action (after doing it) [6]. Daudelin [30] presents one approach of introducing reflective practice. Her work resulted in the following definition of reflection: *"the process of stepping back from an experience to ponder, carefully and persistently, its meaning to the self through the development of inferences; learning is the creation of meaning from past or current events that serves as a guide for future behavior"*.

2.1.2 Reflective Learning and Emotion

Emotions and the motivation to reflect are critical aspects of the reflective process. They can trigger reflection but can also be a barrier to reflection. Experiences might be skewed because they are tightly connected to emotions that prevent an objective analysis. While Kolb[27] said that experiential learning has to integrate *"the cognitive and socio-emotional perspectives on learning"*, these components are not an explicit part of his model.

Boud et al.[4] consider the complete cognitive process including affective aspects. According to their model, reflective learning refers to *“those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations”*. Figure 2.2 shows the reflective process in relation to experiences and outcomes. The reflective process is based on the experiences of the learner, which are *“the total response of a person to a situation, including behavior, ideas and feelings.”* The goal of the three-stage reflective process is to re-evaluate experiences to create outcomes. Outcomes can come in cognitive, affective or behavioral form.

The reflective process itself consists of three steps that have to be repeated for each experience. In the first step, the learner returns to an experience by recalling details about an event or incident. Experiences are often blurred or skewed by the learner’s own emotions. The learner should become aware of emotions without judging them or the experience. In the second step, feelings are evaluated and analyzed. Positive feeling are used to support the reflective process. Articulation of the negative feelings can help to remove them and continue the process. The third and final step, is the re-examination of experiences (i.e., the analysis of the experience in the light of one’s knowledge). These three steps may repeat several times before the learner achieves a clear perspective of an experience and generate outcomes.

The described outcomes are mainly intangible, like the experiences and the reflection process itself. For instance, a new perspective only becomes apparent by articulating it or by changing behavior. There might be outcomes that lack the commitment to action and therefore do not

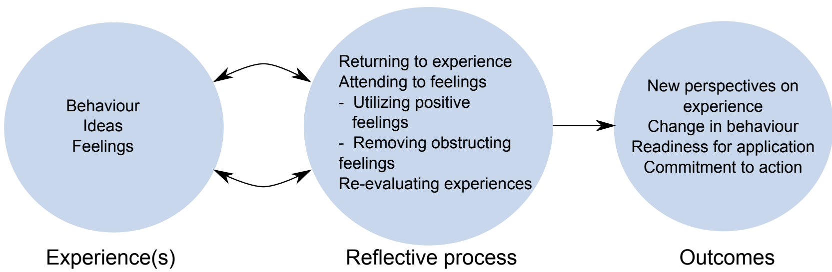


Figure 2.2: Reflection process in context, according to Boud [4]

result in an observable change. However, these changes in the cognitive framework of a learner will influence the behavior in the long term.

The start of the reflection process is a critical point for tool support that initiates the return to experiences. Boud et al. do not explicitly define the beginning of the reflection process because *"most events which precipitate reflection arise out of normal occurrences of one's life."* However, the provided examples can be easily linked to cognitive dissonance theory [31]. Cognitive dissonance theory describes how a mismatch between attitudes and behavior could lead to rethinking of attitudes and experiences. The mismatch is perceived as psychological discomfort (dissonance), motivates a reconsideration of existing attitudes and can trigger the reflective process. The environment can trigger a reflective process by creating an awareness of a discrepancy that leads to a dissonance in the cognitive system of the learner. Examples for such discrepancies are knowledge gaps, unfulfilled expectations or positive surprises such as improvements in productivity or well-being.

2.1.3 Computer Supported Reflective Learning

The previously mentioned theories developed at a time when data capturing meant note-taking with a pen and paper. Reflection was mainly based on cognitive artifacts that had to be reviewed and sorted. Note-taking is possible in only a limited number of work contexts. In the important situations employees are most likely too busy to take notes. For instance a nurse in a hospital has little time to review her tasks at work and even less time to take notes on important events during her shift. Technology can play an important role in capturing event data, structuring that data and sharing it with others. Hence, computer support should become a factor in the theoretical considerations.

The MIRROR project [32] picked up on reflective learning theories to integrate these theories and the current technology. The resulting model for computer supported reflective learning (CSRL) [33] is depicted in Figure 2.3. The CSRL model represents reflective learning as a cyclic four-stage process, similar to the Kolb cycle [27]. The "Plan and Do Work" stage is the source of new experiences and accompanying data. Trigger "Initiate Reflection" and lead to the start of a "Reflection Session," which in turn leads to outcomes that can result in a change of behavior.

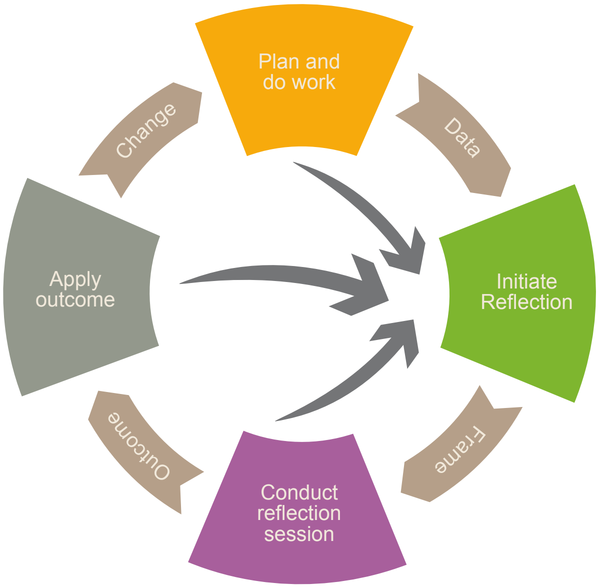


Figure 2.3: Model of computer supported reflective learning (CSRL) by Krogstie et al. [33]

The reflection session can be conducted as an individual, in a group, or on behalf of the organization. This setting, the location, and the more-or-less defined objective form the frame of the reflection session. In the other four stages, new reflection sessions might be triggered that differ in their frame. For instance, a group reflection may trigger an individual reflection by one employee on a related topic. Prilla et al. [34] discuss the transitions between the three levels (individual, collaborative, and organizational) in depth.

The model describes how computers can support reflection at the workplace. The four stages outline the main points in the process, in which technology can facilitate reflection. Data can be captured by technology and become part of the reflection frame. The planning or setup of a reflection session can be supported. The reflection session itself benefits from visualizations and further tools to analyze the experience in depth.

The outcomes of a reflection session can again be stored and turned into tasks to support an actual behavior change. Triggers are the elements that connect different reflection cycles. Data that contrast the learner's perception can act as such a reflection trigger. A more comprehensive overview and discussion of the possible options to support reflective learning by computers can be found in [35].

2.2 Ubiquitous Computing

Ubiquitous computing describes a vision by Marc Weiser [24] of computers that become smaller and embedded in everyday objects. As a result, the computer vanishes from the desk and instead becomes part of it. The user's intentions are inferred directly from the user's actions instead of using the mouse and keyboard, thus enabling users to interact more naturally with computers.

The following sections introduce concepts and research directions in ubiquitous computing that are the technological basis of the sensor systems in our design studies. Context-aware computing involves a plethora of sensor technologies to capture context. Moreover, it defines the technical term context as it is used in this thesis. The proximity sensors in Chapter 6 can be seen as an application of wireless indoor localization technologies. Furthermore, the proximity sensors form a dynamic sensor network that has to optimize media access and duty cycles to minimize power consumption. This is a classic research topic in sensor networks. In addition, the proximity sensors and the ECG sensors used in the first design study are wearable devices. The section 2.2.4 introduces the origins of wearable computing, the role of smartwatches, and the Chronos hardware, which is the basis for the proximity sensors.

Weiser [36] called ubiquitous computing the third wave of computing. In the first wave, big mainframes dominated that were operated from terminals. Computing was centralized and only used by experts. In the second stage, the main frames were replaced by much smaller personal computers. The computational power was no longer confined in computer centers. A much wider range of users could afford and use a computer. Instead of central main frames, computers were standing on a growing number of desks. Ubiquitous computing, as the third wave, follows this trend. Computing capabilities become even smaller and more distributed. The computer itself is distributed in a network of small computational units. These small computational units can cooperate and be combined as needed by the user.

Many of the predictions of ubiquitous computing have come true. For instance, Weiser [24] predicted that there will be three classes of ubiquitous devices, categorized by the size of their display:

- inch scale devices, such as a post it notes
- foot scale devices, such as tablets that mimic the properties of a sheet of paper
- yard scale devices, such as public display that can act as digital blackboards

Today, large public displays are becoming more common. For example, they are found as dynamic banners for advertisements or as information sources in airports and train stations. Tablets such as the Apple iPad are becoming the accessible alternative to PCs. While the variety of mobile phones and smart watches is coming close to the inch scale, new fitness tracking device come in this format.

Size is only one attribute of ubiquitous computing. Devices of different size should be connected and embedded in the environment. The property that is most important for to the user is the interface. The user interface should no longer depend on explicit input but understand a natural interaction and actively support the user without being explicitly asked to do so. This kind of interaction requires that devices are able to sense the current situation by using sensor technology.

2.2.1 Context-Aware Computing

The term context-aware computing was introduced by Ben Schilit [37] in 1994 to describe mobile computing systems that can adapt to their current location, users, and other available devices. When mobile devices move between locations the available resources surrounding them change as well. For instance, a mobile device may use an existing large screen or automatically select a printer based on it's current location. Furthermore, network connectivity itself depends on the available resources. In summary, Schilit's notion of context revolves mainly around location and its implications.

Abowd et al. [38] provide a more comprehensive definition of context:

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

Abowd et al. list four types of context that are especially important: location, identity, activity and time. Consequently, more sensors are required to measure these different types of context. Furthermore, each context requires its own interpretation of the resulting sensor data. The context toolkit provides one approach to manage this complexity, which structures context-aware systems into reusable components [39].

The TECO lab in Karlsruhe developed a variety of sensor technologies to show that context is more than location [40]. The used sensors included: optical/vision, audio, motion, location, bio-sensors, and further, specialized sensors. Their work showed how to combine different sensors and implement the fusion of incoming sensor data. Examples include technologies that have become standard in today's mobile phones and tablets, such as orientation and light-sensitive displays.

The wearIT@work project developed one example for workers in a car factory [15]. The system combines a variety of sensors to recognize work activities. The system recognizes the task currently being conducted and alerts the worker if a critical check was missed.

Context-aware computing has resulted in a wide range of sensors and frameworks for integrating this sensor data in software systems. While the number of sensor technologies is rather stagnating, the implementation of sensors in new application areas is an ongoing challenge. According to Abowd and Mynatt [41] the conceptualization and evaluation of these systems requires not only a sound technological understanding but an in-depth knowledge of the actual usage in real-life scenarios.

2.2.2 Wireless Indoor Localization

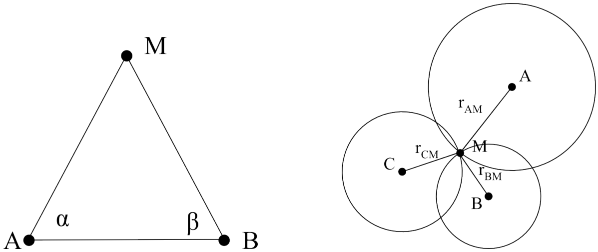
Ubiquitous computing developed a plethora of technologies to determine the location of a device, because location was initially seen as the most important context information. While GPS provides low-cost, precise localization outdoors, the development of indoor localization systems has been an on-going research topic in the ubiquitous computing research community. Although, there have been attempts to use ultrasonic sound [42] or inertial sensors [43], the majority of indoor localization systems rely on radio signals [44].

There are two main methods to estimate location using radio signals: trilateration and triangulation. In both methods, reference locations have to be known and equipped with sensor nodes, known as the reference nodes. Moreover, the object to be located has to be equipped with a sensor, the mobile node. Reference nodes and mobile nodes communicate via radio signals. In trilateration, the reference sensors nodes measure the distance to the mobile node. In triangulation, the reference sensors measure the angle of the incoming signal. Both methods are depicted in Figure 2.4 in a two-dimensional space. Triangulation is used only by few systems such as seen in [45], because measuring angles requires multiple antennas or moving directional antennas for sensor nodes. Hence, most systems use trilateration [46].

The two most used methods to measure distances by radio signals are the received signal strength indicator (RSSI) and variants of the time-of-flight (TOF). This section explains only the underlying fundamentals and their implications for sensor design. Comprehensive summaries of indoor localization technologies can be found in [44] and [46].

The RSSI depends on the distance between the sender and the receiver. The received energy decreases with distance from the sender. The actual decrease depends on the distance from the sender, the frequency of the radio signal the medium, and the properties of both antennas. The following formula by T. Friis [47] approximates the received signal strength P_r by assuming that the radio signal is sent in all directions and that there is no significant attenuation due to the medium.

$$\frac{P_r}{P_t} = G_t G_r \left(\frac{\lambda}{4\pi r} \right)^2$$



(a) Triangulation **(b)** Trilateration

Figure 2.4: Triangulation and trilateration can both be used to calculate the location of a mobile node M. Triangulation requires two points and measures angles. Trilateration requires three points and measures distances.

This formula extends the free-space pass loss formula with antenna properties. A smaller wavelength λ reduces the attenuation, and a sufficient signal strength can be received in a larger distance r from the sender. For a given hardware such as the Chronos eZ430 [48] the frequency and the properties of both antennas G_t and G_r are fixed. In practice, however, antennas do not transmit equally in all directions. Hence, the orientation of the antennas to each other influences this parameter.

The TOF measures the time between the signal is send and received. The method is based on the constant transmission speed of radio waves that can be approximated with the speed of light c. Therefore, a precise measurement of the time the signal is send t_s and received t_r can be used to measure the traveled distance r of the signal:

$$r = c(t_r - t_s)$$

Variations in the received signal strength do not influence the measurement as long as the signal can be received. However, the precise measurement of the time difference is the main challenge in TOF measurement because of the high speed of the transmission. For example, a distance of 50 m results in a time difference of 0.16 µs. Hence, a pico-seconds precise time synchronization between sending and receiving device would be required to measure shorter distances. A common work-around is to instead use round-trip times. The signal is sent to the receiver and immediately sent back. To increase precision, several rounds can be used to effectively multiply the distance the signal has to travel. An implementation of this method has to account for the time to process a packet and send it out again, because this value will multiply as well with each round trip.

In real environments, radio communication deals with reflection and shadowing effects. A human body consists mainly of water and blocks the direct path of a radio signal. The signal may not be received directly, but may be received by reflection from a wall. Furthermore, the same signal can be received from multiple path and multiple reflections. The resulting signal arrives at the receiver several times with varying RSSI values. The electronics used to send the signal may lead to a directionality of the antenna. For instance, if the antenna is placed at the front of the device and a battery is located at the back of the device, the antenna can only send signals forward. This directionality is often intended by design but must be considered when measuring the distance to this electronic device from different reference locations.

Both TOF and RSS measurements are affected and mechanisms are required to deal with these effects. Some system use filters and multiple measurements are used to minimize these effects. Other localization systems do not explicitly calculate distances to reference points, instead using the measured TOF or RSSI values as a finger print for identifying locations. A database of fingerprints for all relevant locations is required to implement this kind of localization. An example of this finger printing is the WiFi-based location mechanism of mobile phones. The location is determined by finding the best matching finger print.

2.2.3 Communication in Sensor Networks

The smallest components in ubiquitous computing are sensor nodes that are distributed in the environment and measure various parameters. A typical sensor node is equipped with a small battery, a microprocessor, a number of sensors and a radio module. Sensor nodes provide the captured data to other ubiquitous devices over the radio module. The resulting network is set up dynamically according to the devices present. The developed protocols are designed for managing the dynamic changes and minimizing the required communication effort to save battery life.

The radio module is the biggest energy consumer on most sensor nodes. Minimizing usage of the radio module is the paramount goal of the protocol design. Communication is limited to few short messages. No lasting connection is established between peers. Ye et al. [49] outlines the main requirements to minimize energy efficient media access in sensor networks. Critical points that must be optimized are: collisions, idle listening, overhearing, and packet overhead.

Protocol implementations will result in a tradeoff between these four points. Collision avoidance is the primary reason to introduce a protocol. If two nodes are sending at the same time on the same channel, communication will be jammed. Resending packets requires more energy and packets can be lost without even noticing. Listening for communication requires an active radio module and requires a comparable amount of energy to sending packets. Listening while nobody is sending wastes energy but is often unavoidable. Similarly, staying awake longer to receive a packet that is addressed for a different receiver can strain the battery. Finally, the sent messages should be as small as possible to minimize the time required for sending the packet. There are three main design ideas to realize low-power media access control (MAC) protocols.

The first approach sends packets with a long preamble. The long preamble signals that the medium will be used to send a packet. Hence, listening nodes must check for packets at an interval that matches the preamble length. As a result, the listening node will always notice that a package is sent without missing any part of the payload. The long preamble reduces the idle listening and saves energy. The main example for this approach is the B-MAC protocol [50]. The drawback of this method is that the preamble is actually a large protocol overhead. The preamble has to be sent with each packet and each receiver will have to receive at least a part of the preamble. Consequently, the sender and the receiver will stay active for a longer time and consume more power than desired.

The alternative and second method are low-power-listening (LPL) protocols. LPL protocols do not use long preambles but repeat the packet or an advertising packet. X-MAC [51] for instance, breaks the preamble down into short advertising packets. Therefore, less data must be received and sent.

Synchronized protocols schedule channels between peers. Each node has receive and send slots that do not overlap with nodes in their proximity. In the best case, peers wake up only when their peers are sending or are ready to receive signals. However, the scheduling requires additional packets, especially if the network structure is changing because of mobile nodes. These changes must be transmitted to all relevant nodes. Hence, a significant part of the communication is spent organizing and reorganizing slots. Furthermore, the clocks on all nodes must be synchronized. This synchronization is necessary in regular intervals because clocks on embedded systems often have a significant clock drift. As a result, receive and send slots may slowly drift until they do not overlap any more.

The time synchronization in distributed networks with a dynamic topology is not a trivial problem. The following formula summarizes the problem by assuming that the clock drift on each sensor node is constant:

$$C_i(t) = a_i t + b_i$$

Therefore, the clock C_i on node i at time t differs from the original time by an offset b_i and a time dependent drift a_i. If the difference between nodes becomes too large, scheduled communication protocols will fail. A variety of time synchronization protocols exists [52] that vary in achievable precision and required overhead. In most protocols, nodes exchange their current time and adjust to each other. This sensor-sensor synchronization can spread the current time through the network but must prevent a constant switching between different times. Another approach is to broadcast synchronization packets that are used by several receivers simultaneously to compare the receive times.

Whichever method is used, the short time to transmit the synchronization message must not be ignored because it is exactly this time difference that TOF measurements have to compare (see Section 2.2.2). The solution is to exchange multiple packets to estimate the difference in clock drift. As a result, the overhead is growing.

In summary, the optimal solution for low-power communication in sensor networks depends on the use case and used hardware. Until now, no general solution has been proposed that is superior across domains and can be applied to highly dynamic network topologies.

2.2.4 Wearable Computing

Ubiquitous computing resulted in smaller devices that can be distributed in the environment to support the user. Steve Mann [53] objected that the technology should empower not the environment, but the user. Hence, technology should be personal and connected to the user in the form of a wearable support system. He described the benefits of wearable computing as follows:

> *"Miniaturization of components has enabled systems that are wearable and nearly invisible, so that individuals can move about and interact freely, supported by their personal information domain."* [54]

One of the first application domains of wearable computing was life-logging by using wearable cameras [54, 8]. Mann used a wearable camera and display to record situations and augment them with digital information [54]. Today, Google Glass [9] is using an approach similar to Mann's, but with a smaller form factor.

Wearable sensors for physiological signals have become available in a wide variety [55]. Electronics can be integrated into textiles to measure motion, temperature, ECG, and respiration. However, the form factors of wearable devices are changing rapidly. Sensors come in new forms to be less intrusive and provide better results than previous devices. Among other options, sensors can be worn as a chest strap [56, 14], on the arm [57], or as a smart wristwatch [58]. A more comprehensive overview of wearable and mobile sensor technologies can be found in reference [10].

The term "smartwatch" refers to wristwatches that are equipped with computer technology to offer functionality beyond time-keeping. The number of models [48, 58] and their capabilities are growing rapidly. Programmable smartwatches allow a user to customize the operation and reuse the form factor for their own applications.

The Chronos eZ430 is a programmable wristwatch offered by Texas Instruments as a development system. The Chronos is sold with a warning that the quality is not up to product standards. The wristwatch uses a CC430 system-on-a-chip (SoC) with an integrated 16-bit MSP430-based low-power microcontroller (MCU) by Texas Instruments and integrates a proprietary low-power sub 1 GHz radio module (CC430). The MCU provides 32 KB flash and 4 KB RAM memory for applications. The hardware platform consists of low-power components and, therefore, a small 3V CR2032 coin cell is enough to power all components. The radio module is configured to transmit with a data rate of 78 kBaud and a radio frequency of the 868 MHz short-range-devices (SRD) band.

2.3 Affect Detection in Affective Computing

The affective state of humans is important information in our daily interactions. If computers were able to sense and react to these affective states the interaction could be more natural and less error prone. Picard [25] defines affective computing as follows:

> *"Affective Computing is computing that relates to, arises from, or deliberately influences emotion or other affective phenomena."*

There is no common understanding in research of what emotions are and how emotions can be measured. Research is based on a variety of models of emotion that have proven to be useful in the past. The circumplex model of affect by Russell [59] is one of the most frequently used models in affective computing because it views emotions by distinguishing only two dimensions. The horizontal axis orders emotions from displeasure to pleasure. Valence has become the common name for this axis. The vertical axis shows the arousal connected to an emotion. Arousal refers to the activation-deactivation level or in other words the energy of a user. Russell refers to the combination of arousal and valance as "core affect". Using the circumplex model all emotions can be understood as a combination of arousal and valence. Therefore, many affective computing systems involve measuring one or both dimensions by analyzing human behavior and physiological states.

In the remainder of the thesis, we will refer to emotions and moods using the broader terms affective aspects or affective context.

Calvo et al. [60] distinguish six different types of signals that are used in affective computing to detect affective aspects. Computers can react to affective phenomena by analyzing:

- facial expressions

- voice (paralinguistic features of speech)
- body language and posture
- physiology
- brain imaging and EEG
- text

Sensor systems and algorithms have been developed and evaluated for each of these signals. In particular, the activity of the heart, muscle activity, and the reaction of the sweat glands indicate the arousal level of a person [61]. Sensors can measure the corresponding physiological signal and algorithms attempt to infer the arousal level. This research direction in affective computing builds on several decades of research in psychophysiology [61]. Today, wearable devices have become available to measure these signals [57, 56, 62, 63]. In the following section, the psychophysiological signals and their analysis will be described, as they have been used in this thesis. A broader overview of current detection and analysis approaches can be found in [60].

The following two sections introduce the background for the first design study in Chapter 5. The ethnographic study used wearable sensors to record the heart rate. The next section explains the general reaction of the activity of the heart to arousal and introduces common methods to analyze this reaction. The prototype in Section 5.3.4 builds on the analysis of the body language. The underlying concepts and features are introduced in Section 2.3.2.

2.3.1 Activity of the Heart

Emotional arousal controls heart rate by influencing the balance between the sympathetic and parasympathetic nervous systems [64, 61]. Both parts of the nervous system control the activity of the heart. Higher activity of the sympathetic nervous system is connected to higher arousal. Conversely, higher activity of the parasympathetic nervous system is connected to a lower arousal. Both systems are always active, but vary in their activation levels over time.

During physical activities, such as sports, the sympathetic nervous system is dominant, resulting in a higher heart rate. The blood moves faster and can transport more oxygen to the muscle cells than during rest. The muscle cells can generate more energy to conduct the required movements. Physical activity is not the only factor influencing heart rate. The human brain attempts to anticipate physical activity and mediates this need by means of the sympathetic and parasympathetic nervous system to the heart. As a result, the heart rate is constantly adapting to the most current situation. Physical effects due to activity and emotional reactions overlap and result in a heart rate that is dependent on the physiological condition of the person.

The heart's activity can be measured by an electrocardiogram (ECG). The electrical signals controlling the heart rate can be measured by electrodes placed on the skin. While Ag/AgCl electrodes with conductive gel are still dominant for medical applications, new dry and noncontact electrodes become available are nonmedical appliances [65]. The gel-based electrodes require careful gel application in the correct amount to guarantee good signal quality. Furthermore, long-term contact between the gel and the skin may cause skin irritations. New dry-electrode devices (e.g. [56]) can be worn as simple chest straps. Alternative methods measure the heart rate by measuring the pulse waves that spread through the body after each heartbeat. Using regular cameras, an MIT team has developed a method to visualize and measure how the pulse wave spreads through the body. Eulerian video magnification [66] is used. The heart rate can also be measured by observing small body movements induced by the pulse waves [67]. Different camera-based methods are used in research and mobile applications to measure the pulse at the finger [68].

The measured heart rate varies because there are changes in the interbeat intervals due to physical activity and physiological, cognitive, and emotional processes. The resulting variability of the heart rate can be analyzed to discern the effects. The following paragraphs explain the most commonly used features to analyze heart rate variability (HRV). The features can be split into statistic, geometric, and frequency-domain methods [69]. The most commonly used methods for automatic analysis calculate statistic features or use features from the frequency-domain.

Hear rate related features are usually analyzed for specified window size. Malik [69] recommends a window size of 5 minutes. An analysis of smaller window sizes can be found in [70]. The analysis in the time domain is based on the time between two normal beats (N = normal beat) of the heart, which is measured in the ECG signal by the NN interval. Typical features in the time domain are:

- SDNN [ms]: Standard deviation of NN intervals in a given interval.
- RMSSD [ms]: The square root of the mean of the sum of the squares of differences between adjacent NN
- pNNx: Count of NN below x divided by the total number of all NN intervals (most common $x = 50$) intervals.

The spectral analysis distinguishes four types of frequency bands shown in Table 2.1. The HF and LF frequency bands have been used for further analysis. The HF power is usually attributed to cardiac parasympathetic nerve activity. The LF power is associated with a dominant sympathetic component [69, 64, 71]. Therefore, the LF/HF ratio is seen as a measure to identify the currently dominant nerve [72, 73]. However, this clear assignment of frequency bands to the activation of the nervous system and the value of the LF/HF ratio for affect recognition are still debated [74].

The methods presented above are affected by physical activity. If users are moving, changes in heart activity are a combination of heart rate increases, which are due to arousal, and changes that result from a change of the physical activity. Therefore, HRV features cannot be interpreted during physical activity. Small changes that are due to movement can be mistaken for arousal reactions.

Myrtek [75] has presented an approach to overcome this challenge. The additional heart rate algorithm aims at isolating affect-related heart rate changes from changes that are induced by physical activity. The algorithm has been deduced from empirical data and uses only the heart rate value for each minute HR_i and acceleration data. In an arousal reaction, the

Abbreviation	**Name**	**Range in Hz**
ULF	Power in the ultra low frequency range	<0.003
VLF	Power in the very low frequency range	0.003 - 0.04
LF	Power in the low frequency range	0.04 - 0.15
HF	Power in the high frequency range	0.15 - 0.4

Table 2.1: HRV frequency bands according to [69]

heart rate rises:

$$HR_i > HR_{i-1}$$

In the algorithm, the heart rate is compared to a sliding mean average $\overline{HR}_i$ of the last 3 minutes.

$$\overline{HR}_i = \frac{1}{3} \sum_{i=1}^{i<4} HR_i$$

The data from the acceleration sensors have to be transformed into the expected increase in heart rate. The acceleration values are filtered and added and transformed into a value ACT_i by using a logarithmic function. The resulting values of ACT_i range between 0 and 200. If these activity values are too high in comparison to a 3-minute average, the calculation is stopped for this minute because the algorithm cannot analyze such sudden changes. The expected activity-related heart rate increase ΔHR_i is calculated by using the variable $CDIV_i$.

$$\Delta HR_i = \frac{90 + ACT_i}{CDIV_i}$$

$CDIV_i$ is adjusted according to the number of arousal reactions that have been found in the last 20 minutes. If less than five reactions have been found, CDIV will be decremented. If more than 10 reactions have been found, CDIV will be incremented. CDIV can vary between 0 and 30 and $CDIV_0 = 23$.

The increase in heart rate has to be higher than the increase that can be attributed to movement. The relation between the actual change in heart rate and the expected change due to physical activity is the additional heart rate ADH at minute i.

$$ADH_i = \frac{HR_i - \overline{HR}}{\Delta HR_i}$$

This algorithm is limited to small changes in physical activity. Explicitly excluded are posture changes, such as sitting down or getting up from a chair. The algorithms adapts to the current type of activity by changing the CDIV value. After rapid changes, the algorithm will need time to adjust to the new type of activity.

Kusserow [76] aimed at overcoming this challenge in his thesis and developed a number of prototypes for specific use cases. The selected use cases (ski jumping and cello concert) are well defined in their physical activity. For example, ski jumpers perform the same movement again and again. A cello player repeats similar movements of the arm during a concert. In these cases, multiple instances of an activity can be compared at different stress levels. The results confirm that *"stress-arousal"* influences the heart rate beyond the physical activity. The application to free-living daily activities showed only an overlap with reported arousal events of 7.8 percent. This result is attributed to the low salience of smaller arousal events.

2.3.2 Posture and Body Language

Already Charles Darwin researched the relation between body language and emotions in humans and animals [77]. Emotional body language (EBL) is a promising venue for further research in affective computing because *"When we see a bodily expression of emotion, we immediately know what specific action is associated with a particular emotion"* [78]. Further research has highlighted the role of EBL in social interaction. Pentland showed how we mimic the body language of each other in communication [79] and how sensors can measure this behavior.

Posture and body language can be measured by either cameras and optical motion-tracking systems [80, 81] or sensors that are worn [82] or embedded in furniture [83, 84]. The sensors used are either accelerometers or pressure mats. Posture while working at a desk has received special attention, because this is a person's typical interaction with a computer. Sensor-augmented chairs [83, 84] provide an unobtrusive means to capture this data. Attention, engagement, and boredom are typical states that can be inferred from body posture [83]. Furthermore, recent research strives to infer information about stress level from posture data [84].

The analysis methods can be split into two broad categories. The first type employs a two-step process. In the first step postures are recognized. In the second step, these postures are translated into affective phenomena. For instance, leaning back could be interpreted as a sign of lower engagement with the task at hand. The second category strives to infer affective aspects directly from the sensor data.

Typical features for the analysis of posture and body language are kinetic energy, consistency, and orientation of sensor or body part. The kinetic energy measured by an acceleration sensor is calculated by summing up all acceleration components and subtracting acceleration due to gravity g:

$$E = \sqrt{a_x^2 + a_y^2 + a_z^2 - g^2}$$

The consistency C quantifies the variability of the energy E over time. It is defined as $1 - \alpha$ where α is the standard deviation of the kinetic energy over time:

$$C = 1 - \sqrt{\frac{1}{N}\sum_{t=0}^{N}(E_t - \mu)^2}$$

These and further features that are specific to the used sensor are used in combination with machine learning methods to link them to emotions [85, 84] and analyze human behavior [86]. A more comprehensive survey of affect-related body language and their recognition can be found in [87].

3 Related Work

This chapter provides an overview of the available reflection support, relevant sensor technologies, and approaches to design such systems. Learning can and is already supported by technology. Technology can help to create, organize, and present learning content. Furthermore, technology can facilitate collaborative learning despite physical distance. Goodyear et al. [88] define technology enhanced learning (TEL) as follows:

> *"Many different types of technology can be used to support and enhance learning. 'Technology' in its broadest sense can include both hardware – such as interactive whiteboards, smart tables, handheld technologies, tangible objects – and software – e.g. computer-supported collaborative learning systems, learning management systems, simulation modeling tools, online repositories of learning content and scientific data, educational games, web 2.0 social applications, 3D virtual reality, etc."*

The following three research fields have developed first approaches to support reflection with technology or technologies that could be applied to this challenge. Sensors increasingly play a role in TEL to adapt and select context. However, there are many more sensor technologies that can be applied to support reflection. Finally, existing research on design methods and models completes the related work.

3.1 Reflection Support

The idea to reflect on captured data is not new. Technologies to support reflection have been developed in TEL, ubiquitous computing, and human computer interaction. The approaches vary according to the key aspect of the corresponding research field. Sensors play a growing role in TEL. However, the main focus in TEL is still formal learning in the classroom.

The MIRROR project is one of the few projects that target informal learning at the workplace. Ubiquitous computing has mainly focused on data acquisition by means of sensors and the management of the data to realize life-logging. Life-logging involves broad capturing of daily activities to create multi-media diaries and memory support. Both are topics that are closely related to reflection support. Recently, reflection has regained popularity as a topic in human computer interaction, but this work mainly addresses how reflection can be facilitated by user interface design and prompting mechanisms.

3.1.1 The MIRROR Project

The work presented in this thesis was conducted as part of the MIRROR project [32], which involved creating a set of applications that facilitate reflection of employees on their work practices. The work on CaReflect (see Chapter 6) and psychophysiological sensing (see Chapter 5) represents only a small subset of the developed applications, which span a wide range of approaches to support reflection and have been developed according to the CSRL model (see Section 2.1.3). This section briefly describes the alternative approaches that have been taken in the MIRROR project.

The work that most closely relates to this thesis is the development of WATCHiT [89, 90] but addresses a different domain. WATCHiT is a modular wearable sensor system for volunteers in crisis preparation. Volunteers must be prepared to react in the event of a disaster, such as an earthquake. In these scenarios, volunteers must cope with challenging, dynamic situations. Therefore, volunteers gain experience during large-scale simulations of such events. Efforts to conduct such an event are large. Therefore, volunteers must learn as much as possible during these events. WATCHiT facilitates reflection by capturing data on tasks. It combines sensors and simple gesture interaction. Figure 3.1 shows the components of the WATCHiT system. Sensor modules can measure heart rate, temperature, noise, and location. The sensor modules can be combined as needed. The gesture interaction is implemented based on near-field communication (NFC). Volunteers can move NFC readers at their wrists to touch NFC tokens at their arms (shown in Figure 3.1-c). The work on WATCHiT complements the work described in this thesis as described in [198]. The modular architecture is beneficial for the dynamic

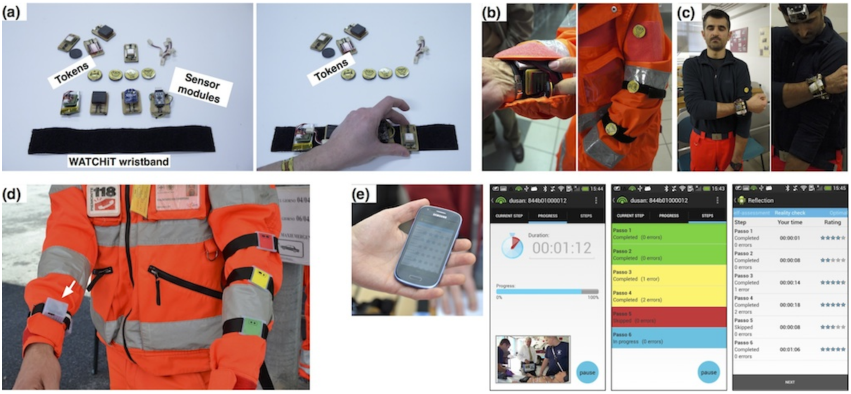


Figure 3.1: WATCHiT prototype: (a) sensors modules and components, (b) WATCHiT sensors worn at the wrist, (c) gesture interaction, (d) sensors and NFC tags worn by volunteer, and (e) mobile application for data analysis

requirements in crisis preparedness, but is not a requirement in care homes and hospitals.

MIRROR developed a wide variety of mobile applications that enable employees to collect data manually [91, 92, 93]. However, each app emphasized a different aspect of reflection. The Live Interest Meter (LIM) app [91] collects data from observers. It is designed for lectures and other public speaking engagements. Listeners can directly provide feedback during the presentation with their mobile phones. The TalkReflection [93] app allows employees to take notes on difficult talks and supports collaborative reflection with sharing and annotation features. The Carer app [92] integrates creativity techniques to help care staff find new alternative reactions to challenging behavior of residents. Care staff can speak or type in case descriptions, and the system finds similar cases and related solutions from different domains. For instance, if a carer is confronted with an aggressive resident, the system may offer solutions, such as a description of how a teacher reacted to an aggressive student.

3.1.2 Life-Logging

Life-logging is one of the initial applications discussed in ubiquitous computing [94, 95]. MyLifeBits [94] is probably the most well-known research project in life-logging. Gordon Bell digitized and recorded all parts of his daily life for several year, including articles, books, cards, CDs, letters, memos, papers, photos, presentations, home movies, videotaped lectures, and voice recordings. A number of tools helped to capture and review these data.

One of the tools used was the SenseCam [8], a wearable camera that automatically takes pictures. A microphone and a passive infrared sensor are used to register changes in audio or light level. In this case, a picture is automatically taken and stored on the SD card. If no event is registered within 30 seconds, a picture will be taken automatically. Hence, the minimal time resolution for a SenseCam picture is 30 seconds. The resulting number of images for a single hour is at least 120. A specialized image browser allows a quick review of a large amount of pictures.

Fleck et al. [96] evaluated the SenseCam in the classroom. Tutors wore the SenseCam during class and reviewed the lessons afterward with their mentors. The triggered discussions indicate that the captured pictures facilitate reflection. The pictures support the returning to experience as described by Boud [4] and triggered discussions of related thoughts and activities.

SenseCam and additional sensors as used in MyLifeBits produce large amounts of data. Blum et al. [95] showed how the data can be filtered. The developed system used audio data and location information to recognize time spans of interest. For instance, if laughter was detected, the related time span would have been marked as “very interesting”. It was assumed that as much data as possible is captured in a first step, and filtered in the second step according to yet unknown requirements.

After reviewing the state of the art, Sellen and Whittaker conclude that life-logging should target specific goals, including reflection and reminiscence [97]. They identified four main challenges for future life-logging applications to be applicable for these specific goals:

1. Selectivity, not total capture
2. Cues, not capture

3. Memory refers to a complex, multi-faceted set of concepts
4. Synergy, not substitution

Hence, the selection of the relevant data that can act as cues is a pending challenge for research. Reflection support has to capture cues and adapt to the specific challenge. Reflection requires a synergistic integration of one's own perception and captured digital artifacts.

3.1.3 Encouraging Reflection

Reflection requires time and effort to review past experiences. The results are often beneficial and can lead to learning. However, especially at work, time pressures and other challenging tasks limit the time that can be used for self-reflection. Human computer interaction research has used three basic methods to encourage reflection: making data on past events available, explicitly prompting the user to reflect, and social facilitation.

Data on past events are the basis of nearly all approaches that have been explored. For instance, most MIRROR apps (see Section 3.1.1) involve capturing data. Visualizations and aggregation can add value to the collected data. Li et al. [98] conducted a survey of available self-tracking tools and their users. Reflection is one of the activities related to tracking tools. Li et al. distinguish short- and long-term reflection. Short-term reflection takes often place directly after or even during recording. Users check their current state and gain awareness. Long-term reflection helps one gain deeper insights. Here, aggregation and visualizations are crucial to provide new perspectives.

Prompting builds on such existing data and actively encourages reflection. Applications that use prompting either simply remind a user or point to relevant data for reflection. Pensieve [99] evaluated prompts with varying content. The results showed: *"Shorter, more general triggers draw more responses, as do triggers containing people's own photos – although responses to photos tended to contain more meta data elements than storytelling elements."* The Echo system [7] uses a self-tracking approach by reassembling an easy-to-use mobile journal. Users can take pictures or videos and annotate them with comments and ratings of their emotional state. Reflection is explicitly facilitated by prompting. Users are explicitly asked to reflect on existing notes. The corresponding reflection screen invites them to review their emotional state and provide the current perspective on the recorded event. Similarly, they can add additional text and pictures. The same event can be reviewed several times, thus resulting in content with multiple reflections. Hence, users can track how their perspectives changed over time. In general, the reflection had a positive effect on the participants' perceived well-being. For instance, there was a tendency to rate experiences with less emotional ratings when reviewing them.

Social facilitation has been one of the main methods to encourage reflection [100], even without technology support. The exchange of ideas and perspectives and the necessity to articulate one's own perceptions in conversations can facilitate reflection. Technology can help to form groups without requiring co-location. Furthermore, technology can capture the current state of a reflection session in a digital manner and allow participants to reflect together, but in an asynchronous manner [101, 93]. For example, care staff that work on different shifts can reflect on incidents using an online forum or a mobile tool. Therefore, the organizational effort to set up meetings is reduced. In MIRROR, Prilla et al. [93] researched technology support for collaborative reflection in public administration, care homes, and a hospital. The insights confirm the potential of this approach in real work settings.

3.2 Measuring Behavior

Besides technology that explicitly targets reflection, research in ubiquitous computing has resulted in a rich variety of sensors that could be applied to reflective learning. In the following, two sensor technologies will be described that are relevant for this thesis. xAffect has been used to process psychophysiological signals in the first design study. Proximity sensors are the basis of the second design study.

3.2.1 Psychophysiological Sensing with xAffect

Emotions play a decisive role in reflective learning, as outlined in Section 2.1.2. One of the first targets of affective computing (see Section 2.3) has been the usage in TEL [22, 102]. This section will relate this thesis to the prior work by Schaaff [102]. Schaaff evaluated the usage of psychophysiological signals such as heart rate and electrodermal activity, to support learning applications. The developed algorithms and tools were available as a basis for this thesis.

The work of Schaaff built on two learning applications. Various studies were conducted to measure arousal and adjust the learning process accordingly. The first application [23] targets financial decision-makers who should learn to regulate their emotions to come to better decisions. A trading game was evaluated that reacts to the arousal of a player by increasing the difficulty of the game with higher arousal levels. The arousal level was inferred by analyzing the heart rate (see Section 2.3.1). Hence, the goal of using the arousal values was to adapt the system. The difficulty of the game increased with the arousal level of the player. Furthermore, feedback on the arousal level was provided to allow the player to react. The second application [103] supports the training of the working memory. Arousal is again deduced by using the heart rate. In this case, however, the design aims at adapting the difficulty of the game to the player. In summary, a number of algorithms has been developed and evaluated to adapt system behavior according to detected arousal levels.

The underlying architecture, xAffect, has been used and developed in collaboration with the author [206]. Schaaff focused on the usage of xAffect for the rapid prototyping of algorithms and their integration in TEL applications. This thesis reused this infrastructure to evaluate the algorithms in a mobile setting and extended it accordingly (see Section 5.3).

The high-level architecture of xAffect is depicted in Figure 3.2. The system is built in a modular way to allow a simple configuration of the data flow. The functionality is split into components that differ according to their roles in the system. *DataSources* are the entry point of data into the system. These could be sensors or network interfaces that receive data from outside of the system. *DataProcessors* are components that modify data. Filters and algorithms are typical examples of *DataProcessors*. *DataSinks* are components that receive data out of the system. In the simplest case, this can be a logfile, a database, or a link to another application. The data flow between the components is managed by the *DataDispatcher*. Each component comes with a description of its capabilities. The *DataDispatcher* connects the components according to this description. The setup defines which components are available and the *LifeCycleManager* loads and stops

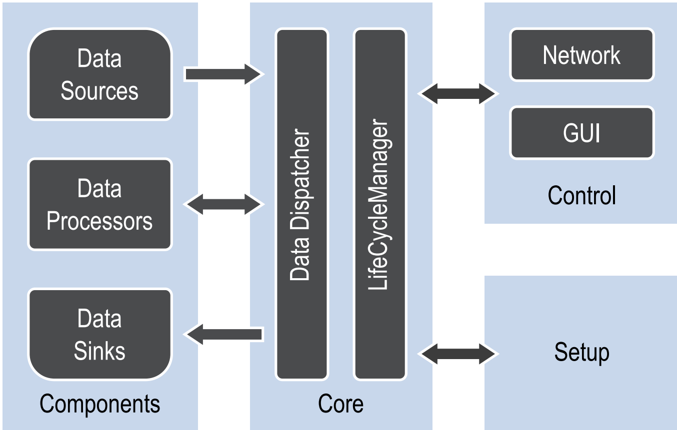


Figure 3.2: xAffect architecture, according to [102]

components. Each component implements a common life cycle, a specified sequence of functional steps starting with the components' creation and ending with its destruction. For instance, one part of the sequence is the training phase during which components already fill their buffers to be ready to process data. The *LifeCycleManager* coordinates the life cycle of all components in a central manner to avoid inconsistencies (e.g., if one component starts to process data while the component that should deliver the data is not yet ready to do so). The overall state of the system can be controlled by a graphical user interface or over the network.

A number of standard components are part of xAffect. A DataViewer visualizes incoming data live on the desktop. The UnisensReader is a source that can read from an Unisens file. The corresponding sink, the UnisensWriter, can write to an Unisens file. Further components include a signal generator to test configurations and a sliding mean filter. Moreover, xAffect comes with a graphical user interface to manage studies and configure the pre-defined combinations of components.

In summary, a number of algorithms for the analysis of heart rate and electrodermal activity (EDA) data have been developed and tested in various studies. In these studies, the psychophysiological context was used to adapt the interactive presentation of existing content. xAffect was used to connect algorithms and games. Furthermore, it is important to note that while the work of Schaaff [102] targeted real work contexts, all conducted studies took place in controlled lab environments according to carefully crafted study designs. Arousal was induced in a controlled manner to compare measurements to the expected state. While the work of Schaaff is the basis for the psychophysiological analysis in this thesis, the focus was on the data analysis and not on the design process.

3.2.2 Proximity Measurement

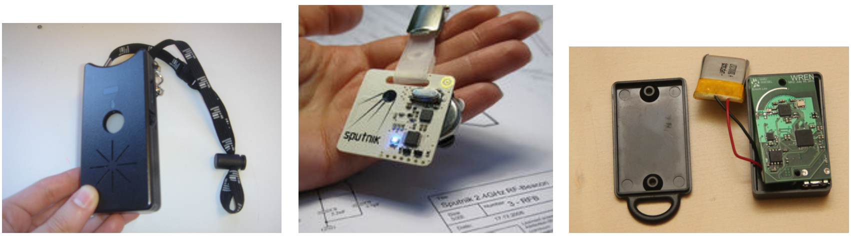
Recognizing co-location between mobile nodes is an application of localization technologies (see Section 2.2.2) to mobile nodes in a sensor network (see Section 2.2.3). Consequently, co-location was initially seen and measured as a by-product of localization [104, 105]. The NearMe Wireless proximity server [104] computed proximity between mobile users by analyzing the exact location of each mobile node. However, as precise indoor location is still a challenge, a variety of technologies were developed to measure proximity directly without knowledge of the location of both nodes [16, 106, 107, 108, 109, 12]. Because they do not need reference nodes, these systems can operate without the localization infrastructure. Therefore, they can operate independently of the environment and location. The developed systems can be split into systems that use dedicated hardware [16, 106, 107] and those that reuse the interfaces of the mobile phone [108, 109, 12].

Systems that directly measure proximity calculate the distance by using the TOF or the RSSI and, based on these values, decide if a node is within proximity. The decision is often based on a threshold value that defines the proximity range. However, RSSI values vary because of reflections and interference of multiple or reflected signals. The multipath propagation of the signal can also distort TOF measurements. Repeated measures can minimize these problems, but will increase the required communication overhead. In summary, a higher precision requires more power and reduces the run time of a sensor. Hence, the determined proximity is only a probability that depends on the trade-off between the precision and power consumption that guided the system design.

The mobile phone as a wearable device provides the required capabilities to act as a proximity sensor and has been used as such [108, 110, 12]. The mobile phone provides a variety of radio interfaces that can be used for TOF or RSSI measurements. Today, nearly everyone carries a smartphone on their bodies. The Virtual Compass system [108] uses WiFi and Bluetooth signals to estimate the distance between mobile phones and laptops. The Bluetooth interface was also used by Eagle and Pentland [109] to measure contacts between students. They built a dataset to analyze student behavior by their contacts. Matic et al. [12] used the WiFi radio to estimate proximity. The benefits of the mobile phone approach are that they are available in large quantities and that they come with rather large batteries. The main problem is that the solutions depend on the actual phone. For example, the used radio chip and the orientation and type of the antenna influence the measurements. Hence, two phones with WiFi may measure different values and the proximity estimation must be device dependent [111]. Furthermore, the implementations are not energy efficient because, the hardware is not designed for proximity measurement and is also used for other purposes. For instance, Matic et al. [12] periodically switch the phone's WiFi module between client and access point mode.

Dedicated devices can deliver more precise results and can be developed with a smaller form factor. An overview of available devices for proximity measurement is shown in Figure 3.3.

The Sociometric badge [115] combines a number of sensors to measure social contact. A Bluetooth module and a 2.4 GHz receiver are combined



(a) Sociometric Badge [112]

(b) OpenBeacon [113]

(c) Wren Mote [114]

Figure 3.3: Available proximity sensors in chronological order of development from left to right

to measure proximity. In addition, an infrared sensor recognizes when two persons face each other and voice detection recognizes conversations. The Sociometric badge was successfully used in a variety of domains [112, 115]. The badge is worn around the neck so that the direction of the infrared sensor conforms to the line of sight. The system captures rich data on social interaction, but the additional sensors (e.g. the microphone) may compromise the privacy of third parties. Furthermore, the system's size is similar to that of a mobile phone.

The OpenBeacon system [113] is an open source initiative to create low-cost proximity sensors for social network research. The system is based on active radio-frequency identification (RFID) and uses a single coin cell to operate over several months. If proximity is recognized, the event is directly transmitted to an access point. Hence, OpenBeacon differs from other dedicated proximity sensors because it requires an infrastructure. The system was used in various settings, including a large conference with more than 500 users [106]. In the conference venue, several access points had to be deployed to measure contacts. Moreover, the access points allow a rough location estimation. While the OpenBeacon system allows large-scale deployments at low costs, infrastructure has to be deployed and managed. OpenBeacon combines mobile devices and stationary access points in a similar fashion as classic localization systems. The access points are powered by cable, whereas mobile devices have to rely on their batteries. Hence, the architecture aims to reduce the load for wearable devices by implementing functionality that consumes more power into the access point. Power-relevant functionality involves, for instance data storage and management of media access (e.g. by synchronizing communication into time slots).

The WREN mote [114] was designed for rapid deployments in large-scale studies. The system uses an 802.15.4-compliant radio to measure proximity and has a rechargeable lithium-polymer battery. The rapid deployment is supported by management racks that can charge and download data from 100 motes at the same time. The system was used with school-age children to research how diseases spread through the social networks [107]. The system should be able to operate for five days at a sampling rate of 20 Hz.

The application domains of proximity sensing are manifold. Envisioned applications include dating services [104], community applications [116], and location-based services [105]. The majority of these sensor systems target human contact research to analyze businesses [112, 115], social contacts on conferences [106], the impact of social interaction on psychology [12], and epidemiology research [117, 107]. In all of these cases, the correctness of a single contact is not decisive because a huge amount of data is analyzed and minor errors do not have an observable impact on the results. The data are anonymized and abstracted from the specific situation. Social network analysis (SNA) algorithms [118] have been used to analyze the social network automatically and deduce the underlying patterns. This research, until now, underestimated the value of the detailed data to measure and understand work processes.

3.3 Designing Reflection Support

The design of CSRL systems has to integrate technical, cognitive, and social processes. Three different research fields have elaborated on providing assistance in the design process. The following chapters provide an overview of the main research direction relevant to the design of CSRL applications.

3.3.1 Design Landscape

Fleck and Fitzpatrick [119] frame a design landscape for reflection support systems. Fleck defines five levels of reflection and suggests how tools can support each of these levels. The work is strongly influenced by the prior work of Fleck using the SenseCam [96].

The first level of the presented model describes the revisiting of experience. The nontechnology support method would be writing a journal. Technology such as life-logging applications (see Section 3.1.2) can help to record experiences. The recorded data are easier to browse, search, and analyze. Moreover, events are often recorded automatically as a side effect of using technology. Sensors provide further means to record information about an event. On the second level, reflection starts with analyzing the data. Originally, this step was mainly supported with reflective questions meant to provoke a deeper understanding of the recorded event. These questions can be integrated as prompts or by using tags. Users are asked to tag experiences according to a reflective question (e.g. “What emotions (positive or negative) do looking at this image provoke?"). At this point, data sharing can force the reflecting person to articulate new aspects of the event. Further insights can be gained from a full discussion regarding any of the questions.

The third level involves gaining a new perspective on the recorded event. Additional data or visualization can provide new perspectives (e.g., by visualizing a trend. Hence, sensors are strongly advocated. This includes biosensors to gain insights regarding the user's own reaction as well as environmental sensors to understand the event from a more objective perspective. Another person can also offer a different perspective. Furthermore, technology can help to reorganize the knowledge by using adaptive visualizations.

On the fourth and fifth levels, new insights are gained in the form of fundamental questions, challenging personal assumptions and leading to changes in behavior. The distinguishing factor of the fifth level is that it explores the wider implications of such insights, which may then lead to new insights. These processes are largely cognitive and difficult to support. According to Fleck and Fitzpatrick [119] the role of technology in reflective learning is mainly to provide material for the reflection and support its exploration.

3.3.2 Persuasive Technology

Persuasive technology is a term coined by B.J. Fogg [120] that describes *"computing systems, devices, or applications intentionally designed to change a person's attitudes or behavior in a predetermined way."* Computers can persuade users in a variety of ways, among which the monitoring of behavior has become the most prominent. Persuasive technology offers a variety of examples regarding how to design for behavior change [121, 20, 122, 101]. Furthermore, a number of design methods [19, 123, 20] have been proposed to guide a successful design.

Fish'n'Steps [121] is one of the first and most-cited examples of persuasive technology that uses self-monitoring to trigger behavioral change. Fish'n'Steps motivates users to live more actively by taking more steps. A pedometer measures the user's steps, then this information is provided as feedback on a public display. A virtual character, a fish, reacts to the number of recorded steps. If the goal was reached the fish grows or additional fishes are added. The system was realized using some "Wizard-of-Oz" components (e.g., the pedometer value was read by a camera that was interpreted by a staff member).

Consolvo et al. [124] recognized that the feedback on the public display is available only once the user stands in front of it. Furthermore, users may not want to share their goals and progress with everyone who can see the public display. Hence, they developed UbiFit Garden, a mobile application that visualizes the step count, adapting the background image of the phone to provide unobtrusive feedback. Figure 3.4 shows the background image. A garden is shown that grows if enough steps are taken. UbiFit integrated a mobile sensor platform to recognize walking, cycling, running, and training machine uses.

Persuasive technology is not limited to step counters and physical activity. UbiGreen [122] applies a technology similar to UbiFit to encourage eco-friendly behavior. A wearable sensor captures the movement pattern and a mobile application allows the user to enter or modify recognized transportation methods. The transport mode is used to calculate the resulting carbon dioxide emissions. The feedback is again visualized on a phone screen. Mobile access to health information (MAHI) [101] supports newly diagnosed diabetes patients to adjust their behaviors to their illness. A mobile application receives glucose data from a blood glucose meter and

Figure 3.4: UbiFit Garden [124]

visualizes the feedback to the patient.

Several design methods and guidelines have been proposed to guide the design of persuasive technology applications such as the ones presented above. Fogg [19] proposes an eight-step iterative design process, because *"many projects are too ambitious, and thus are set up for failure."* Hence, each step is meant to increase the likelihood of success. In each step, Fogg suggests selecting the most simple and promising solution, even if the resulting change in behavior becomes smaller. If a step fails, the designer must go one step back and further simplify the approach. According to Fogg, designers should *"test and iterate quickly."* Fogg emphasizes the practical obstacles and advocates to learn from successful examples rather than underlying theories.

In contrast to Fogg, other researchers start from existing theories to support the design process. Consolvo et al. [20] present a set of guidelines that are inferred from theory. One of the underlying theories is the cognitive dissonance theory [31]. According to this theory, we experience discomfort if our behavior does not match our mental model. Hence, humans either feel urges to change their behavior or cling to internal images of their behaviors that impair their perceptions. Among other guidelines, the systems should be unobtrusive, and controllable. The visualizations of data should be abstract and reflective. In addition, positive reinforcement is suggested as the persuasive strategy.

Oinas-Kukkonen and Harjumaa [123] present a more comprehensive three-step design process called Persuasive Systems Design (PSD). A visualization of the PSD process is depicted in Figure 3.5. The process starts with the understanding of seven postulates behind the design of persuasive technology. The first statement emphasizes that technology is always influencing behavior (e.g., by simplifying a task, it encourages users to perform this task). Persuasive systems are designed to use this influence in a goal-oriented manner. The second point refers to cognitive dissonance theory. Here and in point 6. the model is similar to Consolvo [20]. Point 4 is especially relevant; because it highlights that most persuasion strategies work towards an incremental change of behavior. This links nicely to the minimal changes in behavior that Fogg [19] recommends. Designers of persuasive systems have to understand these postulates to recognize the possibilities and limitations of persuasive technology.

The second step of the PSD process puts the focus on the context of the

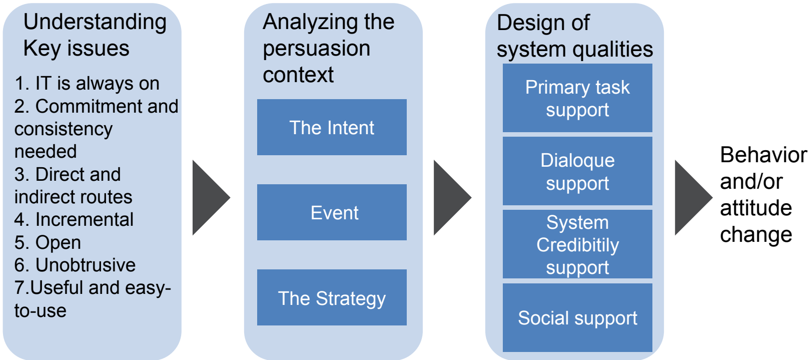


Figure 3.5: Phases in persuasive systems development, according to [123]

target audience. Similar to Fogg and Consolvo, the PSD advocates that the goal of the persuasion has to be as clear as possible. The situation of usage has to be considered which includes a careful analysis of the problem domain. In this analysis, the designer must understand the goals of the user and the user's current progress. Obstacles have to be identified and the definition of intermediate goals can be helpful. Furthermore, the available technology has to be taken into account. This information helps to define the persuasion message and how it can be delivered.

After these two steps, the actual development of the system begins. At this point, designers can select from a variety of persuasion approaches. Self-monitoring is one of the available approaches, but can be combined with other methods, such as reminders, personalization, or rewards. The PSD provides examples for each approach. For instance, the system can gain credibility by referring to people in the role of authority or by providing endorsement from respected sources.

In summary, persuasive technology is designed to achieve and support a predetermined behavior change, whereas reflection is rather exploratory in its nature. Additionally, both approaches differ in their underlying theories. Many of the persuasion strategies that involve the recording and visualization of data are based on Skinner's work on reinforcement learning [125]. A behavioral goal is defined and the current behavior is compared to this goal. Behavior that matches the desired behavior is rewarded with positive reinforcement, but different behavior results in negative reinforcement. This approach involves automation of the new behavior, while reflective learning (see Section 2.1) aims at carefully reviewing experiences. Hence, persuasive technology often aims at minimizing cognitive effort; reflective learning requires cognitive effort to analyze experiences – and deliberately encourages it. Nevertheless, persuasive technology provides approaches to successfully design for behavior change. One could even consider reflection as the target behavior in reflective learning, a behavior that is far more complex than increasing the number of steps per day.

3.3.3 Quantified Self and Reflection

The Quantified Self (QS) community is a group of self-trackers and tool builders that track a wide variety of personal data. They conduct experiments about their lives using mobile applications and sensors. Q-selfers have developed a variety of tools. They are extreme users that invest significant effort into the development and usage of tools, but this extreme usage highlights the limits of such self-tracking tools. An overview of used tools and practices within the community has been presented by Choe et al. [126]. According to Choe, the underlying motivation of self-trackers is not to collect large amounts of data, but to gain insights by reflecting on them. While the effort spent on tracking data and developing tools cannot be directly transferred to reflective learning at the workplace, the encountered barriers and used methods are an inspiration for the design of reflection support for a more diverse user group.

Li conducted two studies [98, 127] to collect requirements from active self-trackers. The results show that the data are reviewed in two ways. In the discovery stage, users explore the data to gain new insights regarding how they could improve their behavior. In the maintenance stage, users have already decided on goals (e.g., to change their behaviors) and are monitoring their progress. The maintenance stage is very similar to the goal-oriented interpretation of data in persuasive technology (see Section 3.3.2). The application should alert users if they do not achieve their goals and help to identify reasons for not achieving them. The discovery phase can be linked to the reflective learning theories presented in Section 2.1. The required features in this phase are similar to the support suggested by

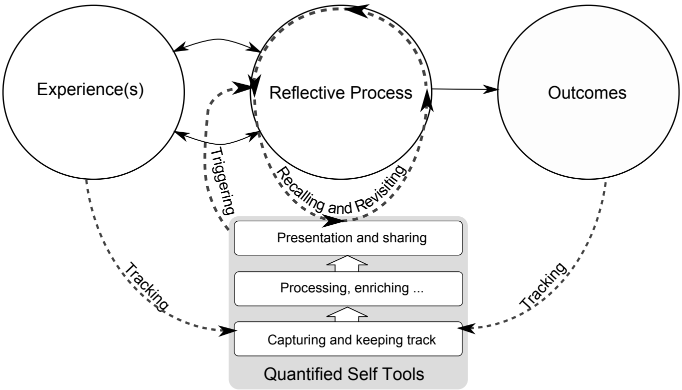


Figure 3.6: Role of the three QS potentials in reflection [204]

Fleck and Fitzpatrick [119]. A wide variety of data should be captured *"anytime, anywhere, and often"* while reducing the effort for the individual. The presentation of data should allow a comparison of data from different sources.

Rivera-Pelayo et al. [204] present a model to integrate reflective learning theory and QS tools. The model, depicted in Figure 3.6, shows how QS-tools augment the model by Boud [4] with support mechanisms.

QS tools can be used to track data related to experiences and outcomes. These data are processed and enriched to generate visualizations and share the data with others. The process of "recalling and revisiting" can be supported by the resulting artifacts. Furthermore, the data can be used to trigger the reflective process. Each support dimension is further dissected to a level that can be supported with QS technology. For example, tracking can be realized by self-reporting applications or sensors.

3.4 Summary

The support for reflective learning by data capturing is currently limited to applications that use self-reporting approaches or camera images, with the exception of the WATCHiT system. WATCHiT complements the work in this thesis, but the focus of WATCHiT is on its modular hardware design.

Although, sensors have been used in life-logging applications, the reflection on sensor data has not been studied. The Quantified Self community is currently reviving these approaches, but the developed tools are targeting the private life. Sensor based CSRL application have proposed in different forms, but we lack implementations that can evaluate their potential to support learning.

This work builds on sensor technologies to capture affective aspects and social contacts that have been developed for different use cases in affective und ubiquitous computing. These promising technologies have to be adapted and evaluated in work environments. For example, xAffect has been originally used to recognize arousal from psychophysiological data in a series of lab studies, but it has not been used in a mobile setting. The available proximity sensing approaches were focused on creating large scale datasets to analyze the resulting social networks. In summary, the sensors were used in experiments and studies but not as tools for employees.

The existing methods for designing sensor based CSRL applications have not been adapted to the current progress in sensor technology. More elaborated design methods have been suggested in persuasive technology. These methods are a starting point for this thesis, but they do not account for the explorative nature of reflection. The work by Rivera-Pelayo is similar in its goals to this thesis and was also part of the MIRROR project [32]. Similar to this thesis, reflective learning should be supported by data capturing. However, the further publications on the Live Interest Meter App [91] and the MoodMap App [191] show that Rivera-Pelayo targets mobile and web-based applications. On the contrary, this thesis researches the usage of hardware sensors to achieve the same goal. Several publications have been written in collaboration to discuss the two different approaches.

The lack of applications and established design methods indicate that the design of sensor based CSRL applications is an open challenge.

4 Designing Sensor Support for Reflective Learning

Sensors can support reflective learning by capturing more data about experiences and ultimately offering new perspectives on these experiences. This chapter analyzes the requirements for such support systems by drawing from reflective learning theory and practical experiences in persuasive technology, CSRL, and the QS community. Building on this background and knowledge about currently available sensor technology, a design space is outlined. The chapter concludes by describing the used design method in the healthcare domain by drawing from the presented design space. Parts of this chapter have been published in other forms [198, 199, 200].

4.1 Requirements Analysis and Approach

While work has been performed using sensors in persuasive technology, this research was limited to few kinds of relatively simple sensors, such as activity sensors and pedometers (see Section 3.3.2). Existing reflection support is limited to self-reporting via mobile applications or image-capturing tools (see Section 3.1). Sensor data have been proposed as additional support, but there are few applications [96, 97].

Reflective learning requires a thorough understanding of the recorded data so they can be used as content for reflection. Self-reporting can achieve this understanding in an easier manner because the reviewed data have been produced by the users themselves. The successful use of pictures as reflection support has been shown using self-reporting [7] and automatic approaches [96]. Pictures are similar to our own visual perception of an experience. Sensor data per se do not have this inherent connection to an experience. Therefore, it can be more difficult to understand and to relate the data to one's own experiences. The design of sensor-based reflection support requires both a user centered-process and a deeper understanding of the role of sensor data in reflection.

Starting with a discussion of the common and scientific understanding of reflection, the next sections will outline the similarities between reflection support and persuasive technology analyzing the underlying feedback loops. The chapter concludes by analyzing the requirements to sensor data as content for reflection in the healthcare domain and discussing the ethical implications of such support systems.

4.1.1 Understanding Reflection

When talking about reflection, the notion of the term reflection varies, not only between scientists (see Section 2.1), but also between employees coming from different domains. If we want to design for reflection we will have to make sure that end-users and developers agree on the defined goals and know the different notions of the core concepts. To this end, we interviewed employees from three different domains about their notion of "reflection" [178]. Three teachers, two business consultants, and four managers were interviewed to research a possible mapping between the common understanding of reflection with the established theories.

Teaching at schools and universities has been the original target domain of reflective learning research. Hence, it is not surprising that teachers highlighted aspects of reflection that are central elements of the theories. For instance, affective aspects were mentioned as a major topic by all teachers. The experience with students is often connected to strong emotions. Teachers try to judge objectively how their actions influenced a conflict and discussion. These are often rather fundamental considerations of a teacher's own teaching method and perspectives on a student.

Managers and business consultants reacted similarly and understood reflection as a method to carefully analyze concepts and their own activities. While being objective is recognized as one goal of reflection, they mainly understood reflection as a method of carefully analyzing not only their own behavior, but also business processes. The outcomes were often new ideas about concepts rather than the re-evaluation of experiences. Managers and business consultants strongly connected reflection to outcomes and to consideration of how these outcomes can be put into practice.

This small study showed the different notions of reflection in the three domains. Even within one domain, the understanding of reflection varied. In summary, designers who work on supporting reflection have first to establish a common notion of reflection. Within this thesis, the term reflection is understood as defined by Boud et al. [4]:

> *“those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations.”*

4.1.2 Sensor Data as Learning Content

According to Fleck (see section 3.3.1) and the CSRL model (see section 2.1.3), captured data can support reflection. The data become the learning content that provides new insights to a user. However, in contrast to traditional learning content (e.g., from a text book), the captured data about an experience do not provide insights or new concepts on its own. The data can be understood only in light of the experience in which the data were captured. Sellen and Whittaker [97] speak of *“Synergy not substitution”* to emphasize that the capturing of data cannot replace the memory and the analysis by a user. The captured data have to relate to the experience of the user to lead to insights. As a consequence, the two main requirements for reflection content are (a) that learners can relate their own experiences to these data and (b) that this process can lead to new insights. Insights can range from an unobservable change in attitude to a change in behavior [4].

CSRL should not be confused with technology that involves only behavior change, such as persuasive technology (see Section 3.3.2). Although, CSRL and persuasive technology both involve supporting and influencing cognitive processes by technology, outcomes of reflection only rarely result in direct observable changes, but rather change the mindset of the reflecting person. Examples are the initial awareness of a problem, changes in attitude, or simply the insight that there is no real problem. This section analyzes the similarities and identifies how CSRL can benefit from the advances in ubiquitous computing and in particular from persuasive technology.

Many applications from persuasive technology (see Section 3.3.2), such as [122, 124], create a feedback loop to influence behavior. The user is constantly provided with feedback regarding current or past behavior to influence future behavior in the desired direction. Such feedback loops are also the basis of many reflective learning theories (e.g., the Kolb cycle [5] or the model by Schön [6]). The model by Boud et al. [4] is not explicitly designed as a loop, but the process of re-evaluating experiences can and probably has to be repeated to learn continuously. As described in Section 2.1, these reflective learning processes target an incremental improvement and learning from new experiences. Similarly, the CSRL model, which serves as the background for this approach, forms a cycle.

Figure 4.1 depicts how sensors and visualizations of sensor data can support reflection and behavior change. The model, similar to the Kolb cycle [5] and the model by Schön [6], is structured in a feedback loop. The naming of three of the four phases is based on the model of Boud et al. [4] because the CSRL model [33] itself is based on this model: *Experience*, *Reflective Process* and, *Outcomes*. The fourth phase, *Behavior*, is not an explicit part of Boud's model, but the required source of experiences to close the feedback loop, similar to the models by Kolb and Schön. In the cycle, behavior results in experiences. During the reflective process these experiences are revisited and analyzed to come to new outcomes. Some of these outcomes will translate into an actual change in behavior that can be analyzed in subsequent cycles.

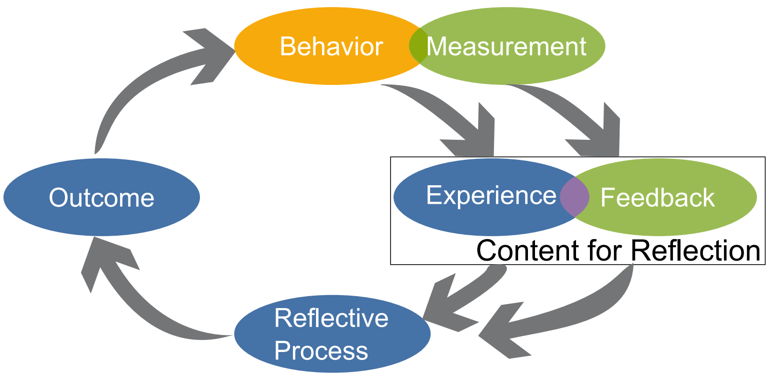


Figure 4.1: Double feedback loop of computer supported reflection

Each transition between phases may fail. Behavior or the impact of own behavior might not be perceived at all and, therefore, does not become part of the experience. An important and relevant experience might go unnoticed because of the amount of experiences that we make every day. As a result, this experience never becomes an item of the reflective process. A reflective process might not be fruitful if it is not creating outcomes and may then turn into rumination. Finally, an outcome that could lead to a change in behavior might be prevented by the circumstances. For instance, an employee who decided to stop working late might suddenly be confronted with a large number of urgent tasks.

Technology can help at various points to minimize error and facilitate successful reflection, influencing attitudes and potentially changing behavior. This thesis focuses on the support of the reflective process by facilitating the transitions from the behavior to the reflective process stage. Further technologies can support the other parts of the cycle. For instance, a system can guide reflection with reflective questions or help to document and communicate outcomes.

The envisioned sensor support creates a second loop with two new components that support the cycle. Behavior is not only perceived by the individual but measured by sensors. In addition, the captured data or a summary has to be provided as feedback to enrich the perceived experience with additional data [203]. The combination of experience and feedback will influence the reflective process, the resulting outcomes, and finally the behavior. The change in behavior is experienced and sensed to create a comprehensive feedback to influence the reflective process. According to Boud et al. [4], the reflective process involves a *"stepping back from experience."* This step to obtain a more objective perspective on behavior can be supported by additional data that provide a different and often more objective perspective.

In summary, a system that is intended to support the reflective process has to close a second feedback loop by (a) capturing data linked to relevant behavior and (b) providing these data as feedback that can enrich the experience itself. The reflective loop and the feedback loop have to be as closely intertwined as possible. This design task is challenging because the internal reflective process can be observed only by its visible or articulated outcomes.

The captured data can facilitate new insights in different ways during a reflection session (see Section 2.1.3).

- Data that provide a different perspective on experience can actually trigger a reflective process by inducing a cognitive dissonance [31].
- In some cases, the data might act as markers to indicate a particular point in time, a particular location or process upon which it is worth reflecting. Although these data do not lead to insights in themselves, they can point to relevant data or experiences.
- Finally, additional data may be required during the reflection session for an in-depth analysis of an event.

The impact of the collected data can be increased by aggregating the data or sharing it with others. The aggregation over time, multiple employees, or events can provide new perspectives on experiences. The resulting high-level perspective can be useful to abstract from the specific event and generalize insights for an individual (e.g. reviewing trends in bio-signals) or collaborative level (e.g. reviewing team performance over one month). Sharing of data enables all forms of social facilitation of reflection (see Section 3.1.3). Users can benchmark their own results in relation to those of their colleagues, or reflect in collaboration with them. For instance, employees could compare performance measures of a specific task to those of a more experienced colleague or to the average performance of all employees. The benefits of collaborative reflection are described in more detail in [93]. However, sharing of data is not always possible, or desired, at the workplace. Especially at the workplace, privacy considerations must be taken into account. Furthermore, colleagues have to be able to understand and interpret the data from their own perspectives.

4.1.3 Requirements Analysis

A hospital and a care home were visited in order to understand the specific requirements in both domains. In local workshops, management and employees presented their perspectives on reflection and how they expect technology to support their work. Developers presented existing technology approaches and discussed them with employees. Building on this rough understanding, ideas for technology support were drafted. A second round of workshops was conducted with employees and developers to discuss and extend these ideas. These discussions provided deeper insights about the requirements. The results of the requirements analysis for each testbed can be found in the corresponding chapters.

The healthcare domain is a particularly promising field for reflective learning applications. Large parts of physician, nurse and care staff education occur in the form of on-the-job training. The procedures during an emergency situation can be explained in a book, but must be experienced firsthand for one to act confident and competent in similar situations. The time to reflect on such important experiences is scarce.

During a workshop in a hospital, a lead physician confirmed that reflective learning is encouraged, but is not sufficiently supported. Technology support is seen as presenting an opportunity to reflect in a more effective manner by providing alternative perspectives on experience. Moreover, the management hoped that the new tools encourage an overall more reflective attitude towards work. Adherence to organization standards and privacy requirements of employees are major non-functional requirements to these tools. Employees are concerned that captured data will be used to assess their work performance. Moreover, the privacy of customers and patients is even more critical in the healthcare domain. These special requirements limit the applicability of well-known sensors, such as the SenseCam [8]. Alternative options are self-reporting approaches, but they are often interfering with daily work or have to be highly customized to integrate with existing tools and applications.

The design of tools to generate the desired content in an appropriate manner must balance the effort to capture these data and the expected benefit. At the workplace, the available time and budget are generally already allocated by the primary work tasks that constitute the business of the employer. Therefore, additional tasks and cost must result in a measurable outcome on the business side. As described previously in Section 2.1, there can be a wide variety of outcomes of reflective learning, ranging from unobservable changes in attitude to an unexpected and sudden change in behavior. The impact is hard to predict and estimated benefits are difficult to prove. For example, if an employee in the sales department decides to act more empathically towards a customer, the impact may be improved sales, but there is no direct predictable link. Hence, the balancing of benefits and barriers for reflective learning at the workplace mainly involves reducing barriers. In some cases, benefits that are side effects of reflection are important as well.

4.2 Design Space

This section organizes the main design decisions and options to support reflective learning by data capturing into a design space. According to Shaw [128], design spaces help to meet requirements by identifying core questions. They provide guidance as well as a framework to compare and analyze existing solutions. Hence, Shaw defines:

> *"The design space for a problem is the set of decisions to be made about the designed artifact together with the alternative choices for these decisions."*

This section is not limited to sensor technology, but describes the choice between sensors and self-reporting applications as one decision among others because this is one of the choices developers face. Sensors are promising tools, but can be complemented and combined with self-reporting approaches [90].

A wide variety of workplaces can be supported with different capturing solutions. Designers have to understand the specifics of the workplace to identify major requirements and limitations. In terms of the CSRL model, the capturing of relevant data during the work process augments an experience and becomes part of the reflection session frame. Furthermore, the captured data and their visualizations can be triggers for a new reflection session. The following three questions must be answered in an iterative process:

- What should be measured? What is the relevant context?
- How can the relevant context be measured at the specific workplace? Which capturing method is appropriate?
- How can the measured data be turned into feedback? How can they be turned into content for reflection?

The following sections will discuss each decision and describe options and related examples. Figure 4.2 shows these decisions as dimensions in the context of the CSRL model. These three questions are interdependent.

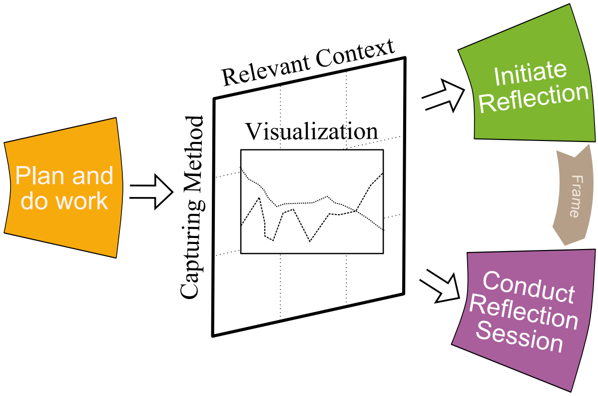


Figure 4.2: Design decisions to turn context into content, in relation to the CSRL model.

The selection of certain context limits the opportunities to capture it. Moreover, there is often a wide variety of context information that cannot be captured at all. The feedback builds on the captured data and the possible feedback method determines which context is optimal. Therefore, the optimal answer to all three questions in a given workplace must be found.

4.2.1 Relevant Context

Workplaces offer a plethora of data that can be captured or is already available in digital form. The selection of data is driven by the considerations explained in Section 4.1.2: learners must be able to relate their own experiences to the data and come to new insights. Furthermore, the design challenges described in Section 4.1.3 highlighted that the goal should be a small subset of data to minimize interference with ongoing work processes. The decision of which data are relevant is a typical example of the *relevance paradox* [129]. While a designer has all data at hand, it is difficult to decide which subset will be relevant in a reflection session because of (a) the unpredictability of context relevance and (b) the subjectivity and need for interpretation of context.

- *Unpredictability of relevance*: Since the outcome of reflection cannot be clearly predicted, (i) it is unclear which context is useful and (ii) more context information has to be captured than will probably be used afterward. Developers can only estimate which data might be useful and, thereby, they have to balance required effort and estimated impact. Which data are useful to reflect on an experience? Which experiences are worth reflecting?
- *Subjectivity and need for interpretation*: It is inherently difficult to identify data that relate to a concrete experience. While hardware sensors and IT systems can capture a growing part of the context, the perception of this context and its interpretation is hard to estimate. The perception by a user depends on existing experiences and biases. Only the user can provide the necessary feedback to select the relevant subset of data for later reflection and, in many cases, this selection is already part of a reflective learning process. Hence, designers cannot solely rely on fully automated ways of capturing (such as hardware sensors or mining of existing data). A combination of multiple sensors may provide additional hints on the relevance of the data, but applications can involve the user into this process.

During the workshops reported in Section 4.1.3, participants repeatedly referenced three types of context in stories on reflection: task context, affective context, and social context. They are promising candidates to support reflection at the workplace with data. This is not an exhaustive list, but can help to identify alternative options. For instance, the location is relevant in many professions, but is mainly interpreted in the sense of social or task context (i.e., "what did I do there?" or "who was I with?") The different roles of these three context types in the reflective process are described below. In addition, example tools and related challenges for measuring this context are provided.

- *Task context* describes activities that are directly related to work tasks. The data are easy to understand and seen as relevant by employees and management. The required data are often already present in existing documentation of work (e.g., as treatment records in a hospital), but might not be available for technical or legal reasons. Activities at a PC can be captured with commercial tools like

RescueTime [130]. Sensors can augment tools and the environment with sensing capabilities that capture task context. For instance, wearIT@work [15] supported workers in a car factory by augmenting tools with RFID tags. All interaction with tools was captured by an RFID reader at the glove. However, capturing task context can be perceived as an undesired performance-monitoring by employees. The resulting resistance of employees can lead to manipulation of data or the corresponding tool may not be used at all.

In summary, task context is relevant in nearly all workplaces. The developed solutions are often tailored to the specific tools or applications, but help to provide better understanding of the selected work processes.

- *Affective context* describes all context information that is related to emotions and moods. Emotions and moods are indicators for relevant time spans during the work day. Important events are often experienced with a higher level of arousal. Furthermore, continuously high arousal may indicate an overburdening of employees. Therefore, affective context can act as a marker for relevant time spans as well as additional data to analyze the impact of emotions on an experience. Moreover, awareness of mood and emotions is important for all employees that directly interact with customers or patients. This kind of work is also called emotional labor [131]. Affective computing (see Section 2.3) has developed a large variety of sensors [57, 62, 56] and self-reporting approaches [191] to record affective context.

  Affective context is useful in a variety of roles and has been a central point in reflective learning (see Section 2.1.2). The available sensors and tools are designed to be used across work contexts. However, data about our emotions are highly privacy relevant.
- *Social context* records the interactions with colleagues, customers, and patients. In interviews, employees often begin explaining an experience by mentioning the people that were present. Persons and their names act as a memory cue for experiences that were made together. Moreover, in many professions, the social contacts structure the day. Physicians, for example, treat one patient after the other and talk with relatives and colleagues. In this case, the social contacts correspond to the task context. Social contacts can

be face-to-face or mediated by technology. Technology-mediated contacts can be analyzed using the corresponding communication tool and social network analysis (SNA) [118]. A variety of sensors and mobile phone applications can record face-to-face contacts (see Section 3.2.2).

Social context is relevant for many professions, is privacy critical, and can be captured by sensors and applications across contexts.

All three types of context can provide a rather objective or subjective view on a situation, depending on the capturing method that will be discussed in Section 4.2.2. The comparison of these possibly conflicting views can induce a cognitive dissonance, reveal new insights, and eventually lead to a change in behavior. A subjective and personal experience is the result of the individual's interpretation of the situation. The return to experience and re-evaluation of an experience in the model of Boud et al. [4] targets this subjective and personal experience. In being subjective, it might not always reflect an actual situation, but is interwoven with personal beliefs and expectations. Subjective and personal experiences are accessible only for tools and others when the learner articulates this knowledge. This articulation might be complicated by the tacit nature of a part of the knowledge at work. If knowledge is articulated, it will generally be subjective because it reassembles the mindset of the learner. This view can be further distorted by the method of recording (e.g., notes might be misinterpreted after reading them at a later time).

Objective perspectives differ from experiential data in their underlying perception of the situation as well as their interpretation of the gathered data. For instance, sensors have only a limited view of a situation as they capture a small subset of an experience in great detail. A learner generally has a broader perspective, but might interpret the perception or even the experience incorrectly. As Moon points out, *"If learners do not learn from experience but from their perception of experiences, there are implications for the nature of guidance required by learners in order to make sense of experience"* [26].

4.2.2 Capturing Method

There are often several methods to capture the same or similar data at a workplace. For instance, affective aspects can be monitored by sensors [57, 56] or self-reporting apps [191]. The best method depends on the requirements of the workplace, because the selection of the capturing method defines the qualities of the resulting data and the effort to capture these data.

- *Quality* of resulting data includes, among other factors, sampling frequency, precision, and accuracy. A higher sampling frequency and precision are beneficial to the point at which no new information is contained. Conversely, the accuracy can be lowered by a systematic bias that does not have to be negative as long as users are aware of this bias. For instance, self-reporting of data introduces the bias of the user into the system. The bias itself makes it easier for a user to reflect on the data. Moreover, the bias may become the topic of reflection as in Echo [7]. An unknown bias, however, may lead to incorrect conclusions in the reflection session. Sensors can introduce a bias as well (e.g. by being positioned in only one of several rooms). The sampling rate may be inappropriate to measure particular events, or the selected sensors may exhibit a general lack in accuracy.
- *Efforts and costs* to capture the data are the main reasons why a system is not accepted by employees or a decision is made against installing the system in the first place. The efforts and costs include: costs to the employer, effort for the employee, and legal constraints. All three can come in varying and unexpected forms. For example, developers are often not aware of the required effort to train employes to use the system. These barriers exist and should be carefully analyzed in collaboration with the end users. Efforts that are accepted in one place may not be accepted in another.

Designers will often encounter a tradeoff between effort and quality of the data. A diary application can be used once per week or every day. The amount of data and the possible insights will differ. Likewise, more complex and expensive sensors can often deliver a higher accuracy. The methods to capture data fall into one of the following three categories:

- *Self-reporting* of data relies on the active effort of a user to report events and the user's own impressions. Mobile applications and blogs have largely replaced the classic handwritten journal. Because of user involvement, the resulting data will be biased by the reporting person. Some QS applications attempt to minimize this bias by restricting the input to a specified structure that can later be analyzed automatically. For instance, the MoodMap App [191] restricts the input to a single click in a two-dimensional space. However, as in the case of the MoodMap App, the collected information is often directly related to the personal subjective experience.

  Self-reporting can be used in a wide variety of scenarios, but requires the cooperation and acceptance of users. The motivation of users will determine the amount and quality of captured data. The user interface plays a crucial role in guiding and motivating the user.

- *Observer-reporting* in its properties to self-reporting, except that the effort to capture the data is moved to an observer. The observer can be a single mentor, for instance an experienced nurse in a care home, or a large group of observers. The feedback from groups is especially valuable because, although the bias from each observer influences the final result, the overall result will contain only an average bias. The aggregation of this feedback can provide an objective external perspective on an event. Customer surveys build on this principle. However, as observers have to capture data by themselves, their motivation is crucial for success. The Live Interest Meter app [91] supports presenters by feedback from their audience, which is motivated to provide such feedback by leveraging on their interest to listen to an engaging talk.

  Observer-reporting can be applied when observers are available and if they can be motivated to share their views on an experience. Observer data can become more objective by aggregating feedback from multiple observers.

- *Automatic* capturing of data is realized by either sensors or monitoring applications. Automatic approaches can capture one detail (e.g. the room temperature) at a high sampling frequency and precision. The user acceptance depends on two conflicting arguments. Sensors and applications remove the reporting burden by automatically recording data, but the monitored person is no longer directly in control of the recorded data. This can lead to feeling monitored by the management and lower the acceptance of such systems among employees. Automatic capturing systems must include a means that enables employees to reclaim ownership of the data. Sensor data are often judged as more objective, but can be biased by the technology or the usage of this technology. These biases are not often obvious to a user and must be communicated clearly.

Automatic capturing methods can deliver a higher sampling frequency and precision. They often provide a very limited perspective, but a maximum granularity on an experience. The costs to design and introduce a sensor system are related to required hardware and software, so they tend to be higher than reporting approaches. Furthermore, hardware and software often limit existing sensors and tools to very specific domains. For instance, the gloves and tools in the wearIT@work prototype [15] are tied to car manufacturing in one factory.

The type of physical tasks conducted by the users will guide the type of capturing method. For example, if users need freedom of movement and cannot use their hands to input any data, self-reporting can be done only with custom input methods [90]. If, however, users can embed the manual capturing of the data in their current working tasks (e.g. because they work at a desk or these data have to be recorded anyway), self-reporting can be the better option.

Although sensors are expected to provide an objective perspective, planned or involuntary interpretation, done by the learner or the capturing method, might skew the view of the perceived situation. Hence, the difference between objective and subjective data are only relative. Conversely, an objective perspective provided by automatic means (e.g., a video or a picture) already includes subjective bias. The person taking a picture with the camera is focusing on a selected aspect and cuts out other relevant elements. Sensors can capture with high precision, but they lack the broader view of a human observer. Additionally, by selecting the monitored detail, a subjective bias is already introduced. However, data that are biased by subjective impressions can become more objective through aggregation across several incidents. For instance, the different perspectives provided by observers are subjective and biased, but by quantifying and summarizing the feedback of multiple observers, a more objective feedback is possible.

4.2.3 Visualizing Captured Data as Content for Reflection

Developers of CSRL applications must visualize context in a form that optimally complements the experience of the user to trigger reflection and lead to new insights. Raw data must be transformed into an accessible form to become content that supports reflection. The resulting feedback is the main part of a capturing system that is visible to the user. Therefore, its quality determines how users will perceive the value of the system. The design of the required visualizations must include three design goals, which in some cases can be conflicting:

- *Simple and aesthetic* visualizations are appealing to users and easy to understand [20]. The cognitive effort to analyze the data is reduced. The visualizations have to trigger and sustain interest in the data. Optimally, a simple message is conveyed that guides the user to an outcome.
- *Surprising and inquisitive* interfaces lead to new insights. They provide new perspectives, spur interest, and foster the analysis of the data. Unusual and surprising data points are emphasized and can result in cognitive dissonance [31].
- *Comprehensive* visualizations allow data analysis in all details. A user can drill down to specific situations and analyze different relations between the data. The visualizations do not artificially reduce the options to analyze and understand the data.

There is a tradeoff between data being simple to understand and the possible depth and impact of insights. Large multifaceted data may promise major insights, but if users cannot understand and relate to these data, a simpler solution should be preferred. This is in line with the design guidelines for persuasive technology by Consolvo et al. [20] and Fogg [19].

Research on information visualization [132] has developed a rich variety of visualization methods. Rivera et al. [203] discuss the different perspectives that are relevant for CSRL. In the following, three of them and their role in reflection are explained:

- *Status* charts provide a quick overview of the data. They relate data to goals or average values. Examples would be the average mood during a day, the number of tasks that have been completed, or the average feedback from customers. Deviations can be a trigger for reflection and lead to an analysis of the underlying reasons.
- *Timeline* charts offer a historical perspective on data. Situations can be analyzed along the timeline and durations can be compared. Timelines can span a few minutes, days, or several years. Long-term timelines can help to identify trends. Shorter timelines can be used to reconstruct specific situations.
- *Comparison* charts help to compare data between several instances of an event. This can be a comparison between different users, trends during different timespans, or between different groups.

In an application, a selection of these charts and further visualizations must be combined to cater to the specific needs in a workplace. For instance, a status or a comparison chart can act as a starting point for the analysis. Further charts can be used to follow up if interesting deviations or differences are found.

The optimal visualization depends on the background of the learner as well as its knowledge of the data. A user-centered design process is required to collect requirements from the end users in an iterative fashion. Feedback on early prototypes guides the selection and arrangement of charts.

4.3 Design Studies in the Healthcare Domain

Two design studies have been conducted with wearable sensors to explore two of the three main categories of context presented in Section 4.2: affective context and social context. From the presented technologies, wearable and mobile systems best matched the requirements of employees in most healthcare professions because their work is not bound to a desk. Carers, nurses, and physicians are moving between patients, relatives, and colleagues. They often share desks (e.g., at the stroke unit two desks are shared by all nurses). Sensors that should be used in this environment must be unobtrusive so as not to interfere with daily work. Wearable sensors (see Section 2.2.4) can measure relevant data in an unobtrusive manner and accompany their users wherever they are. Therefore, they were selected in both studies as the preferred capturing method.

The following sections introduce the selected use cases and describe the common approach that was used in both domains to design prototypes according to the design space presented in this chapter.

4.3.1 Use Cases

Two examples in the healthcare domain in which reflective practices are particularly promising are stress-coping strategies and dementia care. Both require individual solutions that have been tailored for the individual and the situation at hand. Hence, only the individual can identify these solutions because the required knowledge is tacit and very specific. While these insights, such as an identified lack of a particular skill, can be painful in the long-term, the overall well-being will increase. The overestimation of one's own abilities might be motivating and beneficial in the short term, but has negative effects in the long term [133].

Chapter 5 explores the recording of affective aspects within a stroke unit by equipping nurses and physicians with wearable heart-rate monitors. Nurses and physicians, especially in emergency care, work under stressful conditions. The Health and Safety Executive [134] in the UK summarizes that stress has been the second most commonly reported type of work-related illness in the UK for several years. Health and social workers have the highest rate of illnesses across all occupations and industries. Stress coping strategies in the hospital range from breathing exercises to suppression. However, the optimal situation would be to remove the causes for stress. As these causes are specific to the individual, only reflection by the individual can identify them. Although coaching could help, these forms of support are too expensive to be used for all employees.

Chapter 6 describes the development and evaluation of a wearable sensor system to record social contacts in a care home. One reason for the increasing workload in care homes is the growing number of residents with dementia. Caring for dementia patients requires quick reaction to the often unpredictable behavior of elderly residents who suffer from this disease. Dementia is an age-related illness that will become more prevalent with the growing number of elderly citizens. Current projections [135] estimate the total number of people with dementia in the UK at more than 1 million by 2025 and more than 2 million by 2050. Care work in the UK is mainly conducted by care staff without formal training. During their work carers, acquire the skills to achieve National Vocational Qualifications. As a result, care staff members without formal training must deal with the complex behavior of dementia patients. Carers must learn person-centered care, viewing residents as individuals and striving to look at the world from the perspective of the resident with dementia [136]. Many carers struggle in this challenging environment and eventually quit their jobs. As a result, an annual staff turnover of 20 percent to 25 percent is common [137].

4.3.2 Iterative Design Process

Development was conducted using an iterative user-centered approach, as shown in Figure 4.3. Users were included throughout development process, starting with collaborative requirements discussion, testing of early prototypes, and iterative development of requested features. The process consists of three major steps that are aligned to the dimensions of the design space presented in Section 4.2.

In an initial step, the type of context that should be captured is selected. The selection depends on the workspace and underlying goals of employees. Task context helps to improve and compare performance. Affective context is relevant for well-being as well as social interaction with customers and colleagues. Social context supports the analysis of social contacts. Depending on the workplace and the specific type of data, each of the

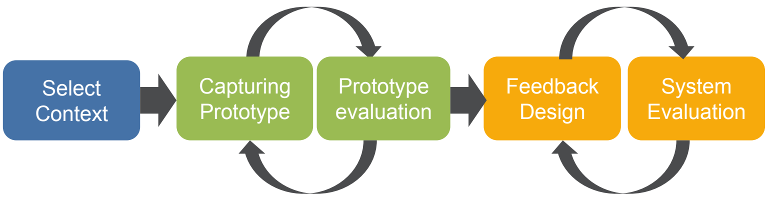


Figure 4.3: Iterative design process

context types may be used for a different purpose. For example, social context of care staff can help to analyze task performance because the time spent with each patient is directly related to the work task. In summary, the selection of context that should be captured requires discussion and interaction with the end users (i.e., the employees).

Based on this decision, prototypes to capture this context are built and tested. A prototype may use off-the-shelf tools and visualizations of the data can rely on the raw data. The evaluation has to take place in the target context to test whether sufficient data can be captured and how much effort is required. The first test at the workplace should take place as early as possible to identify unexpected challenges. For instance, in the care home, several staff members had problems writing or typing. Therefore, self-reporting approaches in the care home should not contain any text entries. Even showing raw data to employees in a very simple form can provide feedback whether the quality and amount of data are useful for reflection. These first reflection sessions were facilitated by a researcher and evaluated with semi-structured interviews. If the developed prototype does not deliver the required quality (e.g., an insufficient amount of data or an unexpected bias) or if the efforts and costs are too high, the prototype must be adapted and tested again.

The discussions in the first step and feedback during the prototype evaluations provide many insights regarding the possible design of the feedback. However, during these tests, employees imagine what the final feedback will look like and which new and deeper insights will become possible with more-advanced visualizations. The first visualization prototypes will most likely not fulfill these expectations completely. Consequently, another cycle of iterations is required. In many cases, the captured data must be filtered and aggregated to minimize the effort to analyze the data.

In the two design studies, the focus is on the second step – the evaluation and improvement of capturing prototypes. In both design studies, prototypes for capturing context have been developed and evaluated.

5 Design Study I: Capturing Affective Context

This chapter explores the support of reflection by capturing affective aspects following the design process defined in Section 4.3. The first step in this process is the decision regarding the relevant context according to the design space defined in Section 4.2. The affective context was chosen because tracking of affective context during the work day can help to identify the experiences that trigger an emotional response. These experiences most likely indicate critical events that are worth reflecting. There are three additional benefits to capturing affective context. First, emotionally arousing experiences are better recalled in the long-term [138]. Hence, tracking of affective context during the work day could help to identify remembered experiences and timespans. Reflective learning support could refer to these episodes and also point to events that might have been forgotten. Secondly, the affective context is relevant in nearly all workplaces. A developed solution has the potential of being generalizable to other work domains. Finally, wearable sensors have been developed in affective computing (see Section 2.3) that do not interfere with work tasks [56, 57, 62].

The conducted design process consisted of three major steps, as depicted in Figure 5.1. The structure of this chapter follows these three steps. As explained above, the decision to capture affective context was the starting point. In a first step, a prototype to capture affective context was developed. To this end, a requirements analysis was conducted, as described in Section 4.1.3. After comparing self-reporting and sensors-based approaches in relation to the hospital setting, an available sensor system was chosen as first prototype. This system was evaluated in an initial ethnographically inspired study to deepen the understanding of affective aspects in a hospital. The study is reported in Section 5.2. The results indicated that a different, more flexible system is required. The

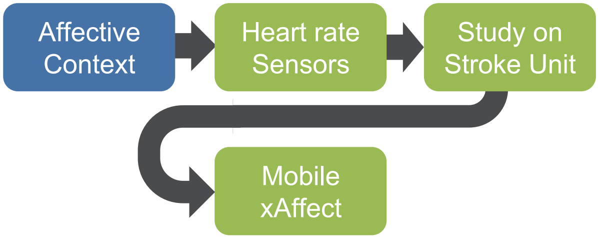


Figure 5.1: Design process affective context capturing

deduced requirements informed the development of mobile extensions of the xAffect system [206] that are presented in Section 5.3. The final section summarizes the main findings on capturing affective context.

5.1 Tracking Affect on a Stroke Unit

The healthcare environment challenges employees with dynamic tasks that must be conducted in a limited time. Each of these tasks might literally be vital for a patient. As a result, employees in social care have the highest rate of stress-related illnesses in the UK [134]. The personal reasons behind this number vary according to the particular workplace and individual mindsets of the staff. Therefore, there is no general solution, so the underlying reasons must be identified. Reflection by the individual staff members is one option to identify the appropriate reaction. Reflective practice is seen as a particularly promising approach in care professions [3]. In addition, research has shown the impact of collaborative reflection on work in healthcare professions [34, 139].

The following section compares automated approaches to self-reporting approaches for capturing affective aspects on a stroke unit. It describes the selection of the used sensor and the evaluation of an alternative self-reporting approach.

5.1.1 Affective Context on a Stroke Unit

A stroke unit is an emergency unit that specializes in treating patients with acute strokes. During a stroke, parts of the brain are cut off from the oxygen supply. If these parts of the brain do not receive enough oxygen, the damage will increase with every minute going by. One of two stroke treatment options must be chosen as quickly as possible. The two options depend on the actual type of the stroke: the blood vessels are either clogged or suffer from internal bleeding. The treatment for one stroke type will worsen the other. Therefore, all activity in the stroke unit focuses on identifying the type of stroke as quickly as possible. One physician summarized this task as “time is brain.”

Physicians and nurses treat patients with life-threatening conditions and must make decisions as quickly as possible. Moreover, they must communicate the patient’s status to relatives. For instance, they must inform relatives of the death of a loved one or about lifelong impairments. The staff members have to cope with the resulting stress and emotional pressure. According to [140], the tracking and reflection of affective aspects could help individuals to raise self-awareness about their well-being and prevent related illnesses. For instance, nurses and physicians could compare the impact of night shifts and stressful events during the day. Trends could be recognized and employees could act earlier on them. The collected data have the potential to identify stressors in the workplace. Moreover, counter measures could be checked regarding their effectiveness. Thus, we hoped to support staff by tracking their arousal levels and supporting reflection on coping strategies.

There are three major challenges when developing and introducing a stress monitoring in hospitals. First, physicians and nurses are not bound to a desk. They move between patients, relatives and their desks. Secondly, the introduction of additional hardware has to overcome several hurdles. Hospitals are careful regarding new devices that require wireless connections, because radio technologies might interfere with medical devices and ongoing measurements. Furthermore, new devices, especially when shared among staff, are seen as possible carriers that can spread infections. Disinfection of devices is a must, because patients are vulnerable to these infections. Finally, data collection of any kind is critical because staff members in hospitals fear legal implications of the captured data. Although they want to protect the privacy of patients, they also are aware that these data could be used against them if they are sued for a mistake. For instance, if a physician recorded and noticed high stress levels before a mistake, she might be held liable because she did not react to these data.

The captured data from a sensor or an application could be a used as evidence against the physician in court.

5.1.2 Self-Tracking Applications

There is a wide range of self-reporting apps that track moods or emotions [140]. Mood-capturing applications are frequently used in the Quantified Self community. Moods and emotions are very personal and differ in their interpretation between individuals. Hence, a self-reporting application can collect data that are correct and easy to understand. Mobile or web-based applications are nearly ubiquitously available. Nevertheless, even a single choice requires a short change of focus to the self-capturing app. Therefore, the number of mood entries will be low compared to a sensor; and in time-critical situations, the application will not be used at all.

To further the understanding of such applications, a simple mood self-tracking app was evaluated in a project meeting [195]. Figure 5.2 shows the interface used to minimize the effort to indicate mood. The MoodMap app is based on the dimensional model of affect by Russel (see Section 2.3). The horizontal axis depicts the valence. The vertical axis shows the arousal. Users can state their mood by clicking at any point in the map. Resulting mood values are represented as a pair consisting of an arousal and a valence value. In this way, mood can be collected with a single click. The arousal and valence values can be easily aggregated and visualized in different forms.

Mood self-reporting is not applicable on a stroke unit. In the hospital, nurses and physicians have no time to indicate their moods and such an application will most likely interfere with their work. Moreover, nurses and physicians do not typically use smartphones and spend only a small amount of time per day at a PC. However, more advanced versions of the MoodMap app have been used in different settings (e.g. call centers, where agents sit at a PC).

5.1.3 Psychophysiological Sensors

Sensors can continuously record psychophysiological signals, not only during routine work, but also during emergencies when there is no time

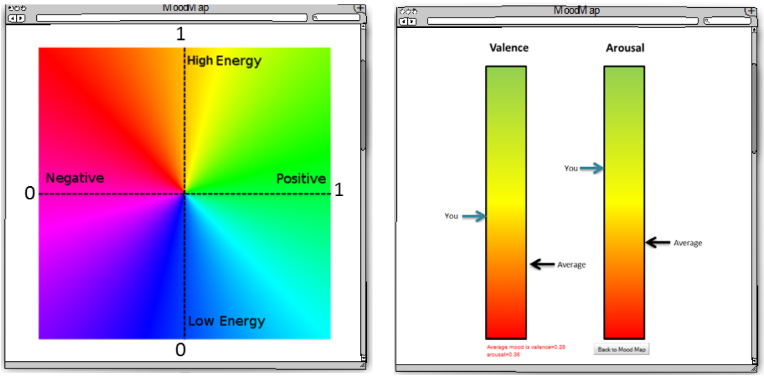


Figure 5.2: Moodmap application prototype [195]

for self-reporting. Wearable sensors can be worn unobtrusively and do not interfere with existing work practices. They can record physiological signals that indicate the affective state of the wearer (see Section 2.3). The individual review of the recorded data can facilitate individual reflection, as explained in Section 4.2.2.

Employees are rather skeptical regarding the use of wearable sensors, according to an initial questionnaire distributed to employees from hospitals, care homes, and IT enterprises [191]. We were interested in events that may contribute to stress-related illnesses. Towards this end, our goal was to measure arousal during the work day. Wearable psychophysiological sensors that are used to measure arousal can be categorized into two main groups according to the recorded signal:

- Sensors that measure the electrodermal activity (EDA)
- Sensors that measure activity of the heart (see Section 2.3.1)

Both signals have been shown to correlate with the arousal state in lab studies [205]. The wearable devices transfer these methods to the field. In this thesis, only heart rate sensors, such as [56, 14], have been used.

First studies with the Q Sensor [57] from MIT and a similar system developed by Philips [62] have indicated the opportunities of measuring EDA with wearable sensors at the workplace. For decades, researchers assumed that affect-related EDA signals can be measured only at inconvenient positions of the body (i.e., at the palms of the hands or under the feet [141]). The two groups have developed wearable sensors that measure affect-related EDA signals at the wrist. Although the initial results are encouraging, the company that was founded to produce and sell the Q Sensor stopped its development and distribution. There are no critical publications available, but the license agreement of the Q Sensor required consent from the sensors' producer to publish research using the sensor. Similarly, the prototype from Philips has been only accessible to a small group of researchers. Therefore, a validation of the reported findings was not possible.

ECG signals can be used to analyze heart rate variability (HRV) to infer the arousal level of a user, as described in Section 2.3. A number of commercial devices from the fitness domain are available, but these standard heart rate monitors for sports [14] often interpolate the heart rate to improve the user experience. The resulting data cannot be used to calculate the HRV. The ekgMove [56] comes in the form of a belt, as shown in Figure 5.3. The ekgMove contains a 3-axis accelerometer and uses dry electrodes to continuously measure the heart rates of staff members. Dry electrodes can be used without prior preparation, offer comparable signal quality [65], and can be removed without traces. After the sensor is started, participants can put on the sensor belt without further introduction.

Figure 5.3: ekgMove sensor: belt with dry electrodes on the left, worn around the chest on the right.

5.2 Ethnographically Informed ECG Sensor Study

The ekgMove was evaluated to test if it can create awareness of one's own arousal level and persuade medical staff to reflect on their own behavior and coping strategies. This first study aimed at evaluating the capturing method and gaining first insights about the possible impact of the collected data. To this end, the existing Movisens hardware and available visualization tools, such as the UnisenViewer [142] were used to record and show ECG and activity data. Observations, questionnaires, and interviews provided reference data and insights about usage and the perceived benefit of capturing these data. The insights gained in this study should help to refine the capturing prototype and collect requirements for visualizations that would have been included in a second prototype. The study reported herein was published in other form in [201].

The remainder of this section describes the method of the study and discusses the results from the analysis of the sensor data, the observation, and the concluding interviews with respect to the captured data, usability aspects, and the subjective potential to learn from these data.

5.2.1 Method

The challenge in evaluating such a system in the field is the lack of reference data. Post-study interviews about stressful phases during a day can yield only a few examples of arousal-related reactions. Furthermore, the interview will be biased by the staff's own perceptions. Therefore, sensor usage was combined with an ethnographically informed study. An adapted rapid ethnographic method [143] was used to collect requirements and opportunities of psychophysiological sensors in healthcare professions. Ethnography is based on the idea that researchers are embedded in the social settings to understand why things happen [144, 145]. In contrast to field observation, which describes what happens, ethnography focuses also on why and how things happen. While traditional ethnography is based on long-term studies, the adapted method compensates the much shorter time frames with (a) a more focused observation scheme and (b) an interview at the end of the study that is used to clarify issues that arise from a preliminary analysis of the data.

Eight nurses and physicians from a stroke unit were equipped with wearable electrocardiography (ECG) and acceleration sensors during their work day to (a) make them aware of arousal by looking at their data and (b) support the recalling of experiences to identify stressors. The ekgMove was worn by nurses and physicians during at least two consecutive shifts. The shift length varied between 8 and 24 hours. The ekgMove recorded movement at the chest and the ECG signal. After each shift, participants recorded their experienced stress levels on a 5-point Likert scale to collect further reference data. They were asked to assign a rating to each hour of their shifts.

Observers followed a selection of participants (three nurses and two physicians, one male and four female) and acted as ethnographers collecting data about: (a) situations that could be conceived as stressful, (b) physical activity that could interfere with the measurements, (c) articulations by participants, and (d) interaction with other staff, residents, and relatives. Each of the three ethnographers was given a doctor's overall to blend into the setting and followed a participant during at least one entire shift. The annotations were made in a traditional notebook, which facilitated the skill to take notes anywhere at anytime, and had a level of detail of about 1 minute.

The recorded data were shown to participants 1 week after the study in semi-structured interviews with their ethnographer. The interviews were planned for 1 hour. Participants first reviewed their ECGs and activity data using the UnisensViewer [142]. The heart rate, activity level, and calculated features, such as step counts, were shown in a timeline view of each shift. The visualized data were only a subset of the recorded data, which were selected to stimulate discussions. For instance, one day with many emergencies was shown and then a very quiet day. Participants were asked to reflect on the visualized days and talk about specific events that came to mind when reviewing the data. Visible changes in the heart rate or activity especially should be explained. A printed report was distributed to participants to support the analysis. The report contained a heart rate histogram, Poincaré plot of HRV, average heart rate per hour, and activity in steps per hour. The interview concluded with a discussion of general questions, such as convenience of sensor usage and the expected benefit from using such a system.

The eight participants from the stroke (four nurses and four physicians) came from all age groups (22-44), including men and women (3:5) who had different levels of experience (1.5-25 years).

5.2.2 The ECG and Activity Data

The sensors recorded 152 hours of ECG and acceleration data. Ethnographers observed 49 hours and annotated with physical activity and important events.

Activity energy and step count were calculated from the activity data. The results show continuous high activity. According to observations, nurses and physicians walked between patients and offices. Figure 5.4 shows the average step count for nurses and physicians during the early and late shifts. Nurses were nearly constantly walking from bed to bed, while physicians spent more time at their desks documenting. However, the observed potentially stressful situations took place while with a patient or on the main corridor. Observations indicate that many more activities cannot be recorded by the used sensor. For instance, nurses lifted patients out of bed or moved a patient's bed while walking.

An analysis of the data using standard HRV features such as SDNN, RMSSD, and SDSD (see Section 2.3.1) were noisy because of the constant physical activity of nurses and physicians. The rare times that participants were not physically active were during documentation tasks or breaks. In these timespans, arousal events were neither registered by the ethnographers nor described in the interviews. An attempt to use

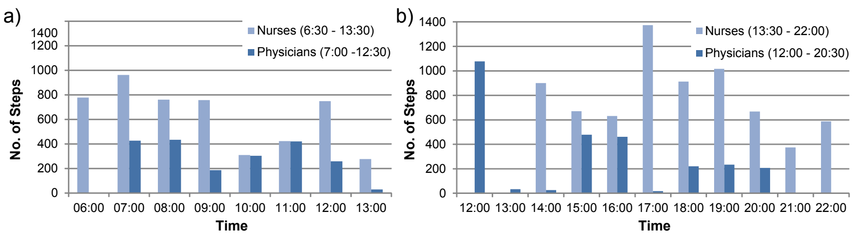


Figure 5.4: Step count of employees: High activity levels hide the cognitive effort. Nurses are walking more than physicians and have fewer breaks for documentation. The figure shows the number of steps for a physician and a nurse during each hour of (a) first and (b) second shift.

the existing additional heart rate algorithm by [75] was stopped. The constant walking at varying speeds, carrying equipment, or lifting patients out of bed resulted in unpredictable rises in heart rate that hid potential arousal-related reactions of the heart rate. As a result, the activity levels were often higher than allowed by the algorithm and no arousal event could be recognized.

Therefore, unprocessed heart rate data of whole shifts were shown in the interviews. Nurses and physicians on a stroke unit are used to seeing these kinds of data. They interpret ECG data every day and can spot significant changes even in noisy signals.

5.2.3 Relevance of Data

According to the observation protocol, the monitored days were rather quiet because there were few emergencies. The average rating in the daily stress questionnaire was 2.38 on a 5-point Likert scale (from 1 being not very stressful to 5 being very stressful). A single participant marked 3 hours as stressful (4), but was the only one to report stress. In the interviews, participants confirmed that the days we monitored were rather quiet with only a few emergencies:

> D1: *"It would have been more interesting for me if it hadn't been so quiet. I was waiting for an emergency to come, but nothing happened."*

Strong reactions, however, can be recognized despite physical activity. The heart rate data contained 15 events that showed an increase in heart rate that could not be explained by the movement data. Seven of these events correlate with observations of potentially stressful situations. These 15 events represent only a small subset of the reactions that were observed. Furthermore, these strong reactions often trigger physical activity (e.g., jumping back or running to help in an emergency).

Figure 5.5 shows an 8-hour shift of a nurse that contains a peak in heart rate and activity 3 hours and 20 minutes after the start of the recording. When these data were shown to the nurse, she immediately remembered this event and requested to see more details, which are shown in Figure 5.6. The detailed view provides an analysis of the event. A patient's heavy muscle spasms surprised the nurse, who was working for a second week on

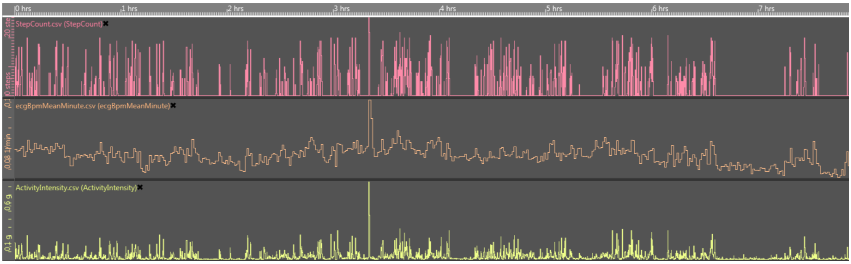


Figure 5.5: Screenshot of the ECG and acceleration data as displayed in the UnisensViewer. The first row shows the number of steps, the second row the heart rate, and the bottom row the general activity. Clearly visible is a sudden peak of the heart rate after 3 hours and 20 minutes.

this station. She sprinted to alarm a physician and returned immediately to the bed. As a nurse, she is not allowed to medicate patients without a physician's permission. Nevertheless, she was about to inject the required medication to help the patient without waiting for the physician. Other more experienced nurses came to calm her down and told her to wait.

The overall event is visible at first glance at the data. However, these extreme events are rare and a stress reaction is normal, although heart rate changes that are caused by the arousal reaction and the physical activity overlap in this example. The recorded data can help physicians and nurses to remember such events and to later reflect on their behavior. Minor stressful events (e.g., a heated discussion while walking) are difficult to capture by using the heart rate of a physically active person.

Two different visualizations were offered in the concluding interview: a printed summary of their day and a zoomable timeline view in the UnisensViewer. The report depicted heart rate and activity data aggregated on an hourly basis. The UnisensViewer is designed for researchers and allows them to explore psychophysiological data along the time dimensions. Users can zoom in and out to inspect data and compare multiple signals at a selected time span. All but one participant preferred the UnisensViewer because they could inspect details and discern the impact of specific events.

> D4: *"Amazing, it is easy to understand."*

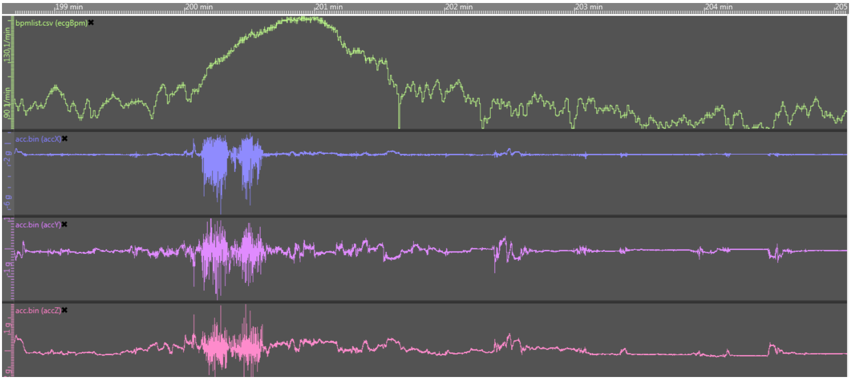


Figure 5.6: Details of the reaction to a sudden emergency: the heart rate shown in the first row rises from 90 beats per minute to 155. This is mainly caused by the intense physical activity shown in the lower three rows. Close examination of the three activity curves shows the two sprints and a very short stop in between.

> N2: *"I like the UnisensViewer more than the graphics. I can see everything that happened there and make a guess."*
> N1: *"Maybe UnisensViewer, then I can exactly see when, what time, something happened. .. With the graphics I can't see, when a seizure occurs, for example."*

However, nurses and physicians are used to reading heart rate diagrams. Different user groups might struggle to understand the amount of data.

5.2.4 Reflection Support

In the interview, participants confirmed that dealing with stress is an important reason to use sensors. Measuring their own physiological data at work was interesting for all of them and the participants expressed their interest in recalling what happened during their work days. Most of them stated that this interest is much higher with respect to their stressful days and that they would like to compare what the measures look like on different days.

N1: *"Yes, it would interest me, especially when I had stress or emergencies."*
D1: *"How often I would use it I can't tell you ... If I had a 24 hours shift with 10 admissions with reanimations"*

The review of the data led to surprising insights because the majority of employees stated that the monitored time span was rather quiet. For instance one nurse said:

N2: *"I thought I was calm but now I see it wasn't like that ...*"

In general, the timeline graphs of heart rate and activity data acted as memory aids. The structure of the day became especially clear to participants. They re-interpreted the data according to their own experiences as defined in the model by Boud [4]. They came up with narrations of the past day and articulated insights and narratives.

N4: *"Yes! I can remember the two patients. They annoyed me"*
N1: "*Yes [it helps me to remember]. I can say approximately when some things happened.*"
D1: *"Yes, it was interesting [the support of the sensors to remember]. It was interesting to see it graphically."*

However, participants did not feel they were able to act on their awareness. They see stress and stressful situations as part of their job.

D1: *"We have to hurry up. On duty you can't do anything against it. What could I do better? You don't think. You are there, and you have to do it."*

Moreover, the reflection of challenging situations is seen critically because it conflicts with existing coping strategies. Nurses are told during their training to strictly separate their private lives from the stress they experience at work. They are told to leave these experiences behind in the locker room with their work clothes when they go home after their shifts because they should not ruminate on their patients problems. Asking them to reflect on their shift and the emotionally challenging situations conflicts with this trained coping strategy. One nurse explicitly argued against reflection to stay focused on the task at hand.

Participants with more work experiences had more developed coping strategies. They also committed to the ideal of "professional' distance," but had developed their own strategies to achieve it. Although some nurses tried to avoid and suppress emotions and emotional reactions during work, others elaborated on concepts of emotion regulation to find the right balance between "cold" and "over engaged."

5.2.5 Potential for Long-Term Usage

Nurses and physician confirmed that the collected data support reflection and can lead to new insights. They liked reviewing their data to get a different perspective on their daily work. They knew this perspective from monitoring their patients, but it was their first time measuring their own heart rate and activity. Activity was seen as more interesting because their heart rates were noisy from the high activity levels. As one participant said:

> D4: *"I don't like staying in hospitals and going to the doctor. I am not the type of person keen on trying new things out but it was actually interesting for me. I would mainly like to know about activity and movement."*

However, participants did not see a potential for long-term usage because the usability of sensors is still insufficient and reflection on stress contradicts with their current coping strategies and training. The sensor belt was used without complaints during the study. However, during the interviews, participants complained that the belt was itchy after some time or that electrodes would stick to their skin and hurt when they got disconnected because of body movement. One participant described the belt as

> D3: *"a badly fitting bra that is a little bit inconvenient but still wearable."*

The insights they gained were limited. They understood that they could know that they were stressed, but could not identify clear reasons. Furthermore, if reasons like an emergency could be identified, participants saw no option to react in a different manner. As a result, they did not know how to act on the insight they gained by reviewing the data. Reflection has to provide clear benefits to ignoring stress at work. These benefits are not clear to participants and each participant expected different benefits that should be considered in the design.

5.2.6 Implications for Design

The ethnographic study led to three major insights: (1) psychophysiological data can facilitate reflection in healthcare professions, (2) sensors and algorithms have to be adapted to the high activity levels, and (3) learners need more context to identify the reasons and patterns behind an emotional reaction.

Nurses and physicians were interested in the data and liked reviewing the data. They remembered their day and found surprising psychophysiological reactions that triggered reflection. Nonetheless, participants did not come to an observable outcome, such as a change in behavior. They did not know how to act on their insights because these insights remained at an abstract level. The levels of acceptance and reflection varied among participants. The most obvious difference was observed between participants with different amounts of work experience.

Similar to Sanches et al. [146], we came to the conclusion that *"it is difficult, sometimes impossible, to make a robust analysis of stress symptoms based on biosensors worn outside the laboratory environment."* The used sensors delivered data at a sufficient quality, but the high activity levels diminished the insights that could be drawn from the data. The HRV algorithms that have been designed for lab settings could not be used. Moreover, the changes in activity types were too rapid for the additional heart rate algorithms. Within the 3 minutes of the sliding window, a nurse had often visited a patient, picked something up, and carried it back. In addition, a variety of heart rate relevant activity could not be recorded at all by the system. For instance, there is a significant difference in the exertion needed to bend over a bed to lift somebody out of the bed or to just check the blood pressure.

The design of algorithms and bio-sensors for work settings must account for the specific challenges at each workplace. An iterative design process using multiple sensors and frequent testing in the target domain is a promising approach to deal with the unpredictability of the workplace setting. xAffect [206] can be a possible base for developing such a framework.

5.3 Mobile xAffect

The following section describes the design and implementation of a rapid prototyping framework for psychophysiological applications under Android. The system is based on xAffect (see Section 3.2.1). The required changes to port xAffect to the Android platform are described. The mobile version of xAffect was used in two supervised theses to create two apps: the Telco App [182] and the Posture App [190]. Both apps record psychophysiological data in telephone conferences which was seen as a promising use case. Participants of telephone conferences often experience arousal, whereas their physical activity is low. These attributes simplify measurements and reduce error sources. The section concludes by describing insights regarding the framework and required changes.

5.3.1 Mobile Psychophysiological Sensing Framework

xAffect is unusual among the available middlewares to analyze psychophysiological data because it is based on Java. Other middlewares [147] are implemented in C++, Matlab, or Python which are programming languages that promise a higher performance to process the recorded data. Java, on the contrary, focuses on a clear, reusable structure of code that can be used across platforms. An especially interesting platform is Android, a operating system for mobile devices. Mobile phones and tablets have become ubiquitous and their processors have reached a performance that is already sufficient for speech recognition. Hence, they are a promising platform for psychophysiological recording and data analysis. In addition, Android currently has the largest market share among smartphones [148].

As explained in Section 3.2.1, xAffect is a modular system that connects three types of components: sources, processors, and sinks. The xAffect core library could be used under Android with minimal changes. For example, available Android versions at the time of development supported only Java 1.6, whereas xAffect was written and compiled with Java 1.7. Only a few code modifications were necessary, which did not affect functionality on Android or Windows. Consequently, the Android and the desktop version can share a common core library. Furthermore, many of the desktop components can be reused under Android, such as signal generators. Components like the UnisensReader and UnisensWriter had to be adapted to the different file system.

Sources that connect to hardware sensors could not be reused because they communicated over a serial interface that is not available on a phone. Hardware that was connected over wireless technologies, such as Bluetooth, could be adapted with minor modification. The MovisensSource that connects to the Movisens ekgMove [56] was adapted to the Android Bluetooth stack. However, while Android phones ship with the same Bluetooth API, the actual Bluetooth implementation differs between manufacturers. An alternative to Bluetooth is described in Section 5.3.2.

Current mobile phones already integrate a number of sensors. In the context of a term paper [189], the available sensors were integrated as sources into xAffect. Figure 5.7 shows the developed example setup that integrates the new available sources. Acceleration and gyroscope can be used to estimate the current physical activity. The magnetic field sensor and the GPS provide information about the current location and direction. The proximity sensor should not be confused with the proximity sensors in Chapter 6. Proximity sensors in a phone are infrared sensors that measure if the phone is currently next to the head (i.e., if the phone is used to make a call). These five sensors and the temperature sensor can be found in nearly all Android phones and can be conveniently accessed

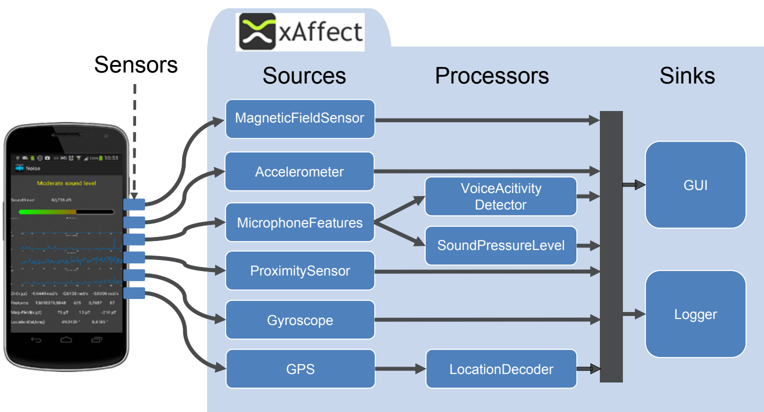


Figure 5.7: Example setup xAffect mobile

by APIs. In addition, these sensors can be used and combined in different ways. For example, if the location of the phone on the body is known, the acceleration sensor can indicate the body posture.

The microphone can be used as a sensor in a variety of applications. For privacy reasons and because of the constrained memory, the recorded audio data are not stored, but are directly processed into a set of features that can be further analyzed. Chang et al. [11] have presented a library that can detect emotions and stress from voice data on mobile phones. Despite an announcement, this library is not available as open source. Inspired by this work, an alternative processor was developed to detect speech activity among other noise.

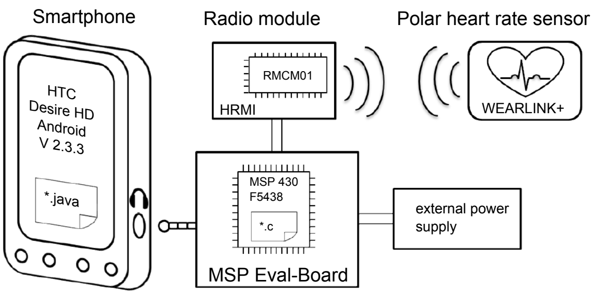
Moreover, xAffect mobile can be used to implement further sources that do not require an additional sensor, but reuse built-in functionality of the phone. The approach explored with this app is also referred to as sensor-less sensing [149]. In fact, sensors are still needed, but no additional sensors have to be worn. Existing sensors are reused or input devices are used as sensors. For instance, the mouse movements [150] or keyboard presses [151] have been analyzed and found to correlate with the arousal level. The upcoming smart watches add more sensors and can help to recognize even more behavior. For example, the Android Wear [152] smartwatch comes with an integrated optical heart rate sensor.

5.3.2 Connecting External Sensors via the Audio Jack

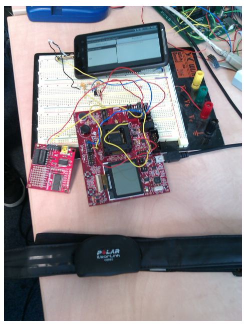
Android already provides a limited number of sensors. A much wider variety of sensors comes as separate devices. For instance, psychophysiological sensors are often separate devices that must be connected to the Android system. A growing number of these sensors implement Bluetooth interfaces, which is integrated on current Android phones as well. The Movisens ekgMove [56] can be connected to the mobile xAffect using Bluetooth. However, there are many more protocols that are used by wearable devices, but are not supported by Android. Proprietary protocols, such as the WearLink protocol from Polar [14], require dedicated radio hardware. If xAffect could connect to such interfaces by a simple hardware extension, a whole range of commercial low-cost sensors could be used. Commercial sensors like the Polar heart rate monitor do not deliver the same data quality as scientific or medical devices, but are much cheaper and are available in large quantities. Simple prototyping applications, such as a heart rate monitor that uses the additional heart rate algorithm, could be quickly implemented.

The standard hardware interfaces of Android phones were analyzed to connect additional hardware modules. The USB interface is present on all phones and has been used. However, it requires a custom firmware and an external power supply. Finally, the audio interface was selected because (a) it is available on all Android phones, (b) access to the audio interface is directly possible from the Android API, and (c) no custom firmware images have to be installed. A hardware system that can communicate over the audio interface can be simply plugged in and can communicate with any app. The audio interface consists of four connections that can be used: microphone, left speaker, right speaker, and ground. The HiJack project [153] had already demonstrated on the iPhone that these four connections can be used to establish a communication and power an external device. Two connections are used as a bidirectional communication channel. The remaining two connections power the hardware.

In a supervised Bachelor thesis [179], a first prototype was developed that is shown in Figure 5.8. A MSP430F5438 was used as the basis of the prototype. The connection to the Polar heart rate chest belt [14] was realized using the RMCM01 chip that implements the Wearlink protocol.



(a) Architecture



(b) Prototype

Figure 5.8: Connecting sensors via the audio jack

Two of the four audio jack connections are used for serial data transmission based on the UART protocol. The audio channel transmits an AC signal that is interpreted by the hardware extension. The data are transmitted using Manchester coding. The developed prototype supports bit rates up to 9600 bits/s. Alternative line code could further increase the available bandwidth.

The prototype is operated by an external power supply. First tests were conducted to estimate the feasibility of powering the hardware over the audio jack under Android, as done for an iPhone in [153]. The test showed that the HTC Desire HD can deliver a maximum of 25mW with a connected resistor of 3Ω. However, a test with a different Android device, the Motorola Milestone 2, resulted in completely different values. The maximum of 10.43mW was reached with 15Ω. These differences indicate that a more complex setup is required to power the hardware with different Android phones.

5.3.3 Telco App

The Telco app is a tablet application that measures the ECG signal and shares the calculated arousal during telephone conferences. The system is based on the ekgMove and existing HRV algorithms to calculate the arousal of the user. The resulting arousal value is displayed in relation to the arousal of all participants. Hence, the calculated arousal values have to be transmitted over the Internet.

Figure 5.9 shows the resulting xAffect setup. ECG data are recorded using the ekgMove from Movisens [56] and by using the corresponding xAffect source. The open source OSEA algorithm [154] detects R-Peaks in the ECG signal, which are the basis to calculate the heart rate. The OSEA algorithm was available as C-Code. After recompiling the code under Android, it could be embedded in xAffect using JNI. The resulting list of R-Peaks serves as input for the HRV calculation. The HRV processor transforms 5-minute windows of the data into the spectral domain using the Fast Fourier Transform (FFT). The output of the processor is the ratio between low frequency (LF) and high frequency (HF) components in the spectrum (see Section 2.3.1). The calculated ratio is displayed in the graphical user interface, logged into the Unisens format for later analysis, and shared over the Internet using the MIRROR spaces sink [155]. The

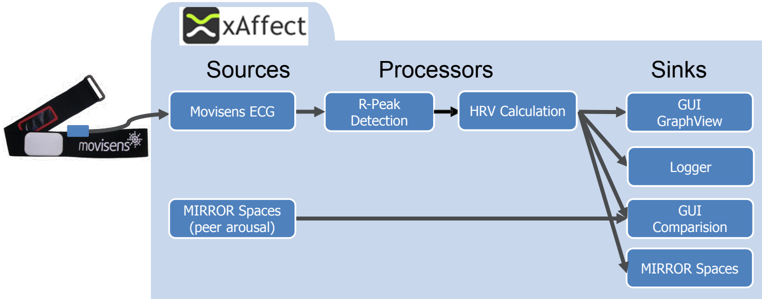


Figure 5.9: xAffect setup for the Telco app

corresponding MIRROR spaces source receives the LF/HF ratio of other participants and displays an average in the GUI.

5.3.4 Posture App

The posture app analyzes the posture of participants during telephone conferences by using two acceleration sensors: one sensor is worn on the chest and the second is the smartphone sensor. The smartphone was assumed to be carried in the trouser pocket at the hip. The goal of the app was to not only recognize the posture, but to also affect related signals. Figure 5.10 shows the resulting xAffect setup and the developed components. Two independent sources gather data from the acceleration sensors. Both sensors deliver signals with a variable sample rate and have to be resampled to a common constant sample rate by the resampling processor. The resampled signal was analyzed by a self-organizing map [156] and the results of the analysis were visualized in the GUI and stored in a file.

The required resampling was a challenge for xAffect in this use case because xAffect assumes a fixed sample rate for high bandwidth data. In this particular use case, the acceleration signal from the ekgMove is received with a varying delay because of the Bluetooth radio transmission. Moreover, Bluetooth packets and the contained data might be lost. The sample rate of the Android sensor was configured to 25 Hz, but on a Samsung tablet, the resulting sampling rate was 25.5 Hz. The same

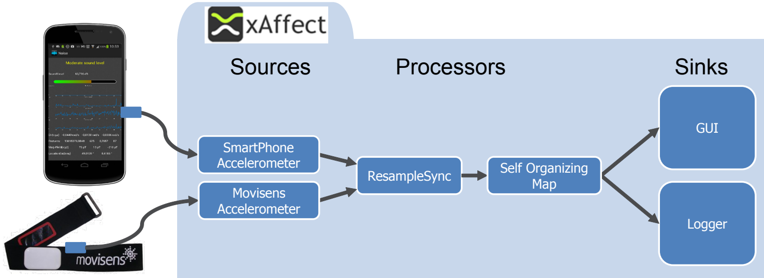


Figure 5.10: xAffect setup for Posture app

configuration on a smartphone resulted in exactly the desired 25 Hz. The ekgMove was configured to send acceleration values at 64 Hz. The resulting sampling rate of the data stream varied between 64 and 65 Hz. A processor was developed that resamples and synchronizes the two incoming data streams. A buffer and an additional look-ahead buffer are used to collect data and inspect whether the current sample rate is too high or too low. Accordingly, samples are deleted or interpolated. To minimize the impact on the data, interpolated or deleted samples are evenly distributed across the buffer (e.g., if only one sample is missing, it is added in the middle of the buffer). A 10-second buffer and 2-second look-ahead buffer have been used. Hence, the processor will result in a 12-second delay of data processing.

The self-organizing map used 11 features that are listed in Table 5.1. The features are sorted by the corresponding physiological signal and have been selected by literature review and exploratory data analysis. The posture features use mainly the data from the chest strap. They try to identify if a person is leaning forward, backward, or sideways. The angle between the two sensors is a good indicator of the current posture. Energy and consistency measures have been good features in the sociometric badge [86]. They should, for example, recognize if a user is nervous and is constantly moving. Finally, the breathing rate was estimated by analyzing the movement of the chest sensor. This feature cannot be calculated during general body movement and is held constant in these moments.

The self-organizing map (SOM) [156] was the first artificial intelligence

Physiological Signal	Feature	Sensor
posture	posture angle	both sensors
	mean x acceleration	chest
	mean y acceleration	chest
	correlation between x and y acceleration	chest
	correlation between x and z acceleration	chest
	correlation between y and z acceleration	chest
	mean x acceleration	chest
activity	movement energy	chest
	movement energy	smartphone
	consistency of energy	chest
	consistency of energy	smartphone
breathing rate	filtered chest movement	chest

Table 5.1: Analyzed features in the Posture app

algorithm that was implemented in xAffect. The implementation had to overcome three main challenges. First, clustering and classification algorithms usually work on features instead of a raw data signal. However, xAffect is engineered around the concept of time-oriented data, as defined by the Unisens standard [142]. Multidimensional data, as required for clustering and classification algorithms, could only be implemented as Unisens data signals with multiple channels. Each frequency of a calculated spectrum would be treated as a channel. This possibility was not obvious and, therefore, was not used. As a result, the feature calculation was integrated into the SOM implementation. The implementation was validated against the Matlab SOM implementation.

The second challenge was the required training data. Until now, the *train* method was used to calibrate thresholds and fill buffers. This method is called directly before the classification starts and is designed to calibrate algorithms to the incoming data (e.g., to fill a buffer with ECG data to start an FFT directly at the beginning of the data recording). However, clustering and classification algorithms need large amounts of training data. While these data could be streamed from a file using the UnisensReader

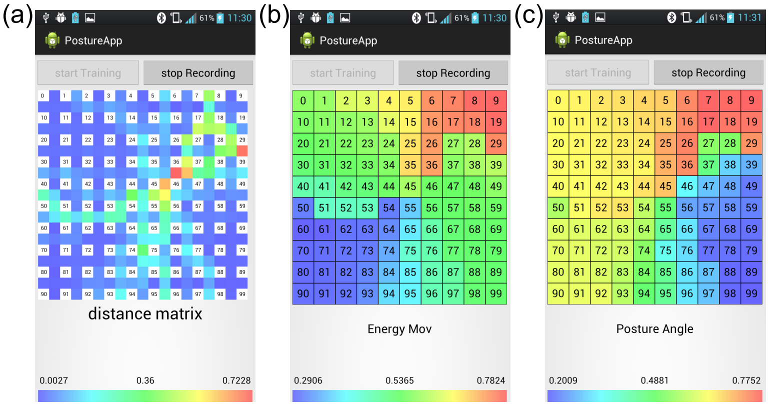


Figure 5.11: xAffect setup for Posture app

Figure 5.12: Implemented visualizations of SOM: (a) distance matrix, (b) energy distribution of movement data in SOM, and (c) distribution of posture angles in SOM

component, the data would be streamed at the normal processing speed (e.g., loading two hours of acceleration data would take two hours). Instead, the training data are read directly from a file.

The last challenge was the visualization of the clustering data due to the architecture of Android and the specifics of the desired graphic depicted in Figure 5.12. The resulting multi-dimensional output of the SOM was again too complex for the xAffect data structures. Therefore, the visualization was part of the SOM processor. The visualizations could not reuse existing charting libraries because of the underlying data structures. Moreover, Android prevents background threads from directly accessing, and thereby blocking, the user interface thread. A so-called handler has to be used to modify the user interface.

5.3.5 xAffect Redesign

The first prototypes revealed the limitations and some misleading concepts of xAffect 1.0. The developed components and especially processors were not reusable as intended. Processors became highly complex and specialized. For instance, the developed SOM processor [190] integrated feature extraction, SOM algorithm and visualization. Each of the calculated features would have been useful as a separate processor, but the supported data formats were not fully understood. Reusable small components were the key design goal of xAffect. Developers should be enabled to rapidly build prototype by selecting from a growing amount of components. Instead both prototypes reimplemented a Fast Fourier Transformation and struggled to implement a data viewer.

A redesign was required to make xAffect easier to use and the prototype faster on Android. The redesign intended to guarantee backward compatibility by making only minor changes to the xAffect core and developing optional components that show how standard problems on Android can be solved. The DataDescription was extended by an optional ID field. The DataDescription is normally used to match outputs and inputs of components automatically. In the PostureApp, two sources deliver acceleration data of different types. To distinguish between them, the developers misused the content class field (e.g., by using SMARTPHONE_ACC instead of ACC). This leads to components that accept only these custom content classes and, hence, cannot be reused. The ID field allows differentiation between data streams of the same content class from multiple sources. The same problem would also occur with network sources that receive data streams from multiple clients.

The new optional components that come with the mobile version include filter processors, a data viewer sink, and sources for the Android sensors. While the Android desktop version comes with a data viewer, Android developers had to build their own data viewer components. However, desired visualizations can be fairly complex and specific to the use case. Therefore, the developed AndroidDataViewer offers a number of interfaces that can be extended and a small set of default visualizations that can be used for early prototypes. These default visualizations include a graph viewer based on the GraphView library [157]. The underlying interfaces and sink implementation realize the communication between xAffect and the user interface. In addition, a set of filters was developed because filtering noise or analyzing specific spectral parts of a signal are common tasks. The filters include a Butterworth and a Wavelet filter that can be parameterized to the desired frequencies. These components are not part of the xAffect core, but are available as options. The developed filters are platform independent and can be used in Android as well as on the desktop.

5.4 Summary

Table 5.2 provides an overview about the developed candidates and evaluations that explored the design space for the capturing of affective context. A first test showed that self-reporting approaches require too much effort to be used in the healthcare domain. Therefore, the majority of the applications are based on sensor technologies. The ethnographic study created a broad understanding of the challenges of sensor-based approaches in a stroke unit. As a consequence, xAffect was extended to support the rapid prototyping of mobile capturing apps. Two prototypes, Telco app and Posture app, evaluated the developed system and triggered a redesign of the mobile xAffect. Although the developed prototypes have not be evaluated in the field, they provide a number of insights for the capturing of affective context in challenging environments.

At this point, no sensor solution could be identified that can directly measure arousal in environments with highly variable activity levels. Although specific situations and settings with lower activity levels can be

Candidate	**Data**	**Technology**	**Evaluation**
MoodMap	moods	self-reporting web application	project meeting
Movisens ekgMove	ECG and activity	existing sensor	ethnographic study on stroke unit
Telco App	ECG	prototype based on xAffect	formative
Posture App	posture and body movement	prototype based on xAffect	formative

Table 5.2: Capturing prototypes for affective context

targeted, multi-sensor systems may help to overcome the observed challenges. Multiple sensors and algorithms must be evaluated. The first captured data and early feedback are crucial to adapt quickly to domain-specific requirements. xAffect mobile reduces the time to a first prototype that can be used to gather the required data. Algorithms and sensors can be quickly exchanged. The component-based architecture facilitates the reuse of existing components. Moreover, xAffect already integrates all Android sensors and a growing number of external sensors are available. In addition, an approach to connect proprietary hardware via the audio jack has been implemented.

Furthermore, general insights regarding the usage of affective context have emerged. The ethnographic study involves measuring arousal as a negative element that should be avoided. However, the knowledge about one's own arousal levels is useful for reflection only if stressors and reoccurring patterns can be identified. Otherwise, employees are not able to act on this knowledge. In this case, knowledge about stress is experienced as an additional burden. The data must be analyzed in light of the conducted tasks and situations that led to an emotional reaction.

In the interviews, nurses and physicians referred to patients, colleagues, and residents. Their narratives were structured by these social contacts and the contacts mirrored their daily work. The observation protocols confirm that the daily work of physicians and nurses can be summed up as series of social contacts with patients, colleagues, and patients' relatives. Data on these contacts would allow reflection on work patterns and, therefore, may provide the needed clues to identify stressors. An approach to capture these contacts is outlined in Chapter 6.

6 Design Study II: Capturing Social Context

Social contacts are an essential part of daily work. Especially in health professions (e.g., carers in a care home), social contacts with patients, colleagues, and relatives define the workday. Capturing these contacts could help to quantify daily practices and provide an objective perspective on work processes in a care home. How much time is needed to care for a resident? Is this effort increasing over time? These types of questions can be answered by reflecting on quantitative data about the interaction between residents and carers in a care home. Hence, this design study focuses on capturing social contacts in care homes by means of sensors. The work described within this chapter has been published previously in different forms in [194, 202, 198].

Dementia is currently a pressing issue in care homes [158]. Therefore, we analyzed the options to capture the social contacts in dementia care. The selected solution builds on new wearable proximity sensors that record the co-location of carers and care home residents. We developed two types of prototypes, which have been iteratively refined according to evaluations in the target domain, as shown in Figure 6.1. The proximity sensor prototype described in Section 6.2 was used to measure acceptance and assess the feasibility of the selected capturing approach. Building on the successful

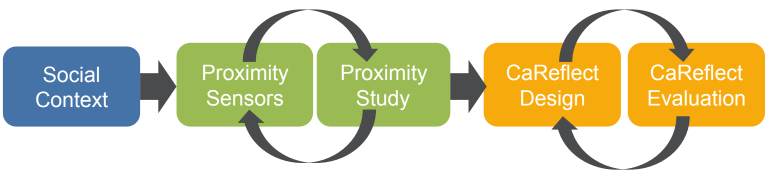


Figure 6.1: Design process social context capturing

evaluation in Section 6.3, we developed CaReflect–an application that visualizes the collected data and manages the study process as described in Section 6.4. After two iterations, the developed application was successfully evaluated in a care home, as described in Section 6.6.

6.1 Measuring Dementia Care

Dementia is an age-related illness that affects the cognitive abilities of mainly elderly patients. People with dementia lose temporal and spatial orientation and, therefore, experience reality differently. In many cases, a dementia patient's different realities are hard to understand and do not match our reality. Hence, people with dementia react in unexpected ways, including challenging behavior, depression, or apathy [159]. Carers who want to react properly must learn the resident's biography and see the world from their perspective. Reflective practice is seen as a promising approach to acquire this knowledge.

Observational frameworks, such as *dementia care mapping* [160] and the *short observational framework for inspection 2* (SOFI 2) [161], are used to ensure care quality. Both approaches are based on trained observers who visit a care home for 2 to 5 days. SOFI 2 is currently used by the Commission for Social Care Inspection in the UK to *"capture, in a systematic way, the experience of care for people who use services who would otherwise be unable to communicate this to an inspector"* [162]. The observers confirmed that care staff will try to improve their behavior during this period. However, they also notice that care staff will quickly default to their normal care practices [160].

The data collected by the observational frameworks provides detailed feedback and opportunities for improvement. However, the costs are high and the quality of the feedback depends on the observer. Automatic capturing approaches, such as sensors, could deliver quantitative feedback on care quality at lower costs. Hence, they could be used more often to induce and facilitate a continuous improvement of care practices. Moreover, a continuous monitoring is possible as well as monitoring on the weekend when staff levels are often lower. However, to be applicable, such a system has to be easy to deploy and use in a care home.

6.1.1 Dementia Care in the UK

Caring for people with dementia is challenging. Care staff have to deal with unexpected reactions and behavior (e.g., residents that become suddenly aggressive or try to leave the care home). Social contact and engaging activities are important to maintain cognitive skills as long as possible. However, care funding is limited. Therefore, new care approaches like the “Butterfly method” are recommended to provide many short social contacts to people with dementia instead of only a few prolonged interactions [163]. Furthermore, care has to be personalized. To react correctly to challenging behavior, carers must know the details of the biography of the person suffering from dementia (e.g., who the person is that the resident is talking about). This knowledge can help to stimulate them or calm them down, depending on the situation.

The majority of care staff have received no formal training, but learn the profession during care. They are supervised by experienced carers and registered nurses who take up medical tasks. In most homes, carers and nurses work a three-shift system. The numbers of carers and nurses vary according to the shift and care home size. A common pattern is that there are only a few nurses, but one nurse has to be on each shift in the care home. The low-education level of care staff and the challenging requirements lead to a wide variety in the quality of the provided care. In addition, these challenges are the reason for a staff turnover of more than 17.3 percent in 2013 [137], which again affects care quality. Hence, it is likely that inexperienced carers will have to deal with the unexpected behavior of dementia patients. To provide high-quality care, these carers must learn about dementia care and about the residents as quickly as possible.

Care homes in the UK are often not purpose-built, but converted from former usage. The floor layout in the three visited care homes is complex because existing houses have been connected and integrated with new buildings. The IT infrastructure is often limited to a few computers and laptops that are used for documentation. The computers are either in the office or used in the common rooms to document during care. WiFi is not available and would be expensive because of the architecture of the care homes. Furthermore, care homes are skeptical regarding technology and want to focus on their work with residents. The value of IT infrastructure to support care is often doubted, although this is currently about to change [194].

The deployment of IT infrastructure in care homes is a challenge that should be avoided, if possible. In general, IT usage and expertise are low. Staff in care homes want to focus on care and do not want to spend much time with technology or perform maintenance tasks. In fact, we observed that carers delegated documentation tasks to younger colleagues to avoid using the computer. In summary, the requirements are:

- The system should neither make assumptions on IT infrastructure nor require the prior installation of such.
- The system should be easy to embed in the care process and minimize the time spent with technology.
- Maintenance tasks, such as changing batteries and configuring devices, should be simplified and, if possible, avoided.

These high-level requirements translate into challenging technical requirements for the final solution.

6.1.2 Sensor-Based Care Measurement

Sensors can measure the co-location of care staff and residents, which can be interpreted as care activities. In initial discussions, carers suggested that care should be registered, if the distance between carer and resident is less than arm's length. However, observations showed that carers rarely have time to sit down for a longer time with residents. Instead they move between residents and, even when caring for a resident, frequently leave the immediate proximity to pick up materials. As a consequence, all contacts within an approximate 2-meter distance have been defined as care activity.

The distance between two actors can be calculated either by localizing all carers and residents in the care home or by measuring the distance between two actors directly. Both approaches rely on radio technologies to estimate the distance between actors or locations. A full localization requires beacons or access points that cover the relevant locations and wearable or mobile devices for all relevant actors. Measuring distances directly, however, requires no beacons or access points, just wearable or mobile devices for all relevant actors. In the resulting point-to-point network, co-location can be calculated without additional access points. Therefore, the sensors are the only infrastructure. They are mobile and can be rapidly deployed in any care home.

Several systems have been developed to measure co-location and analyze contacts (see Section 3.2.2). Most approaches examined the resulting structure of the social network graph with social network analysis (SNA) [118]. Fully distributed systems that measure co-location are made up of small, wearable, battery-powered devices. To be used in a care home, the wearable devices must:

- Be as small as possible to minimize interference with care routines. Carers should be able to wear the sensors in an unobtrusive manner. Residents should be equipped with a small device because dementia patients may become nervous or confused, if they notice an unknown item on them.
- Be as energy efficient as possible, because (a) changing batteries may put an extra challenge on care staff, which may result in nonfunctioning sensors and (b) smaller sensors require smaller batteries (i.e., coin cells). The optimal system would run on a single coin cell for at least 5 days, which is the maximum time span of SOFI observations.
- Use an existing hardware because multiple sensors with stable hardware are required. In a research context, it would be challenging to produce large quantities of sensors in consistent quality.
- Measure all contacts that last longer than 10 seconds. This value was agreed upon by care staff to minimize contacts that are recorded when carers pass by residents; this increases the battery life of a sensor.
- Be sufficiently affordable because a sensor for each carer and resident is required. Moreover, it is very likely that sensors will be lost, washed, or require replacement for other reasons. Care staff should not be concerned about losing sensors, but should use them as available tools.

As described in Section 3.2.2, none of the developed systems is applicable or has been applied in this scenario. Therefore, a new solution was developed.

6.2 Proximity Sensor Prototype

The proximity sensors are based on a programmable wristwatch with a radio module. A custom firmware was developed and installed on all watches to turn them into proximity sensors. The design and implementation of the sensor firmware was intended to maximize the battery life. Towards this end, a low-power mechanism to detect proximity was developed. An asynchronous low-power listening protocol reduces the required radio communication and increases the time a sensor can sleep. The initial firmware version was developed in a Bachelor's thesis [185]. The developed firmware maximizes the lifetime of the system by (a) minimizing energy consumption according to the particular platform and (b) minimizing the required memory to store as many contacts as possible.

6.2.1 Hardware

The programmable Chronos eZ430 wristwatch (see Section 2.2.4) was chosen as the hardware platform because it is small, includes a low-power radio module, and can be programmed. The sensors must be small and lightweight to be wearable during daily work. The battery will make up the main share to the overall weight, because the sensor consists of only a few components. A smaller battery reduces weight, but also reduces the battery life. This balance poses a challenge to the power consumption of the sensor. The proximity detection must be implemented by the firmware in a way that minimizes power consumption.

We developed alternative badge formats that evolved as shown in Figure 6.2. The initial watch format did not comply with the regulations in the care home. Carers are not allowed to wear jewelry because it could hurt a resident's skin, which is fragile and heals slowly. The third and final format of the sensor prevents access to all buttons because residents should not be able to stop the recording by accident. Furthermore, the system is splash-proof. In one of the studies, a sensor was washed by accident, but continued to work. Consequently, all management of the sensor (e.g., starting and stopping recording) must be implemented over the radio module, which is also used for proximity detection. Control and data signals have to share the same communication channel.

The watch is powered by a single coin cell. However, coin cells are

Figure 6.2: CaReflect hardware iterations: (a) original Chronos eZ430 watch, (b) prototype of the wearable badge used during the first study, and (c) final robust badge case

not designed to provide a high current, which would be necessary for continuous active radio operation, even with low-power modules like the CC430. Consequently, coin cells are easily overloaded, although a coin cell can theoretically provide enough power for several hours of active radio operation. This well-known problem [164] results in a significantly lower capacity and finally a sudden drop in voltage. This brownout would lead to a crash and reboot that would repeat until the battery is empty or, depending on the firmware, has time to recover. Texas Instruments added three capacitors in a parallel circuit with a total capacity of 141 μF to reduce the problems of the coin cells [48]. The capacitors can provide the high current for short transmissions instead of the battery. The capacitors are slowly loaded by the battery when the SoC does not need much power.

Tests showed that it is a matter of a few seconds until the voltage will drop suddenly, if the radio module of the Chronos watch is in a continuous transmitting or receiving state. At this point the capacitors are completely discharged and the above-described problem occurs. Hence, we had to optimize power consumption not only regarding the maximum time of operation of the system but also to avoid a brownout due to longer time spans with higher power consumption. The power consumption of the sensor depends mainly on the usage of the radio module and the MCU. Table 6.1 shows the power consumption of the radio module and MCU in the states that are relevant for our firmware. The values have been confirmed by our own measurements.

The MCU provides various low-power modes to reduce energy consumption. The final system uses the lower power mode LPM0. In this mode, the CPU is off, but all clocks (e.g. the CPU clock) are still running. Therefore, the energy consumption of LPM0 is heavily dependent on the frequency of the CPU. In consequence, we reduced the frequency from 20MHz to 8MHz. One of the earlier implementations used the low-power mode, LPM3, of the MCU during the sleep phases. While this version was stable on approximately 50 percent of the sensors, a brownout occurred on the other 50 percent of sensors between 10 minutes and 3 hours after the start. The mode LPM0 resulted in a stable version for all sensors, but the MCU consumes approximately 300 times more power during the sleep phases than in LPM0. This reduces the battery life by 40 percent.

6.2.2 Low-Power Distance Estimation

The selected solution is based on a decentralized approach that needs no local installations of additional power supplies, coordinating access points, or a central server. Figure 6.3 depicts an example setup of four sensors worn by three residents and one carer. Resident 1 and carer A are in

Component	**State**	**Current consumption (mA) at $V_{DD} = 3.0V$**
Radio module	TX (-8.2 dBm)	15.8
	RX (active transmission)	17.24
	IDLE	1.7
	SLEEP/WOR	0.0002
Microcontroller	Active (8 MHz)	1.55
	Low-Power Mode LPM0	0.64
	Low-Power Mode LPM3	0.0021

Table 6.1: Power consumption of the Chronos eZ430 hardware according to [165, 166]

proximity. Hence, we assume that carer A is providing care to resident 1. Although resident 2 is very close, she, unlike resident 2, is not within proximity. However, if carer A starts to walk, there will probably be a short period of proximity between resident 2 and carer A. The resulting patterns of co-location and assumed care activities can be captured without a localization of each participant. This distributed sensing approach is similar to [106, 12, 115].

There are two main options to estimate the distance between sensors and to decide whether or not a sensor is in proximity: TOF and RSSI. Both mechanisms have been discussed in the context of indoor localization technologies in Section 2.2.2. The TOF is not applicable for our case because the distances are too short and the clocks of the Chronos eZ430 are not precise enough. The firmware would have to send a high number of packets back and forth to accumulate a significant time difference that can be measured by the Chronos. Moreover, frequent synchronization packets would be needed. In contrast, RSSI-based methods can be realized with only a few sent packets. However, RSSI signals tend to vary across several packets and multipath propagation may result in a large variation. Therefore, filter mechanisms are used that require additional processing, therefore increasing power consumption.

In our system, proximity between two sensors is detected based on the practical limited range of a radio module. According to Friis (see Section 2.2.2), the free-space path loss of the signal rises with the square of the distance until the signal-to-noise ratio (SNR) makes it impossible to

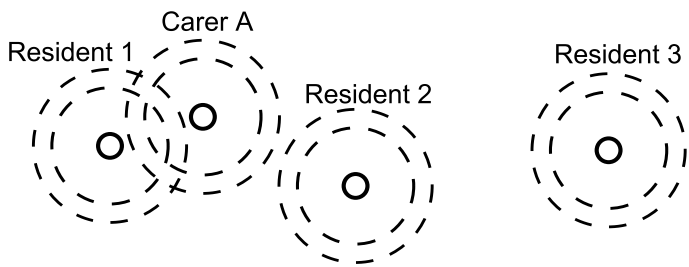


Figure 6.3: Schematic example of the proximity between three residents and one carer

receive a sent packet. The underlying assumption of this approach is that the noise will be relatively constant for distance measurements. Moreover, the orientation of the sensors will influence the distance. Finally, the body will shield the signal. We assume that the effects due to orientation and shielding will dominate because the Chronos has only one antenna. Therefore, the presented approach allows only a raw estimation, but any further effort to be more precise will be obsolete because of movement and shielding effects.

The signal strength of the Chronos radio module can be reduced to limit the range to the desired proximity distance (3m in our case). This has three advantages:

- Data transmissions with a smaller signal strength require less energy.
- Sensor nodes that are not within range do not have to keep their radio modules activated to receive packets that are later filtered.
- Furthermore, the radio module is already detecting signals from background noise in the RX state by monitoring the signal strength. In our approach, we reuse this process instead of implementing an additional filter at a later stage, which would require additional processing power.

In summary, this proximity-detection method is similar to RSSI-based methods. It requires less computing power, but is also less precise. Moreover, the reduced range decreases the noise induced by other sensors that are not in the desired proximity distance or are received by reflection. However, first tests showed that the range of a radio module fluctuates, so further filtering is necessary. A sensor can receive a packet from a sending sensor that is clearly outside of the proximity area (false-positive detection), but single packets are lost, even though a sensor is clearly within the proximity area (false-negative detection). The reduced signal strength reduces the number of false positives.

The false-negative detection was tackled by reducing the send power of the radio module only so far that it was able to send slightly past the proximity range. Furthermore, multiple packets are sent and received so that single missing packets can be filtered. This increases the false positive rate (i.e., more out-of-range sensors are detected). The system uses a simple filter to detect sensors that are not in the desired proximity area by analyzing the packet loss over several packets. Sensors in the system send several packets in a row and an amount of n_{recv} packets must be received to detect that the sender is within proximity range (e.g., if four out of five packets have to be received, a packet loss of 20 percent will be used to discriminate between sensors within proximity against sensors that are too far away). The mechanism can be understood as a simple low-pass filter. An energy-efficient filter will require a minimal number of packets to reduce the time sensors have to send and receive packets. The required number of packets and the required power can be minimized by choosing a filter that allows only one missing packet. The filter counts the number of packets and the contained sequence numbers and rejects a signal if more than one packet is missing before n_{recv} are received.

A high value of n_{recv} would reliably filter packets of sensors outside of the proximity range, but each sensor would have to send and receive more packets. A lower value would decrease the power consumption that is necessary to detect proximity, but would also increase the error. This tradeoff was evaluated in a pre-test. Figure 6.4 shows the proximity detection for different filter parameter values for a maximum radio range of approximately 4 meters. Even if one packet of the n_{recv} packets is not received, the sensor will still be detected as being in proximity. The filter significantly increases the robustness of the detection algorithm.

A value of 10 packets decreases the error to approximately 10 percent in the range of 3.5-4 meters. The number of packets must be doubled to decrease the error to almost 0 percent. Sensors within 3 meters are

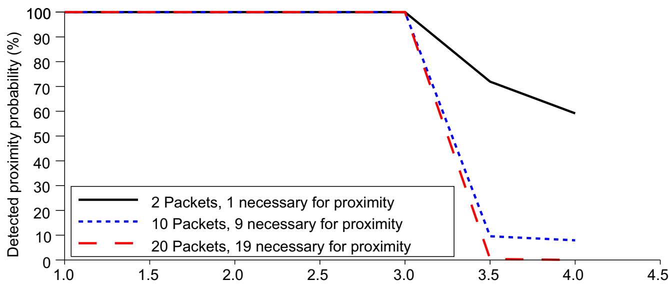


Figure 6.4: Proximity detection for different filter parameter values

reliably detected as being in proximity range. Based on these results, we decided that 10 packets decrease the error to an acceptable level, but 20 packets require too much energy in the longer send/receive states for only slightly decreased error rates.

6.2.3 Asynchronous Low-Power Listening

Based on the existing research on sensor networks (see Section 2.2.3), an asynchronous low-power listening protocol was defined to minimize the required communication between sensors. The times when sensors are sending and listening must be coordinated. Without staying in the receive state for too long, sensors must detect if another sensor is sending proximity packets. The design of the protocol was guided by reports of care staff and care home managers on their daily work practices. Drawing from their experience, we assume that in a worst-case-scenario, 10 people are within the range of a sensor. These rare situations could be, for example, a crowded meeting or a group activity in a common room. Most of the time, however, there is only one or no other sensor within range, because carers work with one resident, fetch material, or walk between residents. Consequently, we assume that, on average, one other sensor will be within range. The resulting requirements to the protocol are:

- low-power
- few contacts
- limited amount of continuous radio module activity (listening and sending) to prevent battery damage
- scalability (30 or more sensors)
- unlikelihood of more than 10 sensors in communication range

The Chronos default firmware already implements the SimpliciTI protocol [167] that can be configured to support star-topologies and peer-to-peer networks. The protocol is simple and efficient, and the corresponding code is available as open source. In our implementation, sensors take turns in broadcasting their addresses. They act as beacons so all sensors in the range of the radio module can receive the same packets. All sensors can determine if they are in proximity at the same time, because they can all process the same packets. Receiving sensors do not respond so as to reduce the communication overhead and avoid collisions with broadcasting sensors. Furthermore, the distance between sensors may have changed again. In summary, the communication is based on unidirectional broadcasts; the sending sensor does not know who receives a broadcast or if packets are received at all.

After first experiments with SimpliciTI, a simplified communication was developed based on the hardware layer of SimpliciTI. Figure 6.5 shows the original packet format and the shorter simplified format used. SimpliciTI packets contain several fields that are not needed for the above-described method, such as the destination address or the port. By removing unnecessary fields, the packet overhead can be significantly reduced. The send proximity packages contain only the ID of the sender and the broadcast ID, a simple counter to indicate the sequence of packages. This sequence number can be used to filter packages that arrive late or even twice because of reflections. If a packet is received by another sensor, this will indicate that the watch is in the proximity area.

Figure 6.6 shows the three different phases of the proximity protocol for two sensors. These phases are designed in a manner that allows all sensors to receive broadcasted proximity packets within the specified 10-second interval. No prior synchronization of sender and receiver is required.

Most of the time, sensors are in the *WOR phase*. In this phase, sensors are mainly sleeping to save energy, but check regularly if another sensor is broadcasting. To this end, the radio module stays in RX for $t_{rxTimeout}$. If the radio module receives at least one packet in t_{packet}, the sensor switches into the *RX Phase* and waits for t_{rx} to receive 10 packets. Sensors in the *RX Phase* put their radio module in the receive state long enough to receive at least 10 packets.

PREAMBLE	SYNC	LENGTH	MISC	DSTADDR	SRCADDR	PORT	DEVICEINFO	TRACID	PAYLOAD	FCS
RD*	RD*	1	RD*	4	4	1	1	1	n	RD*

PREAMBLE	SYNC	LENGTH	MISC	TRACID	SRCADDR	FCS
RD*	RD*	1	RD*	4	4	RD*

*RD: Radio dependent populated by lower stack levels or handled by the radio itself

Figure 6.5: Packet structure of SimpliciTI [167](top) and simplified proximity protocol (bottom): unnecessary fields have been removed to reduce the packet size

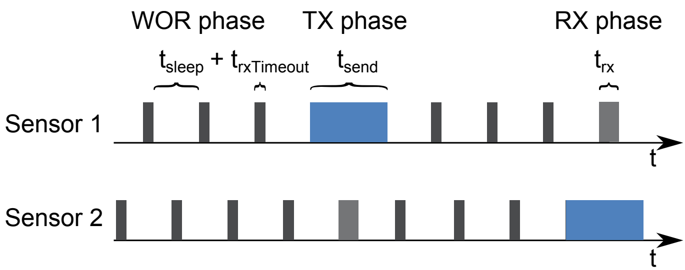


Figure 6.6: Asynchronous low-power listening: the system is sleeping for specified intervals t_{sleep} but periodically checks for a short time $t_{rxTimeout}$ if another radio module is active. If it is, the sensor attempts to receive 10 packets in t_{rx} to estimate proximity. Every 10 seconds, the sensor broadcasts its own ID for a longer period t_{tx}.

The proximity is evaluated afterward, based on the packet loss within these 10 packets, which can be deduced by analyzing the transaction IDs. The transaction IDs increase for each send packet. Hence, missing packets can be identified. The *TX phase* is reached once during the 10 seconds sampling interval. Sensors continuously send proximity broadcasts for the entire duration t_{sleep} between two WOR checks.

$$t_{tx} > t_{sleep} + 10 \cdot t_{packet}$$

Consequently, other sensors in the WOR phase can detect the transmission of this sensor and start to receive the packets in the *RX phase*.

The radio module of the eZ430 Chronos includes a Wake-on-Radio (WOR) feature that puts the radio module in regular intervals into the receive mode to detect if another radio module is sending [168]. The WOR reduces the power consumption significantly. If a packet is currently transmitted, the radio module will stay in the receive state to receive the whole packet; otherwise it returns immediately to the sleep state. The implementation of the *WOR phase* was realized using the WOR feature of the hardware. If a packet is detected, the sensor will switch to the *RX phase* to stay in the receive state until at least 10 packets are received or a timeout occurs.

The key to optimize the power consumption using the above-presented communication lies in the correct timing of the different phases, which depend on the specifics of the used platform. Furthermore, collisions (i.e., multiple sensors sending at the same time) should be avoided or solved as fast as possible. The timing or duty cycle has to optimize the duration of listening phases $t_{rxTimeout}$, sleep phases t_{sleep}, and the send interval t_{tx}. The longer the sensors broadcast, the longer they can sleep. An optimal value t_{sleep} and t_{tx} had to be found to minimize the power consumption and to balance the power consumption for sending and receiving. The length of these times implicitly defines the number sensors that can broadcast in the 10-second sampling interval without a collision.

Using our implementation on the Chronos eZ430, the radio module needs about 2.7 ms to check if the channel is clear and one packet is transmitted. A delay of 0.7 ms was introduced to fix stability problems with the power source, which would otherwise occur during this phase. During the 0.7 ms, the MCU waits in a low-power state before it transmits the next packet. Consequently, the sensor needs a total time of $t_{packet} = 3.4$ ms to send a packet.

The Chronos platform limits the possible values for $t_{rxTimeout}$ for the WOR feature to a divisor of powers of 2 of t_{sleep}.

$$t_{rxTimeout} = \frac{t_{sleep}}{2^n}$$

The divisor 64 results in the lowest possible $t_{rxTimeout} = 3.9$ ms, during which one packet can always be successfully received. The sync word of a packet has to be received during this time to detect a transmission of another sensor.

Figure 6.7 shows the different currents that are required depending on the value of t_{sleep}. The calculation is based on the current consumption values listed in 6.1and the following formula:

$$I_{interval} = t_{tx} \cdot i_{tx} + n(t_{sleep} \cdot i_{sleep} + t_{rxTimeout} \cdot i_{rx})$$

with

$$n = \frac{10s - t_{tx}}{t_{sleep} + t_{rxTimeout}}$$

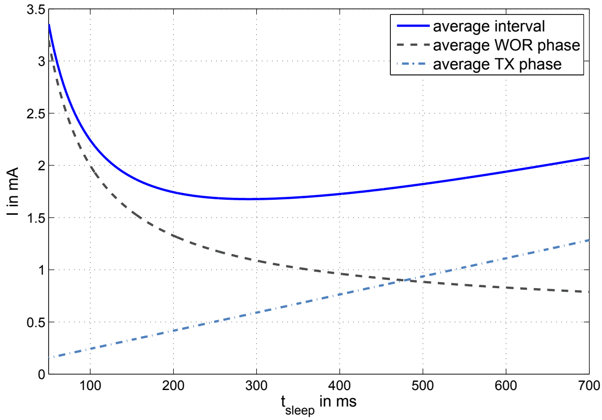


Figure 6.7: Current during proximity measurement for different values of t_{sleep}

In *RX phase*, the radio module has to stay in RX for at least to receive 10 packets. However, it is very likely that the radio module enters RX in the middle of the transmission of one packet, which cannot be received then. To avoid this problem, the value was adapted:

$$t_{rx} = 11 \cdot t_{packet} = 37.4ms$$

As a result, we set t_{sleep} = 250 ms and t_{rx} had to be doubled to 7.8 ms. During the TX phase, sensors transmit broadcasts for t_{tx} = 290 ms. The additional 40 ms are sufficient to broadcast 11 additional packets, because sensors in the RX detection phase check every 250 ms (t_{sleep}) if another sensor is broadcasting by using the WOR feature. If such a receiver detects the broadcast at the end of the 250 ms interval, it will still receive 10 packets to calculate the packet loss and can decide if the sensor is in the desired proximity area.

As sensors act asynchronous, two sensors may attempt to broadcast their proximity packets at the same time. Therefore, sensors check before the transmission of every packet if another sensor is currently broadcasting to avoid collisions and overlapping send intervals of two sensors. If a sensor attempts to enter *TX phase* and detects the transmission of another sensor, it will stay in sleep mode for a random amount of wait intervals t_{wait} before trying again. The seed of the random generator is the measured signal-strength value of the background noise when the proximity firmware is started.

The duration of a single wait interval t_{wait} after a collision is set to 400ms, because 400ms is the lowest integer divisor of 10s that is bigger than t_send. The time t_{wait} has to be bigger than t_{tx} because otherwise the sensor might restart broadcasting before the other sensor has finished sending. In summary, the sensors use 250ms WOR intervals, but 400ms intervals for sending after a collision, which results in 25 possible send intervals during the 10s proximity interval after a collision. A further advanced CSMA algorithm is not necessary because there are 25 slots for up to 10 devices. Furthermore, in most situations, only one or two devices are in range, lowering the collision probability.

In general, a synchronized protocol implementation would have been preferred, because synchronized protocols can nearly eliminate packet collisions and, therefore, reduce the required packets. However, the Chronos watches have a significant clock drift. The most precise of the available clocks has a frequency error of 0.002 percent, which would lead to a drift of 1.5 seconds per day. Moreover, the low-power optimized clocks have a much higher error [169]. During one day, the clocks on two sensors differed already by more than a minute. Therefore, frequent synchronization of clocks would be needed. However, the wearable sensors form a highly dynamic network. Moreover, most sensors have only contact to one sensor at a time. In the worst case, one sensor would be disconnected long enough to fall out of the assigned time slot. Asynchronous protocols provide an alternative because they do not require synchronized clocks. Therefore, communication partners can start to exchange packets without synchronizing. Furthermore, they are particularly suitable for small, broadcast-based networks.

6.2.4 Data Logging

The detected proximity contacts must be logged in the limited memory of the eZ430 Chronos, which has a flash memory of only 32kB. Approximately 8000 bytes are available in the flash memory for data logging because the firmware code is stored in the flash memory as well. If one or more sensor contacts must be stored every 10 seconds, the memory requirements

Start 9 bytes	Contact 1 4 bytes	Contact 1 4 bytes	...	Contact n 4 bytes	End marker 3 bytes

Figure 6.8: Memory structure of stored contacts

will grow rapidly. Therefore, a compact data structure was necessary to store a large number of entries. The chosen data structure is based on the original data-logging project for the wristwatch by Texas Instruments and is depicted in Figure 6.8.

The sensors assume a fixed sampling rate of 10 seconds. As a result, the current time for a contact can be calculated by knowing the start date of the recording and the number of scans since the start time. The start date and time are stored at the beginning of the recording in a short header. The current scan number is stored in a 14-bit counter. This allows recording a contact up to 1.8 days after starting. After 1.8 days, the counter is reset and a marker is placed in the stored data to indicate this reset. If proximity of the same sensor is recognized repeatedly in consecutive scan intervals, the proximity contacts are stored as a single event by indicating the length of a contact.

As a result, a single proximity entry comprises the counter indicating the start time, the duration of the contact, and the unique ID of the sensor. The whole data structure of a contact requires 4 bytes of space for each contact. The duration of a contact can be as long as 2.8 hours, but longer contacts will be concatenated when the data are read. One byte is reserved to save the ID of the detected watch. The current implementation utilizes only 5 bits to store up to 32 different proximity sensors in a study. The remaining 3 bits are reserved to either support more sensors or to allow larger time stamps.

6.2.5 Accuracy and Battery Life

The accuracy of the proximity detection was tested using three sensors. All sensors were placed in a distance of approximately 50 cm of each other on a flat surface and not moved until the end of the test. The results show that the proximity detection is stable. In only 2 percent of the complete tests, proximity cannot be detected, although the sensors remain in proximity range all the time (false-negative detection). In the study, however, the recognition rate will be lower because bodies shield the signal and movement will lead to asymmetries in the recognition. Nevertheless, each proximity contact is logged by at least the two sensors that are in proximity. The probability that both sensors do not recognize each other at the same time because of time drift is approximately 0.04 percent. Our test showed that the false negatives of two sensors are independent of each other. Hence, the logged data are redundant and errors can be filtered.

Under extreme conditions, 10 sensors might be within close proximity. This rare setting results in increased battery consumption because each sensor has to receive broadcast from 9 other sensors. Furthermore, the continuous drift of the TX intervals and the increased amount of sensors result in a higher number of collisions and each collision will result in a false negative. We conducted a stress test to evaluate the effects on the sensors. The error rate increased until reaching 20 percent. However, because of the redundancy of the capturing, up to 8 other watches captured the correct data. Furthermore, the high load strains the battery and can induce a voltage collapse after 2 hours and reduces the lifetime of the battery overall.

The limited range of radio communication may increase the appearance of a problem known as hidden nodes. This problem will occur if two sensors are in proximity of a third sensor, but not in proximity of each other. Both sensors may send at the same time because they do not notice each other, but effectively jam the proximity signals to the third node. While this situation can occur, the frequent movement of participants will resolve the situation. Furthermore, both sensors can receive the proximity signal of the third sensor. Therefore, the redundant capturing of proximity can solve this issue.

The sensors were tested using a new coin cell with a capacity of 230 mAh. Two sensors were placed 1.5 meters apart to ensure a constant exchange of packets. We measured a maximum lifetime of 180 hours. All data were stored on the sensor and was analyzed after the data had been downloaded. In 98 percent of the 10-second intervals, both sensors triggered a proximity event. This error rate remained constant during the test.

The system is intended to be used for a week, or even shorter time spans, in care homes, similar to such observation frameworks as SOFI 2 [161]. The maximum battery life of a sensor depends on the average number of sensors within range. If sensors notice an active broadcast, the 10 packets containing the ID must be received every 10 seconds. After talking to experienced carers, we decided that, on average, one sensor is within range during a shift because carers spend a significant share of the shift on their own (e.g., preparing medication or walking between residents). Therefore, the battery life was measured by two sensors. Both sensors had new coin cells with capacities of 230 mAh. They were placed on a flat surface at a distance of approximately 50 cm. One sensor acted as a reference. The sensor's battery was changed at regular intervals and the captured data were analyzed. The tested sensors lasted for more than 6 days, or 140 hours, using one coin cell. In practice, the 140 hours must be reduced to 135 hours to retain enough energy to download the sensor data with the sensor manager. If the battery is empty, the data can still be recovered. However, the entire memory has to be downloaded and analyzed. We developed a tool to simplify this process, but if the sensors crash during a write operation, there can be invalid contacts that have to be filtered.

6.3 Proximity Sensor Evaluation

The study reported in [202] used the initial version of the sensors as a probe in a UK care home that is specialized in dementia care. The goal was to learn about the acceptance of the system and practical application barriers. Privacy considerations especially were expected to be a potential barrier. The study collected a first set of data from the carers in their daily work and mirrored it back to them after the shift. The underlying hypothesis was that carers are able to understand and interpret the data. Hence, visualizations of the data can trigger reflection about critical events.

6.3.1 Method

All carers and residents on a small ward were equipped with proximity sensors during three morning shifts. Observers were not permitted during the study. The focus of this study was to measure the acceptance and usefulness of the proposed system for carers. Both can be deduced from the captured data and concluding interviews, which focused on such questions as: Are carers willing to distribute and wear sensors for a limited time span? Can the data provide sufficient insights for carers? Which insights can be gained by further analyzing the data?

After discussion with experienced carers, the minimal distance to trigger a proximity event was set to 1.5-2 m. However, this is the maximum distance if two sensors face each other. If carers or residents are side by side, the range drops to 1 m. An additional sensor was placed at the laptop computer that was used for documentation tasks. The sensors were disguised with soft and colorful material in the form of a brooch. Carers expected that residents would cooperate more easily if everybody on the ward wore the same kind of colorful badge. Residents remained on the ward during the study. All carers, except for one, worked only on this ward. Hence, the captured data provide a complete picture of the daily work activities of carers and the care received by the residents.

All participants were required to sign consent forms to take part in the study. The carers distributed the sensors to residents and among each other on three consecutive mornings. The morning shift started at 7:00 and ended at 14:00. After the shift, the data were read from the sensors using the management software that ships with the Chronos eZ430. This software can connect to one watch at a time to set parameters and download recorded data. Researchers had to manually stop every sensor and start the synchronization. The mapping between sensors and carers was documented in several lists. Small changes to the software allowed data storage directly in the Unisens format [142]. Each available sensor was stored as a separate event entry type. The events contained the start and end of a contact as well as the ID of the contacted sensor. The unprocessed data were visualized directly using the UnisensViewer. Figure 6.9 depicts a screen shot of example data that were shown to carers. In short interviews, carers were asked to explain their own raw data (e.g., long periods of documentation or providing care to multiple residents at the same time). This interview was the only possibility to verify the recorded data because observation of care staff was not permitted.

6.3.2 Results

During the 3 days, 15 sensors were used. Two sensors failed on the first day and one on the second day because of residents pulling at them. The sensors were initially attached at residents' and carers' chests. However, residents were confused by this new item and grabbed for them. Therefore,

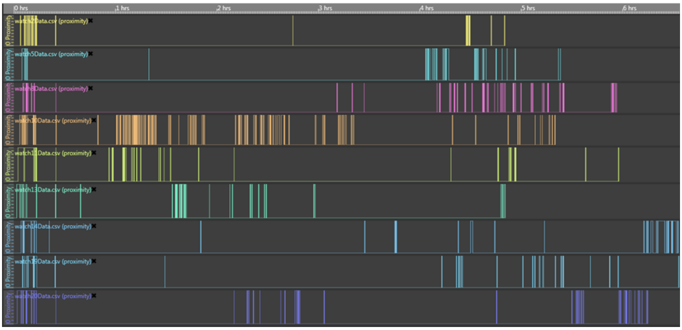


Figure 6.9: Example of raw data of one carer as shown in the UnisensViewer [142] to carers after the shift. Each row shows a resident and each colored bar in a row indicates proximity to this resident.

the carers decided to place the sensors at the hip or under a shirt. A total of 41 successful measurements and 290 hours of data were recorded.

In average 44 percent of a carer's shift could be matched to a specific resident or a documentation task. The detailed results for the carers are listed in Table 6.2.

Carer 4 has a lower time share because this carer worked only part time on the monitored ward. The remaining 56 percent of the carer's time was spent walking between residents or caring from a distance. Moreover, the 10-second sampling interval may miss short contacts. When the results were visualized to carers after the shift, carers could recognize behavior from the raw data and started to discuss care practices. The added value was more important to carers than their privacy concerns.

The time spent on documentation was lower than expected. Initially, documentation effort was seen as the most tedious and time-consuming task by all carers. When asked before the study, the carers and the care home manager estimated that 30 minutes of care will result in 10 minutes of documentation. However, the proximity sensors showed that, on average, only 12 percent of the proximity events for a carer are related to the documentation laptop. This percentage varies across carers. A new carer spent only 4 percent on documentation, while an experienced carer spent 21 percent on documentation. Carers confirmed in the interviews that

Carer	Day	Shift duration in hours	Covered by sensor events in % of shift	Measured documentation time in % of shift
Carer 1	1	8	35%	n/a
Carer 2	1	3.5	42%	n/a
Carer 3	1	8	45%	n/a
Carer 1	2	8	48%	17%
Carer 4	2	8	25%	16%
Carer 5	2	8	43%	2%
Carer 6	2	8	59%	8%
Carer 1	3	8	48%	21%
Carer 7	3	8	48%	13%
Carer 8	3	8	48%	10%

Table 6.2: Recognized proximity of a carer to at least one resident or the laptop during the shift

experienced carers support novices in documentation tasks. Documentation times are the most reliable values because the distance between carer and laptop is small and there are few movements. Carers attributed this differences to the experienced burden of documentation. Furthermore, if a day was busy, carers had to stay longer to finish the documentation. As a result, documentation is often the last impression of a busy day.

The time one carer provided to a single resident varies significantly. The differences between residents are visible in the raw data illustrated in Figure 6.9. Some residents had only a few short contacts (e.g., the data shown in the first row). Other rows show unusually long and frequent contacts (e.g., the fourth row). Interview partners confirmed the correctness of the recorded data, but had expected more contacts. They felt that they were busy with residents all day. Although the data were shown in an anonymized manner, carers immediately recognized individual residents by their data: *"This must be [name of a resident]"*(pointing at a row).

Carers explained that one resident with few contacts stayed in bed during the majority of the shift. Unusually long resident contacts were caused by residents sitting next to a carer who worked on documentation tasks. These reactions prove that the data reflect the actual workday and the raw data can be understood by carers. However, this raises new concerns regarding the anonymization of recorded data.

As mentioned above, nearly half of the overall shift was not registered by a sensor. Three residents (R5, R7, R8) and the laptop account for the majority of the time connected to proximity events. This is a typical pattern among all carers; two or three residents demand the majority of the carer's attention. Although some of these differences between residents are due to the shift organization, few residents keep all carers busy each day. During the 3 days of evaluation, there were only slight variations in the residents requiring more care. These emerging patterns have to be analyzed in a longer study.

Figure 6.10 shows an example of the captured data. The break taken by the carer between 10:30 and 10:50 is clearly visible and could trigger privacy concerns. All participants said that the benefits of the system outweigh their privacy concerns. Carers even suggested more critical approaches (e.g. long-term monitoring using the sensors or finding ways to present these data to relatives). Only one carer voiced privacy concerns when asked. After a detailed description of the system, these concerns were resolved.

After the last interview, a short group discussion between three carers took place. The carers discussed, for example, whether attendance to a resident will be influenced if the resident stays in bed all day. This might lead to insufficient attention to residents in the sleeping rooms, because most residents are in common rooms. The captured data were seen as beneficial to understanding care practices and discussing possible improvements. Carers could identify time spent with residents and when contacts took place. This allows the identification of typical patterns and reflecting on reasons for resulting time shares.

The recorded data were analyzed to generate a report for each carer, which contained an overview and a timeline for each day. The timeline shown in Figure 6.10 visualizes one shift of a carer during the study. The carers meet in the morning in the common room to plan the day. Afterwards, residents are washed, dressed, and brought into the common room, if possible. From 10:00 to 13:00, the majority of residents are in the two common rooms. Multiple contacts between carers and residents are now likely (e.g., when two residents are sitting at the same table). After lunch, most residents return to their rooms. This timeline is similar to the data presented in the study, as show in Figure 6.9 but the aggregation in 5 minute intervals provides a quick overview.

The data were further analyzed regarding the symmetry of recorded contacts. If sensor A is within proximity of sensor B, both sensors should record a proximity event. Small differences have been expected, because of the asynchronous recording and constantly changing proximity. In this study, the recorded proximity between two sensors differed up to 20 percent. This is due to the different maximum communication distance between sensors because of sensor orientation and body shielding.

The carers brought up several challenges that need to be addressed. The sensors need a more stable, waterproof casing to survive the challenges in a care home. All buttons should be removed or deactivated. Therefore, the data transfer from sensors has to be fully automatic. The analysis of the data can be improved to shed light on specific care practices. A connection to the official documentation system would be beneficial, because task and non-task-related proximity events cannot be distinguished.

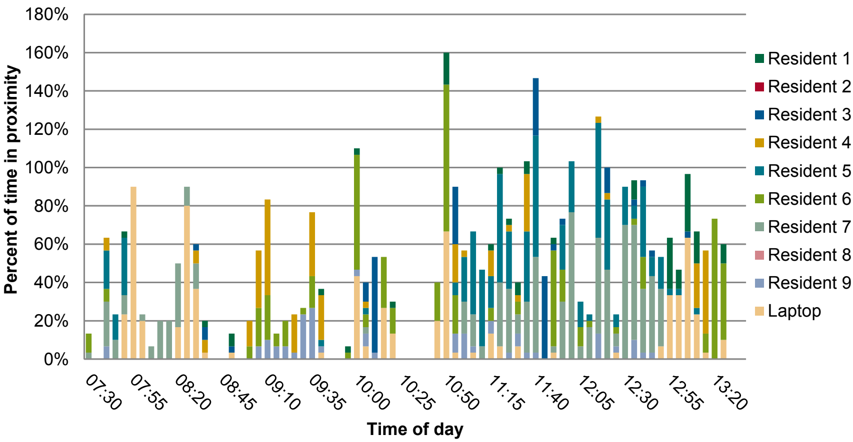


Figure 6.10: Timeline of a carer's shift showing the time spent in the proximity of a resident within 5-minute intervals.

6.4 CaReflect Prototype

CaReflect, a visualization and management software, was created in response to the requests in the first study. The developed CaReflect sensor-management software can configure, start, and stop sensors. When a sensor is stopped, recorded data are downloaded automatically and stored in a database. A variety of visualizations facilitate the inspection of the data according to requests from end-users. Figure 6.11 depicts the final system as it was used in the studies that are reported in Section 6.5. The

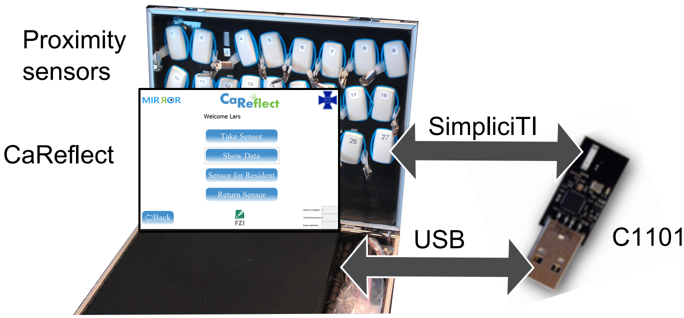


Figure 6.11: CaReflect suitcase with sensors and management laptop

suitcase contains all required components to conduct a study. The laptop starts with the CaReflect software that uses the CC1101 USB access point to communicate with the sensors. The system could be used without additional equipment. The used USB access point comes with the Chronos watch. A custom firmware has been installed to communicate with multiple sensors simultaneously using the SimpliciTI protocol [167] that was part of the pre-installed Chronos and access point firmware.

6.4.1 Architecture

The CaReflect manager is a Java-based software that consists of five major components that work together to realize the required functionality. The components and their interaction are depicted in Figure 6.12. The *Study GUI* is the visible part of the application. It is a full screen application that starts automatically when the study laptop is started. Users can

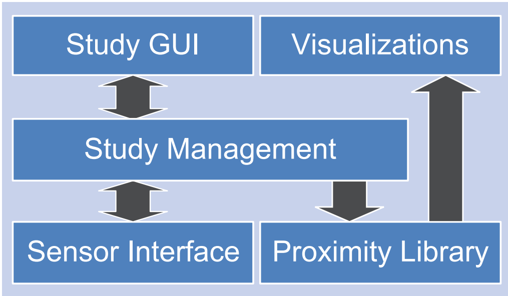


Figure 6.12: CaReflect architecture

create accounts, take sensors at the beginning of the shift, and return them after the shift. All activities can be performed with two or three clicks. However, users must sign in to take a sensor and again later to review their data.

The *Study Manager* monitors which sensors are available and their status (e.g., if they are already assigned to a carer or resident). It is controlled by the *Study GUI* and communicates with the sensors using the *Sensor Interface* to check and change the status of sensors as well as reading proximity data. Data are stored using the *Proximity Library*, with special attention to the arising privacy aspects of working with such sensitive data. Finally, the analysis and visualization component provides the tools to draw charts and browse through the collected data.

6.4.2 Sensor Interface

The *Sensor Interface* communicates with the sensors by using the Java Native Interface (JNI). The interface to the proximity sensors is based on the CC1101 USB wireless interface. The CC1101 acts as SimpliciTI access point. Sensors request a connection and receive a channel ID to be addressed later on. After sensors are assigned a channel ID, they will start polling for new commands in 500ms intervals. Supported commands are reconfiguration commands or a request to read the current configuration of the sensor. The original firmware supported only one connection to a single sensor at a time. The access point firmware was extended to support up to 10 sensors at the same time. More sensors and corresponding channel IDs could not be stored because of memory limitations. Sensors still connect in regular intervals, but an independent command queue is maintained for each sensor. The command protocol of the USB interface was analyzed and re-implemented as a new C++ driver to communicate with the access point based on a serial interface. The main difference from the original driver is the possibility to specify the sensor that should receive the send command. The driver was packaged as a dynamic-link library (DLL) and made available to the Java application by using the JNI. The communication between sensor and access points follows strict timing rules. The access point has to react within milliseconds to prevent a timeout. If other sensors attempt to connect at the same time, the connection is reset and a send command is ignored. As stated above, the transmission of commands to the sensors is error prone. A growing number of sensors in the receiving distance of the access point increase the likelihood of a packet collision and a transmission error. Therefore, every command that does not result in an immediate response, such as a reconfiguration command, is wrapped into read commands to ensure that the new values are correctly set. Furthermore, connected sensors were extended with a new sleep command that increases the poll interval to 10 seconds. As a result, sensors in sleep state are slower to react to commands, but collisions can be reduced.

Reading of data does not follow the poll-command pattern described above. The download state is started by a dedicated download command. The sensor starts to send packets in shorts bursts and does not expect a reply. The sensor DLL has to read the receive buffer continuously to prevent data loss. Transmissions of other sensors may lead to lost packages as well. If more than 80 percent of the packets within a burst are missing, the burst is completely restarted. If fewer packets are missing, these packets are requested by requesting only the missing packets. However, these single packets are not transmitted in burst mode.

6.4.3 Data Storage and Visualization

Data storage was implemented in a separate library, *Proximity Library*, to hide implementation details. The library contains not only functionality to read and write data, but optimized methods for often-used queries. For example, the question of which users have been in contact with a specified sensor at a specified time can be implemented in a more efficient manner if the specified time and sensor can be used directly to filter the data. Otherwise, a large number of objects has to be created in memory to filter them in code later on. Supported storage formats include MySQL, HSQL, and a file-oriented format based on Unisens [142].

The analysis component shown in Figure 6.13 builds on the supported queries of the Proximity Library to show visualizations that aim at triggering and facilitating reflection. Three types of visualizations are supported: timelines, pie charts, and bar graphs. Timelines aim at reconstructing the work process in chronological order to facilitate recollection of specific timespans. Pie charts were requested by care staff because carers can see how they spent their time at a glance. Bar charts are available for in-depth comparison (e.g., how much attention was overall provided to each resident). Although carers can use the visualization component to create all three kinds of graphs for different kinds of data, the system uses a small subset of the graphs by default. Carers requested that a pie chart

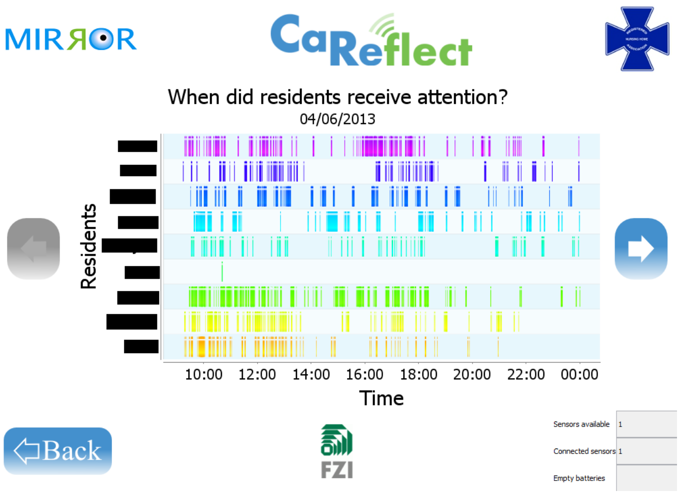


Figure 6.13: CaReflect user interface: visualizations of all care activities during one day

provide a brief overview of their day before any other graphs are shown. Carers can analyze the data in detail by showing a timeline of the day. Bar charts are used only to compare documentation effort or the overall time spent with residents by all carers.

6.5 CaReflect Evaluation Method

CaReflect was tested in UK care homes to evaluate the acceptance of sensors and the underlying concept by care staff. Furthermore, the impact on learning and reflection should be measured. Evaluation tools can only measure articulated insights and ask for feedback, because reflection is a cognitive process. A desired behavior change requires long-term monitoring, which is neither the intention of the developed system nor accepted by care staff. Hence, researchers have to stick to questionnaires and anecdotal evidence (see Section 1.3).

Three studies were conducted in three different care homes with a varying number of participants and sensors. The following sections describe the used method and results.

6.5.1 Procedure and Tools

All three studies listed in Table 6.3 followed a similar procedure. On the first morning, the care home staff received information about the study and the used system. After collecting signed consent forms from all carers and the relatives of selected residents, care staff distributed the sensors to residents. Each carer was asked to create an account and take a sensor.

Care Home	CaReflect version	Duration in days	Workshop	Measured contacts
Mansfield	0.5	3	yes	47532
Risby	0.8	3	canceled	51255
Wren Hall	1.0	4	no	45749

Table 6.3: Conducted studies with CaReflect in care homes.

Directly after the shift, care staff were asked to return the sensor and review the data during an interview. Providing this time was difficult for carers because residents were more demanding on some days. Furthermore, carers wanted to go home on time (e.g., to pick up their children from school). Carers were asked about their expectations and their knowledge of the last shift, before using CaReflect Carers had to estimate the amount of documentation and the residents they spent the most time with. A short questionnaire was used to quantify the findings.

Observations took place during the usage of the CaReflect application and especially during the review of recorded data. Carers were encouraged to think aloud during the review of the data. Considering the early stage of the app, the researcher provided guidance for using the app. Observations during the shift were not possible.

Questionnaires used a 5-point Likert scale (1 strongly disagree, 2 disagree, 3 neutral, 4 agree, 5 strongly agree) and a few free text questions. Questionnaires had to be succinct because carers had only limited time to fill them out and were often confused by more complex questions. Questionnaires in the first two studies were intended to evaluate acceptance and uncover possible flaws; the questionnaires in the final study additionally measured the impact on learning and reflection.

Interviews after each shift focused on collecting feedback regarding the impact of the data, the used visualizations, and the willingness to share the captured data. The quality of the data and the impact of the visualizations were evaluated by interviewing carers before and after using CaReflect.

In concluding interviews, carers were questioned regarding their impression of the App, their willingness to share the data, and possible applications of the app. These semistructured interviews provided more insight from care staff. Care staff were asked to explain why and how they used the system. Managers reported on the perceived value of the system for the care home.

The Net Promoter Score (NPS) [170] was used to measure whether participants would not only use the system, but actively recommend the system. The score is used by many companies to measure the impact of their products. Participants score the likelihood of recommending the used product on a scale of 0 to 10. All ratings below 7 are counted as detractors (D). All ratings over 8 are counted as attractors (A). Ratings between 7 and 8 are neutral and, therefore, ignored. The NPS is the difference between attractors and detractors divided by the number n of participants.

$$NPS = \frac{A - D}{n}$$

This is a very strict measurement, but the results measure the possibility of using the CaReflect prototype as a product.

In addition, we planned to conduct workshops at the end of the study to stimulate the articulation of insights between care staff without the influence of a researcher. However, in most care homes, only one or two carers can leave a shift to take part in a workshop. Therefore, workshops have to take place after work and conflict with carers' daily practices. As a consequence, only one workshop was conducted in the Mansfield care home. If a workshop was not possible, concluding interviews were conducted with carers and management.

The first study was a feasibility test. In subsequent tests, we improved the application and aimed at increasing the number of participants and covered shifts. The final experiment was again conducted in the Wren Hall care home, the home where the initial test of the proximity sensors took place. In the final study, reports for each carer and the overall care home have been created using CaReflect. We collected overall feedback in short interviews and questionnaires 1 week later. For each of the available carers, approximately 20 minutes were spent to feedback their own data, show some of the aggregated data slides, and administer the end-of-trial questionnaire. Carers were asked about the usefulness of CaReflect for triggering reflection, the usability of actual physical sensors, and privacy concerns. In a concluding interview with the manager, the organizational-defined objectives were reviewed and the possible impact of the CaReflect system regarding the planned objectives was discussed.

6.5.2 Participants

The selection of residents and locations that should be equipped with sensors was left to care staff. However, the number of sensors was limited. Research advised care staff about successful patterns, such as equipping only one ward or lounge, but in the end the carers decided. Care staff selected nine residents to compare typical care patterns across residents.

For example, carers selected some residents requiring intense care and other residents they felt might be neglected. Additional sensors were placed in the office where care staff work on the documentation, on the medical trolley, and in selected locations. The distribution of sensors in all studies is summarized in Table 6.4. The number of staff members with sensors increased with each study because the number of available sensors increased from 20 to 30 and the time needed to download the data and configure a sensor was optimized. The number of contacts grows with the number of participants and the length of the study. In the final study, all carers were equipped with sensors. Hence, all care activities related to the selected residents were captured.

In the study in the Mansfield care home, 19 proximity sensors captured 630 hours of data. During the 3 three days, 21 carers conducted their work as usual while wearing sensors. Fifteen interviews were conducted with staff members (10 male, 5 female), all of whom were working full time. Researchers interviewed one nurse, seven care assistants and seven senior carers (a total of 15 interviews) with a varying degree of professional experience ranging from 6 month to 19 years. The impact of these data on collaborative reflection and the organization was researched by arranging a workshop comprised of three carers, one manager, and the assistant manager. The management view of the data was shown as an introduction. Insights gathered from the interviews were discussed.

The studies in the Risby and Wren Hall care homes were conducted in collaboration with Tracoin, a project partner in MIRROR. In the

Care home	**Care Staff №**	**%**	**Residents №**	**%**	**Locations equipped with sensors**
Mansfield	21	65%	9	30%	nurses' office
Risby	30	50%	9	20%	nurses' office and 2 medical trolleys
Wren Hall	45	100%	9	17%	nurses' office and 4 medicine cabinet areas

Table 6.4: Participants and sensor coverage in conducted CaReflect studies.

Risby study, staff pointed to the importance of the medical trolley as an indicator of a typical care task. The study resulted in many contacts and observations showed the reaction of carers. Carers wanted to evaluate whether they stick to the desired “butterfly care” mode. Data indeed showed many short contacts. However, questionnaires and interviews were not accepted because there was not enough time. Consequently, no demographic data are available and a planned workshop was canceled.

In the final study in the Wren Hall care home, 28 sensors were allocated to nine residents, five locations (four medicine cabinet areas and the nurses’ office), and to 44 carers over a 4-day period. The carers were monitored for one shift or a number of shifts, depending on the roster, with between 3 and 12 carers being monitored at any one time.

6.6 CaReflect Evaluation Results

The results of the three experimental studies are presented as a whole because they shared a common method and produced similar results. The studies led to insights on the achieved quality of the data, which is crucial for acceptance by carers. However, a high data quality can lead to privacy concerns if the sensors are understood as a monitoring tool. The impact on learning and the resulting insights are significant results that can be appreciated only if the data quality is sufficient and privacy concerns are not a barrier to using CaReflect. The support of the care home management is essential for successful usage. Moreover, the feedback of managers is directly linked to the commercial potential of such a solution.

6.6.1 Data Quality and Relevance

Overall, more than 45,000 contacts were captured during the four days of the final study in Wren Hall. The distribution of recorded sensor events over time is shown in Figure 6.14. No sensors were distributed to carers during the first two night shifts. Documentation is only a small share of the contacts that have been measured. The four days do not provide a consistent pattern. Moreover, there were large differences in the recorded contacts between residents and carers each day.

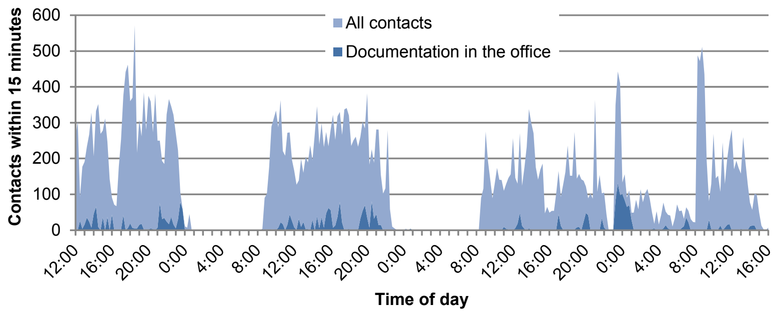


Figure 6.14: Timeline of all contacts and time in the office in the Wren Hall study.

Carers in all care homes could recognize their shift from the CaReflect visualizations and confirmed the correctness of the data as shown in Table 6.5. Differences between received care across residents were explained by care staff. The data shown on a day-by-day view often stimulated the carer to provide a narrative of this specific day (e.g. *"this was the day Allan died"* or *"this was the day Doris didn't want to get up"* or *"this was the day I spent ages in the office talking to John's daughter,"* etc.). Carers did not only recognize their timelines, but started to reflect and discuss their behavior with each other during their shifts. Moreover, they developed their own ideas on how to use the system.

Questionnaire item (5-point Likert scale)	**Study**	**Number of participants**	**Mean and standard deviation**
CaReflect helped me to collect information relevant to reconstruct work experiences	Mansfield	16	3.87 (SD=1.26)
The graph showed my work properly.	Wren Hall	40	3.77 (SD=0.92)

Table 6.5: Perceived data quality in Mansfield and Wren Hall studies

The correctness of the data also showed if there were errors in the data. For example, on the last day in Mansfield, one sensor was accidentally given to a carer instead of a resident. The two carers that first reviewed their data on this day immediately recognized the mistake. They could even name the carer, who took the sensor, and reconstruct what had happened. Similarly, in Wren Hall, carers immediately identified a resident whose sensor was not working properly.

In summary, CaReflect helped participants to capture data during the shift to later reconstruct work experiences.

6.6.2 CaReflect Usage and Usability

Carers and senior staff had only a few minutes to analyze the data. The general pie chart was often of more interest to them than the timeline. However, the inspection of the timeline visualization triggered more descriptions of particular events. Observed carers were interested in seeing the absolute and relative time given to residents, spent with other staff, and at various locations, particularly "the office." A number of carers said it was difficult to remember all their contacts over an 8-hour shift, particularly when encouraged to work in the "butterfly" mode (i.e., a large number of small contacts, rather than large blocks of a single contact). Carers used the CaReflect visualizations to explore the following questions:

- How was their time was spent differentially with specific residents?
- How much time was spent with another (specific) carer?
- How much time was spent with a specific resident?
- How much time was spent in the office or on documentation?
- How many different contacts were made over a shift?
- How much time was "undocumented"?

Observed team leaders and senior carers were especially interested in the overview graphs, showing the aggregated time given to individual residents from all carers. This aggregated level of data was seen as useful to reflect on performance, equity, volumes of service, frequency of contact, need, and amount of "doubling up" given for heavy, difficult, or highly dependent residents.

In Risby, carers quickly glanced at the data and were satisfied that they had as many short contacts as prescribed by the “butterfly care” method. CaReflect was used as a check and not as a reflection tool.

Participants suggested in the interviews that wearing a sensor, and knowing others are wearing one, could affect a worker’s behavior–in this case giving residents with sensors more attention. This effect may have influenced the result of the study, but can be a starting point for a long-term coaching approach based on self-tracking.

When asked for possible improvements, 90 percent of comments related to the number of sensors. More sensors would allow the capture of more data. In particular, all residents should be equipped with sensors. Furthermore, the user interface should include more flexible filter mechanisms. The mechanisms are already available, but have been removed from the carer interface to simplify the operation. One carer suggested reducing the range of the sensors to 1.5m to improve precision. Another carer would have preferred a self-reporting approach because the data are *“too vague, you might pick somebody up who is standing next to you. Carers should rather press a button to document their work.”*

6.6.3 Learning and Overall Acceptance

The impact on learning can be inferred from the quantitative results in the questionnaires (see Table 6.6) and the variety of gained insights. Furthermore, learning something new is the main benefit of CaReflect for staff members and, therefore, essential for overall acceptance. During this short time, carers did not have enough time to change their behavior. There is only a slight indication that changes are planned in Wren Hall, whereas the carers in Mansfield slightly disagreed with this statement. It should be noted that we simplified statements like “I made a conscious decision about how to behave in the future” into more direct statements like “I now have an idea what I could change” for the study in Wren Hall because 5 out of 16 participants in Mansfield said they did not understand this statement.

Figure 6.15 shows the perceived impact on reflection measured by two questions in the Mansfield study. The majority saw a benefit in looking at the data. It helped carers to reconstruct work experiences and made them reflect: *“It makes you think how much time you spent with whom.”*

Questionnaire item (5-point Likert scale)	Study	Responses from participants	Mean and standard deviation
CaReflect helped me to reflect on experiences from work.	Mansfield	16/16	3.87 (SD=1.20)
I gained a deeper understanding of my work life.	Mansfield	11/16	2.90 (SD=1.70)
I made a conscious decision about how to behave in the future.	Mansfield	11/16	2.30 (SD=1.55)
I learned something by looking at this data	Wren Hall	40/40	4.03 (SD=0.55)
I have now an idea what I could change.	Wren Hall	40/40	3.61 (SD=0.82)

Table 6.6: Perceived impact on learning in Mansfield and Wren Hall studies

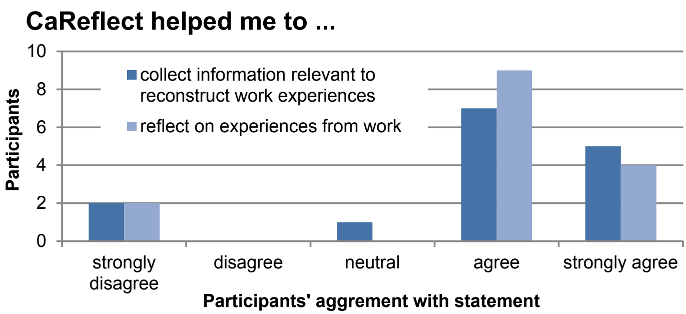


Figure 6.15: Reported impact on reflection

Furthermore, carers wanted to *“compare with others to see how much time they are spending.”*

When asked for examples of insights, carers talked about the time spent on documentation or the differences between residents. The latter was the main insight that was common across all care homes. Carers already knew that some residents need more attention. The data, however, quantified this impression and showed that the differences are bigger than expected. Difficult residents required not the expected 50 percent more effort than other residents but rather 4 to 5 times as much. For example, several carers found out that *“I have spent a lot of time with resident A.”* This resident A required the main share of the attention from three different carers. Care staff explained that it is definitely not their intention to give everyone the same amount of attention – on the contrary, the individual needs of each resident will be different, and judged accordingly. These needs vary from person to person and from day to day. When a resident was dying, they received almost constant attention, whereas others were happy on their own for periods of time. The knowledge of these individual needs is used to evaluate the CaReflect data – this is where reflective learning occurs most clearly. For example, one carer noted that a particular resident with a sensor seemed to get more attention than usual and responded by being more alert and brighter. In addition, low numbers were seen as a warning: *“I makes you think that people don’t get enough attention.”*

The time spent on documentation was often overestimated. In Wren Hall, one carer said *“I thought I spent more time in the office.”* Nurses, however, have to do the main share of the documentation. One nurse was surprised how much time she spent in the office: *“one hour [16 percent of time] spent in office!”* This amount of documentation was disappointing for a care coordinator: *“too much time spent at computer during morning shift”*

Several articulated insights are specific to the care home or the carer. For example, one carer was surprised how much time she needed to assist a resident with meals and wanted to discuss with colleagues about their experiences. In another example, one carer became aware of how the current care organization in teams is more dynamic than planned: *“We don’t only work in teams but mix and match more than we think. I think that is good.”* Other participants could not estimate how much time they spent with a resident before using CaReflect, but said afterward that it is useful to know.

The results of the questionnaires of the final evaluation in Wren Hall are shown in Table 6.7. If the feedback is split into experienced carers with at least 5 years of experience (12 out of 41) it becomes apparent that experienced carers see more benefit in using CaReflect. They believe that they learned more by analyzing the data and more of them had an idea how to change their behavior. The concluding interviews emphasized these differences (see Table 6.8). This result was surprising because we expected young carers to be more attracted to new technology. However, the existing knowledge and the wider perspective of more-experienced staff appears to be decisive. This would explain why the biggest difference between both groups can be seen in the question that asks for insights the team level. One member of the group that could not see the benefit of the app argued that they know their work and do remember what has happened: *"I already knew that."*

Figure 6.16 shows the overall feedback summarized by the loyalty metric NPS in the Mansfield and Wren Hall care home. A value above 7 is seen as positive. The majority of Mansfield users are distributed around the border between neutral and a positive view. Most carers used the sensors only once or twice because of the shift system. The short usage of the tool was one reason for this feedback (e.g., a critical carer said, "*I don't know enough about the tool to recommend it.*") The overall net promoter score in Wren Hall was negative (NPS -29 percent). This result is due to the

Questionnaire item (5-point Likert scale)	**All staff members (n=40)**	**Experience $\geq$ 5 years (n=12)**	**Experience <5 years (n=28)**
I learned something by looking at this data	4.03 (SD=0.55)	4.18 (SD=0.39)	3.96 (SD=0.60)
I have now an idea what I could change.	3.66 (SD=0.85)	3.77 (SD=0.91)	3.61 (SD=0.82)
Did you see anything surprising in the data?	2.00 (SD=0.7)	2.1 (SD=0.83)	1.90 (SD=0.62)

Table 6.7: Responses from the daily end-of-shift questionnaire in Wren Hall split by experience

Questionnaire item (5-point Likert scale)	All staff members (n=17)	Experience $\geq$ 5 years (n=7)	Experience <5 years (n=10)
I am satisfied with CaReflect.	3.82 (SD=0.62)	3.71 (SD=0.92)	3.9 (SD=0.54)
I have an idea now, how we can improve our work as team.	3.36 (SD=0.68)	3.57 (SD=0.50)	2.20 (SD=0.75)
I would like to use the app again.	3.71 (SD=0.75)	3.86 (SD=0.64)	3.60 (SD=0.80)
NPS	-29%	0%	-50%

Table 6.8: Concluding questionnaire responses in Wren Hall split by experience

many young detractors (5) among inexperienced carers (10 of 17) who did not yet see a value in the collected data. Experienced staff members (7 of 17) were neutral (NPS= 0 percent). These experienced staff members were comprised of all nurses and care coordinators.

Participants from Wren Hall stated in the concluding questionnaire that they were satisfied with CaReflect. Eighty-two percent said they would like to use CaReflect again with their team. Only 24 percent would also

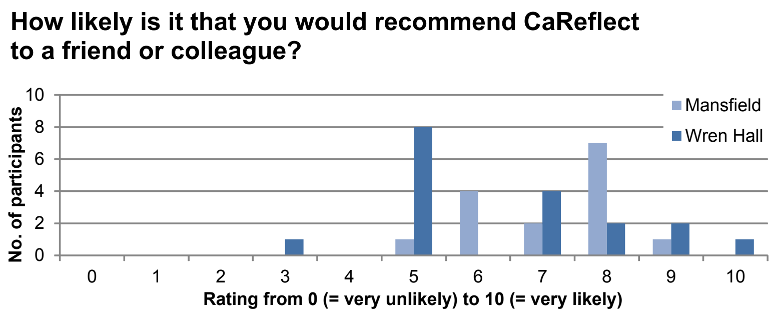


Figure 6.16: Feedback for Net Promoter Score (NPS) [170] in the CaReflect studies in Mansfield and Wren Hall

use it individually. The interviews showed that carers are eager to compare their data to others and improve together.

6.6.4 Privacy

The results of the initial study with proximity sensors, reported in Section 6.3.2, indicated that carers had surprisingly few privacy concerns. In the CaReflect studies, we aimed to verify this impression in other care homes.

According to a survey conducted in Mansfield, privacy concerns among carers were very rare. Figure 6.17 shows that 12 out of 15 interviewed carers would share the data with everybody. The researcher deliberately pushed for more critical responses (e.g., by suggesting that these data be made available to all relatives of a resident or even put it on the Internet. Finally, carers were reassured that any privacy concerns would be taken seriously and would lead to changes in the system. Carers responded by saying *"I have nothing to hide"* or *"this is good proof . . . when talking to relatives."* One of the carers who wanted to share the data only with management said *"I don't mind sharing it internally, but not external. It is private,"* but was actually more concerned about the privacy of residents. She feared that residents will be embarrassed if their relatives knew how much care they receive. Furthermore, carers asked for an option to share and compare their shift data to other carers. They expected that such a comparison would yield more insights.

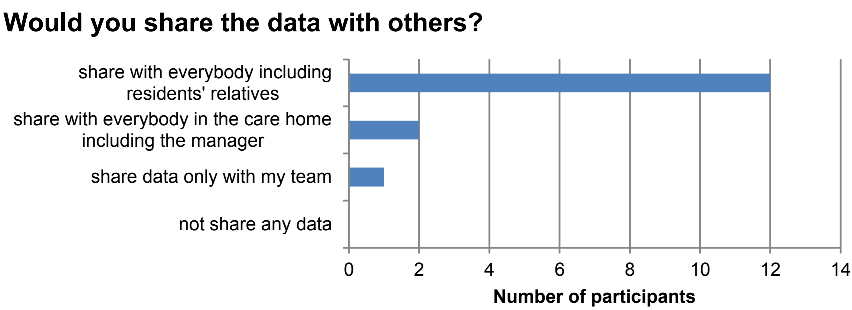


Figure 6.17: Sharing preferences by carers in the Mansfield study

The vast majority of carers of the final evaluation stated in the concluding interviews that they had no concerns about privacy and expected their data to be seen by the senior carers. Certainly this was seen as important information by the seniors to monitor individual and team performance–amounts of time given to residents, amounts of time spent with other carers, or at "locations" (e.g., the office, amount of time "unaccounted.") Whether this positive attitude towards the system is specific to the particular setting in the UK must be tested further.

6.6.5 Management Perspective

Both the owner of the Mansfield care home and the assistant manager saw huge potential in the CaReflect technology to improve care and collect evidence to discuss increased funding. However, because only a minority of residents and only some carers were equipped, these insights were not yet possible. The management was eager to test the system again with all carers and participants to gain a complete picture. Furthermore, the test period should be a full week of 5 to 7 days and the focus group should be part of the regular training days at the care home.

The concluding workshop in Mansfield was dominated by the care home manager, who attempted to derive direct insights from the gathered data. Only one experienced carer actively participated in the discussions. This carer saw a potential for CaReflect and spoke about the recorded collaboration across the team. In addition, this carer asked to use the sensor for a longer time period. The manager highlighted interesting aspects from an organizational perspective:

- Time spent on documentation across carers
- Overall effort by resident (including required time to double up)

Furthermore, care staff used the questionnaire to suggest an introduction of CaReflect to use *"for reflective practice and planning of care,"* and hope that they *"might get additional funding"* using the data. The data could also complement handovers by enabling *"communication with colleagues without having to meet face to face."*

In the concluding discussion with the manager of the Wren Hall care home after the final evaluation, the approach was seen as very positive and several ideas emerged:

- The aggregate of all individual and team reflections are seen as beneficial to the organization as a whole, for improved performance, better overview perspectives on service contacts, volumes, timings, and for coordination purposes. Access to the level and detail provided by the sensor database addresses many organizational issues.
- The question of how often staff are "doubling up" was seen as particularly important, because the government provides a higher level of subsidy for residents who require 1:1 care, or who regularly require two carers to provide proper care. Estimates are given, but they change over time, and no exact data are currently collected on a routine basis. Management wanted to know whether there are residents who are funded for 1:1, but do not get that level of care. The idea has been considered further to provide data to allow differentiated levels of funding/charging according to the levels of basic support required.
- CaReflect can be used to provide reassurance to relatives and evidence of good service. It is often said that residents will not remember staff carer visits over the previous hours, and will claim they have been alone "all day." It is also the case that for care home inspectors, "if it's not documented, it hasn't been done." Both scenarios are a source of concern for busy carers who do not have time to write everything down. The data provided by CaReflect provides this reassurance to relatives and evidence to inspectors, showing the details of the provided care–who, when, and for how long. This is seen as a major source of evidence of performance for management and individual carers–all without manual input.
- The data could also provide a perspective on care as seen by residents. The data could help carers understand how their care could be interpreted as neglect. For example, a tool could ask a questions like, "Imagine you are alone all day and see care staff only for 15 minutes per day. How would you react if your carer has no time to talk to you?"

Finally, all managers in all four care homes showed interest in a product based on the used technology. The issue of how to market this "product" was raised by the manager in Wren Hall. He wanted to know, if CaReflect is it going to be sold, rented, or part of a consultation exercise? Is it for monitoring care levels for relatives, or reflective learning for staff, or is it used as evidence for differential payments to government?

6.7 Summary

A system to measure social contacts in a care home has been developed and evaluated. We have presented an approach to turn the eZ430 Chronos hardware into a proximity sensor that captures proximity for up to 180 hours, powered by a single coin cell. The proximity sensor technology has shown potential to support reflective learning by carers in four studies, listed in Table 6.9. The sensor application and the corresponding study platform were refined in several iterations and tested in three different care homes.

The strength of CaReflect is the simplicity of the concept. Therefore, it allowed the carers, without technical background or training, to adapt the system by placing sensors not only on residents, but at places that are relevant to their work. Hence, they were able to capture data not only about their interaction with residents, but also about time spent on documentation. In this study, the number of available sensors was limited

Sensor firmware version	**Visualization and management**	**Evaluation in care home**
0.8	UnisensViewer and Chronos Management Software	3 morning shifts in Wren Hall
0.9 (increased stability)	CaReflect 0.5	3 days in Mansfield
1.0 (automatic data download)	CaReflect 0.8	3 days in Risby
1.0	CaReflect 1.0	4 days in Wren Hall

Table 6.9: Capturing prototypes for affective context

and not all ideas could be realized. Carers strongly requested more sensors to equip every resident and more places. This behavior provides insights for the design of capturing solutions. They should be easy to use and adapt. If users can adapt a solution to the needs of a workplace, they can build their own custom solution and are more engaged.

The care staff in the selected care homes gained new insights from analyzing the data. They learned about the different care needs of residents, were able to better understand the time spent on documentation, and also gained personal insights. They requested more options to compare their data with their colleagues' data, to take even more from the data analysis. Moreover, they used the data to talk to their managers about care planning and processes. Managers saw many options to use CaReflect in their homes and asked for an available product. In summary, carers and managers confirmed the positive impact of CaReflect. However, the three care homes volunteered for these studies and were convinced of their high standards. Other care homes that have more potential for improvement are most likely hesitant to quantify their problems because they fear the initial negative reaction of relatives and future customers.

CaReflect has been built for and successfully evaluated in care homes that specialize in dementia care, but the chosen approach can be applied to care homes in general.

7 Discussion

We have presented an approach to the structured design of sensor support for reflective learning. Based on the presented design space, we created and evaluated prototypes that capture affective and social context at work. This chapter critically reviews the conducted work and its implications for the design of sensor support for reflective learning.

7.1 Designing Reflection Support

The developed design space (see Chapter 4.2) provides guidance for the design of CSRL applications that capture data. The design space outlines key questions and common challenges. The current design space was used to classify the existing applications in MIRROR (see Section 3.1.1). However, the majority of the developed applications rely on self-reporting. The single other sensor-based approach in MIRROR, WATCHiT [90], complements the conducted design studies. While WATCHiT was developed for a different domain, the design space can be used without modifications as reported in [198]. However, a solution for a different domain may reveal additional aspects.

However, the design space can help only to structure the design process, designers still face challenging requirements. The optimal solution is tailored to an existing work process and provides data on topics that have been continuously debated. Designing such a solution requires an in-depth understanding of the work processes and open issues. The following aspects can prevent acceptance in the target domain in relation to a specific reflection support system:

- Costs to the employer
- Effort for the employee
- Legal constraints

Costs for the employer include the costs to implement the system, the costs to maintain it, and the costs for employees to operate it. Implementation costs include the investments to buy the required system components and install them in the workplace. In addition, employees will require training to understand, accept, and use the system. Maintenance costs include the use of external services and additional employee effort (e.g. changing batteries, backup services, or user management). The costs to operate the system include mainly the time employees must spend on the system.

Minimizing the effort for employees is crucial to gain acceptance. Reflection is not the employees' main task. There is a widespread fear of neglecting important work-related tasks if data capturing requires too much time or effort. Complex user interfaces discourage usage. Moreover, some employers advise employees to use these tools during their breaks and spare time. In these cases, few data will be recorded. In the worst case, additional management pressure will result in the generation of data. In practice, there is often a tradeoff between implementation and maintenance costs. More sophisticated systems require less interaction by employees, but a simple note-taking approach places a high effort on employees, which might prevent them from doing their original work tasks. In the best-case scenario, the desired data are already collected in existing information systems and can be made accessible for reflection. For instance, developers add new visualizations into the system or implement functions to export the data into an analysis framework.

Legal constraints limit the usage of sensors and the resulting data in many work domains. In most workplaces, recording data about a customer requires the customer's consent. Tools like the SenseCam [8] automatically take pictures of everyone in the user's vicinity, including served customers or patients. Asking for consent is at best difficult. Therefore, the law in most European countries would not allow such systems in the workplace. In the healthcare domain, patient data are protected by similar laws. It is difficult to argue why reflection tools have to access these data, because according to the law, such access must be tightly connected to the patient's well-being.

7.2 Data Quality and Risk of Misinterpretation

The quality of the recorded data is the basis for the reflection session. The developed sensing solutions deliver only a limited perspective on the experience and will always contain a small number of erroneous measurements. The design of sensor systems should filter such errors or minimize them in the first place, but it is inherently difficult to distinguish unusual data from erroneous measurements. The person that knows most about the underlying event and can distinguish errors from significant changes is the reflecting learner.

The initial ethnographic study has highlighted the challenges of measuring arousal in healthcare environments. Existing arousal algorithms that are solely based on heart rate cannot be applied in environments with high and volatile activity levels. The arousal-related components of the heart rate cannot be distinguished from the frequent changes that are due to movement. In addition, critical situations such as the emergency shown in Figure 5.6 triggered physical reactions. The fight-or-flight system is activated, but can be translated into activity. Hence, it can be debated how much arousal in these active environments matches the common notion of stress (i.e., an activation of the fight-or-flight system that is undesired). Therefore, raw heart rate data was shown to nurses and physicians, who attempted to interpret the data on their own. A subset of the participants came to insights about their stress that could not be confirmed by looking at the data. The data acted as a trigger to review own experiences during the day. However, there is a risk that the data was misinterpreted. For instance, the data may convince participants to assume they were stressed during the day. Consequently, they re-evaluate their experience and the outcome is a wrong negative perception of the day.

The proximity sensors can recognize proximity between care staff and residents. However, the data is only an indication of social contacts. The proximity range of sensors varies according to body shielding and the properties of the antenna itself. Hence, contacts might not be measured (e.g., when standing behind a resident). As a result, the proximity range was increased to 3m. Therefore, the number of false-positives increases (i.e., residents that are 3 m away are still counted as contacts). Furthermore, the constant movement of carers will continuously change the distance and the transmission range of the sensors. If the range is reduced during the one moment in the 10-second sampling interval when sensors broadcast their ID (e.g., because there is an arm in front of the sensors), no proximity will be detected. These limitations apply to all measurements and do not systematically affect particular residents. Moreover, the captured data were still understood by care staff and stimulated reflection. The sensors provide a raw picture of the processes in a care home, but it is more than was available before.

The correct introduction of the developed applications is crucial to prevent these problems. Users should be aware that errors are possible and how they occur. CaReflect benefited from its simple concept. Carers did not only learned to operate the system, but they also understood why it may not work. If users know about these errors, they will become more critical regarding the collected data and the system. However, it is the best option to prevent misinterpretation.

7.3 Impact on Learning

Reflective learning is a cognitive process that can be observed only by its outcomes or if the thoughts in the reflection process are articulated. As stated in Section 2.1, outcomes can vary from unobservable changes in attitude to a change in behavior. Questionnaires and interviews can inquire about attitude changes. Observations and sensor data can measure behavior changes.

Our questionnaires and the reported insights indicate that participants learned by analyzing their data. However, many participants may rate systems more positively in questionnaires and interviews, to please the researcher. The reported insights and the narratives during the review of the data provide more evidence. Moreover, the care homes in our studies volunteered to use sensors. According to local partners, the visited care homes deliver a high standard in their care. Care homes that struggle more to care for all their residents might provide more interesting findings but would not volunteer to reveal them. Therefore, the visited care homes are most likely particularly motivated to prove their care quality and learn about further improvement options.

During the short time of the studies no actual change in care practices was expected or measured. A second study in the same care home could have revealed such changes. However, an observable change in behavior is only one of the possible outcomes of reflection. Furthermore, the majority of the few planned changes are difficult to measure because the underlying values (e.g, on the time spent on documentation) vary from day to day depending on the behavior of residents. Moreover, measuring behavior change is difficult. Klasnja et al. [171] explain that *"demonstrating behavior change is often infeasible as well as unnecessary for a meaningful contribution to HCI research."* According to their argument, behavior change is a long-term process, which would require studies of several months or years. A time frame of multiple years does not account for the rapid pace of development in this area. This also applies to the broader field of technical-support systems that do not focus on a medical impact, but rather on the applicability of a new technology itself.

7.4 Ethical Aspects

The capturing of data at the workplace inevitably triggers ethical concerns regarding privacy and the danger of manipulating behavior against the will of employees. Designers have to be sensitive to these issues because the system is either not accepted at all or is used in an undesired manner. It is not enough to aim for a beneficial impact. Designers have to clearly see and discuss the risks as well.

Privacy is the most prominent concern, even before the introduction of a data-capturing solution. Sensors and self-reports can be abused by the management as a compulsive monitoring tool beyond legal boundaries. Even a system recognized and accepted by employees as being beneficial can turn into monitoring tool. Designers must be aware of how their tool can be used in an undesired manner. These threats have to be clearly communicated to employees and management. Therefore, it is required that employees are able to delete and manipulate data to react to pressure from the management. If the only option is to not use the system at all, employees who do not comply can easily be identified by the management.

Privacy concerns vary across domains and countries. For instance, in MIRROR [32], employees in a German hospital objected to a questionnaire that asked for their gender, while carers in a UK care home suggested placing cameras in all rooms. In general, the attempt to preserve the privacy of employees is appreciated in all workplaces. However, designers should be prepared for a wide variety of attitudes. If the opinions on privacy vary, the opinion with the strongest request for privacy should guide the implementation, because the awareness of data-capturing can and will influence the behavior of employees, especially if employees do not accept the use of such data for reflection purposes.

The ethical aspects of persuasive technology have been discussed since its inception [172, 120]. Two major guidelines for ethical persuasion systems are (a) to clearly communicate persuasion goals and (b) to select these goals carefully (i.e., to avoid persuasion that may lead to unethical goals). Berdichevsky and Neuenschwander [172] refer to this second guideline as the golden rule for the ethical design of persuasive technology, phrased as follows:

> *"The creators of a persuasive technology should never seek to persuade anyone of something they themselves would not consent to be persuaded of."*

In CSRL, there is no persuasion goal. One might assume that, as a result, the ethical challenges in persuasive technology do not apply to CSRL. However, CSRL applications use a similar (and sometimes the same) technology as persuasive technology. Furthermore, it is clear that technology is never neutral [123]. Therefore, CSRL apps may have persuasive properties that have not been explicitly designed. In this case, the ethical problems of manipulating behavior become relevant and the proposed solutions, such as clearly selecting and communicating persuasion goals, cannot be applied. This is an inherent risk that can be explored only by testing the system in the field. For instance, during a test of CaReflect, staff noted that carers devoted more time to residents that were equipped with a sensor. The carers proposed to give a sensor to all residents. This reaction outlines that sensors will direct attention towards the monitored aspects. As a result, other aspects might be neglected.

8 Conclusion

Reflective learning provides the theoretical basis to support professionals to learn from the growing amount of sensor data that are recorded in our daily life. Mobile applications and wearable sensors have been developed for nurses and physicians on a stroke unit and care staff in care homes to facilitate reflection. The systems have been evaluated with end users in their workplace setting. The lessons learned have been condensed into a design space that guides the creation of sensor based reflection support. The conducted studies have shown that the design approach creates sensor systems that support learning and reflection in workplace settings.

8.1 Summary

The goal of this thesis has been the development and evaluation of sensor-based reflection support for professionals. If professionals can combine data collected from their work with their own experiences, they can learn more quickly and effectively on the job. Computer supported reflective learning (CSRL) is a new field that so far has been based on self-reporting applications and social facilitation.

A design space has been devised to guide the conception and incremental implementation of sensor based CSRL applications. Developers of CSRL applications face the challenge to select, capture, and visualize context in a form that optimally complements the experience of professionals in order to support reflection and learning. The developed solutions have to match the requirements and constraints of the particular workplace setting. We propose an iterative development approach that structures the design process according to three decisions: the identification of the relevant context, the selection of the appropriate capturing method, and a choice of visualizations that trigger reflection and can lead to new insights on work practices. While the design space helps to identify challenges and alternative implementations for specific scenarios, developers still have to analyze carefully the specific requirements of the workplace setting.

Two design studies developed and evaluated CSRL applications that capture the affective and social context in healthcare environments. The healthcare domain has been chosen as an example because (a) reflective practice is seen as promising in this field [3] and (b) it is one of many non- or little computerized work environments. Wearable sensors have been used and adapted in a user-centered design process to the requirements of the particular workplace. The developed systems have been tested in a stroke unit and care homes to evaluate the impact in a realistic setting.

The first design study explored the capturing of the affective context of nurses and physicians on a stroke unit. An ethnographic study examined the use of wearable ECG sensors to support reflection on arousal and stress. The results outlined the challenges to capture and process affect-related data in workplace settings. A more flexible approach to collect and analyze data from multiple sensors was required. Consequently, a rapid prototyping framework for mobile systems that capture and analyze sensor signals has been developed and used to rapidly prototype two new mobile applications.

The second design study investigated the capturing of social context in care homes. A wearable proximity sensor was developed and refined in a series of studies. In comparison to existing systems, the proximity sensors are optimized for low-power consumption, the system is completely distributed, and does not require an additional infrastructure. The system estimates proximity using an adapted communication protocol that minimizes communication overhead and power consumption. CaReflect is an application that builds on this technology and is the first application that uses proximity sensors to analyze care practices. It provides quantitative data about the time spent with each resident. The system was refined and evaluated in four studies in care homes. Care staff and residents have been equipped with proximity sensors to infer the time they spent together. While the sensors cannot measure all contacts during a shift, participants confirmed the correctness of the data and discussed the findings with each other. Moreover, carers and management appreciated the sensors as versatile tools to analyze their care practices as well as measuring other practices in the care home (e.g., the amount of documentation or the time needed for medical rounds with the medical trolley).

Our evaluation results show that employees can directly understand simple visualizations of the raw recorded data. Social and affective data act as memory cues and trigger participants to narrate experiences that relate to the data. However, users must feel able to react to the gained awareness to appreciate the system. For example, the visualization of arousal and its interpretation as stress in the ECG study was perceived as a burden, but social contacts recorded by CaReflect were regarded as a new method to learn more about one's work practices. Care staff reconstructed specific situations and discussed alternative behaviors and organizational changes, such as bringing residents to the common rooms as often as possible. The outcomes have been diverse, as predicted by reflective learning theory. The insights and feedback collected during evaluation of CaReflect provides initial evidence that carers can learn from sensor data. CaReflect indeed provides content for reflection and thus turns context into content.

The evaluation of the developed sensor systems and applications has shown that sensors can support reflective learning. The recorded context can become the content of new personalized and workplace-specific learning applications. The successful implementation of such applications is challenging but the elaborated design space can guide developers to build similar systems for more use cases and more sensors.

8.2 Outlook

The number of sensors used in daily life is increasing. For example, mobile phones are equipped with more sensors to differentiate themselves from the competition. Reflective learning and learning from sensor data can build on this development to create learning support for the wide range of domains and skills, augmenting traditional learning content. Sensor-generated content is, by definition, highly specific to the domain in which it was recorded and is personalized because it is interpreted in relation to the personal experience of a learner. Further applications can be developed to leverage on this potential. Applications that require online processing of psychophysiological data on mobile devices can built on xAffect and the growing number of available xAffect components.

The CaReflect platform has shown the potential of proximity sensors in care analysis. However, the hardware of the system is still a development platform. Therefore, new platforms are being considered that are available in large quantities and include the latest low-power technology. The Bluetooth 4.0 standard is a promising candidate because of its low-energy options. An initial analysis of the power consumption of Bluetooth 4.0 and its proprietary competitor protocols [173] states that Bluetooth low-energy still needs more energy to scan for other devices. However, Bluetooth 4.0 adoption and the number of available modules are growing rapidly and the energy consumption of hardware modules are sinking. The upcoming Bluetooth 4.1 specification [174] will facilitate the implementation by introducing a dual mode topology that no longer requires a host device. Hence, Bluetooth low-energy based sensors are evaluated as an alternative technological basis for the proximity sensors.

The developed applications cover only the first step in the reflective process. Further applications could help to record the outcomes of reflection and help to sustain lessons learned by reflection. The presented approaches use a single sensor, whereas future multi-sensor approaches could lead to richer insights. Automatic data analysis methods could help to identify relevant patterns from the resulting richer set of data.

This thesis is a first step to explore how sensor data can become learning content. Capturing context can create learning content for specialized domains and situations that either cannot be addressed by traditional forms of learning or are not profitable for professional content developers. The creation of learning content from captured data is a promising new application domain for sensor technology.

List of Figures

List of Tables

Bibliography

[1] M. Eraut and W. Hirsh, *The significance of workplace learning for individuals, groups and organisations.* SKOPE, 2010.

[2] J. Dewey, *Experience and Education.* London & New York: Macmillan, 1938.

[3] M. Jasper, *Beginning Reflective Practice (Foundations in Nursing & Health Care).* Nelson Thomas Ltd, 2003.

[4] D. Boud, R. Keogh, and D. Walker, *Reflection: Turning Experience into Learning.* New York: Routledge Falmer, 1985, ch. Promoting Reflection in Learning: a Model., pp. 18–40.

[5] D. A. Kolb, *Experiential Learning: Experience as the source of learning and development.* Englewood Cliffs, N.J.: Prentice Hall, 1984.

[6] D. A. Schön, *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* Jossey-Bass, 1987.

[7] E. Isaacs, A. Konrad, A. Walendowski, T. Lennig, V. Hollis, and S. Whittaker, "Echoes from the past: How technology mediated reflection improves well-being," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2013, pp. 1071–1080.

[8] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and W. Ken, "SenseCam: A retrospective memory aid," in *Ubicom 2006*, 2006.

[9] T. Scott and S. Star, "Google Glass teardown," 2014, [accessed 20-October-2014]. [Online]. Available: http://www.catwig.com/google-glass-teardown/

[10] M. Swan, "Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0," *Journal of Sensor and Actuator Networks*, vol. 1, no. 3, pp. 217–253, 2012.

[11] K.-h. Chang, D. Fisher, and J. Canny, "Ammon: A speech analysis library for analyzing affect, stress, and mental health on mobile phones," *Proceedings of PhoneSense*, vol. 2011, 2011.

[12] A. Matic, V. Osmani, A. Maxhuni, and O. Mayora, "Multi-modal mobile sensing of social interactions," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*, may 2012, pp. 105 –114.

[13] "fitbit," fitbit, [accessed 20-October-2014]. [Online]. Available: http://www.fitbit.com/

[14] "Polar - listen to your body," Polar electo, [accessed 20-October-2014]. [Online]. Available: http://www.polar.fi

[15] T. Stiefmeier, D. Roggen, G. Ogris, P. Lukowicz, and G. Tröster, "Wearable activity tracking in car manufacturing," *IEEE Pervasive Computing*, vol. 7, no. 2, pp. 42–50, Apr. 2008.

[16] D. O. Olguin, J. A. Paradiso, and A. S. Pentland, "Wearable communicator badge: Designing a new platform for revealing organizational dynamics," in *ISWC 06*, 2006.

[17] D. A. Norman and S. W. Draper, "User centered system design," *New Perspectives on Human-Computer Interaction, L. Erlbaum Associates Inc., Hillsdale, NJ*, 1986.

[18] Y. Rogers, K. Connelly, L. Tedesco, W. Hazlewood, A. Kurtz, R. E. Hall, J. Hursey, and T. Toscos, "Why it's worth the hassle: The value of in-situ studies when designing ubicomp," in *Proceedings of the 9th International Conference on Ubiquitous Computing*, ser. UbiComp '07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 336–353.

[19] B. J. Fogg, "Creating Persuasive Technologies : An Eight-Step Design Process," in *Persuasive '09*. ACM, 2009.

[20] S. Consolvo, D. McDonald, and J. Landay, "Theory-driven design strategies for technologies that support behavior change in everyday life," in *CHI '09*. ACM, 2009, pp. 405–414.

[21] K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic, and E. Duval, "Context-aware recommender systems for learning: a survey and future challenges," *Learning Technologies, IEEE Transactions on*, vol. 5, no. 4, pp. 318–335, 2012.

[22] S. D'Mello, T. Jackson, S. Craig, B. Morgan, P. Chipman, H. White, N. Person, B. Kort, R. el Kaliouby, R. Picard *et al.*, "AutoTutor detects and responds to learners affective and cognitive states," in *Workshop on Emotional and Cognitive Issues at the International Conference on Intelligent Tutoring Systems*, 2008.

[23] P. J. Astor, M. T. Adam, P. Jerčić, K. Schaaff, and C. Weinhardt, "Integrating biosignals into information systems: A NeuroIS tool for improving emotion regulation," *Journal of Management Information Systems*, vol. 30, no. 3, pp. 247–278, 2013.

[24] M. Weiser, "The computer for the 21st century," *Scientific american*, vol. 265, no. 3, pp. 94–104, 1991.

[25] R. W. Picard, *Affective computing.* MIT press, 2000.

[26] J. A. Moon, *Reflection in learning and professional development.* Routledge, 1999.

[27] D. A. Kolb and R. Fry, *Theories of Group Processes.* London: John Whily., 1975, ch. Towards and applied theory of experiental learning., pp. 33–58.

[28] A. J. Faria, "Business simulation games: Current usage levels—an update," *Simulation & Gaming*, vol. 29, no. 3, pp. 295–308, 1998.

[29] D. M. Irby, "Three exemplary models of case-based teaching," *Academic Medicine*, vol. 69, no. 12, pp. 947–53, 1994.

[30] M. W. Daudelin, "Learning from experience through reflection," *Organizational Dynamics*, vol. 24, no. 3, pp. 36–48, 1996.

[31] L. Festinger, *A theory of cognitive dissonance.* Stanford Univ. Press, 1957.

[32] "MIRROR - reflective learning at work," The MIRROR project, [accessed 21-October-2014]. [Online]. Available: http://www.mirror-project.eu

[33] B. R. Krogstie, M. Prilla, and V. Pammer, "Understanding and supporting reflective learning processes in the workplace: The CSRL model." in *EC-TEL*, 2013, pp. 151–164.

[34] M. Prilla, V. Pammer, and S. Balzert, "The push and pull of reflection in workplace learning: Designing to support transitions between individual, collaborative and organisational learning," in *21st Century Learning for 21st Century Skills.* Springer, 2012, pp. 278–291.

[35] B. R. Krogstie, M. Prilla, D. Wessel, K. Knipfer, and V. Pammer, "Computer support for reflective learning in the workplace: A model," in *Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on.* IEEE, 2012, pp. 151–153.

[36] M. Weiser and J. S. Brown, "The coming age of calm technology," in *Beyond calculation.* Springer, 1997, pp. 75–85.

[37] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on.* IEEE, 1994, pp. 85–90.

[38] 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.* Springer, 1999, pp. 304–307.

[39] A. K. Dey, "Understanding and using context," *Personal and ubiquitous computing*, vol. 5, no. 1, pp. 4–7, 2001.

[40] A. Schmidt, M. Beigl, and H.-W. Gellersen, "There is more to context than location," *Computers & Graphics*, vol. 23, no. 6, pp. 893–901, 1999.

[41] G. D. Abowd and E. D. Mynatt, "Charting past, present, and future research in ubiquitous computing," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 7, no. 1, pp. 29–58, 2000.

[42] A. Ward, A. Jones, and A. Hopper, "A new location technique for the active office," *Personal Communications, IEEE*, vol. 4, no. 5, pp. 42–47, 1997.

[43] P. Robertson, M. Angermann, and B. Krach, "Simultaneous localization and mapping for pedestrians using only foot-mounted inertial sensors," in *Proceedings of the 11th International Conference on Ubiquitous Computing*, ser. Ubicomp '09. New York, NY, USA: ACM, 2009, pp. 93–96.

[44] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 6, pp. 1067–1080, 2007.

[45] G. Jin, X. Lu, and M. Park, "An indoor localization mechanism using active RFID tag," in *Sensor Networks, Ubiquitous, and Trustworthy Computing, 2006. IEEE International Conference on*, vol. 1. IEEE, 2006, pp. 4–pp.

[46] S. Gezici, "A survey on wireless position estimation," *Wireless Personal Communications*, vol. 44, no. 3, pp. 263–282, 2008.

[47] H. T. Friis, "A note on a simple transmission formula," *proc. IRE*, vol. 34, no. 5, pp. 254–256, 1946.

[48] *eZ430-Chronos™ Development Tool User's Guide*, Texas Instruments, [accessed 09-October-2014]. [Online]. Available: http://www.ti.com/lit/ug/slau292f/slau292f.pdf

[49] W. Ye and J. Heidemann, "Medium access control in wireless sensor networks," in *Wireless sensor networks*. Springer, 2004, pp. 73–91.

[50] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, ser. SenSys '04. New York, NY, USA: ACM, 2004, pp. 95–107.

[51] M. Buettner, G. V. Yee, E. Anderson, and R. Han, “X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks,” in *Proceedings of the 4th international conference on Embedded networked sensor systems*, ser. SenSys ’06. New York, NY, USA: ACM, 2006, pp. 307–320.

[52] F. Sivrikaya and B. Yener, “Time synchronization in sensor networks: a survey,” *Network, IEEE*, vol. 18, no. 4, pp. 45–50, 2004.

[53] S. Mann, J. Nolan, and B. Wellman, “Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments.” *Surveillance & Society*, vol. 1, no. 3, pp. 331–355, 2002.

[54] S. Mann, “Wearable computing: a first step toward personal imaging,” *Computer*, vol. 30, no. 2, pp. 25–32, Feb 1997.

[55] P. Bonato, “Wearable sensors and systems,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 29, no. 3, pp. 25–36, 2010.

[56] “Movisens - ECG- and activity sensor,” Movisens GmbH, [accessed 20-October-2014]. [Online]. Available: http://www.movisens.com/

[57] M. Poh, N. Swenson, and R. Picard, “A wearable sensor for unobtrusive, long-term assessment of electrodermal activity,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1243–1252, 2010.

[58] “Pebble smartwatch,” Pebble, [accessed 14-August-2014]. [Online]. Available: http://getpebble.com/

[59] J. Russel, “A circumplex model of affect,” *Journal of Personality and Social Psychology*, pp. 1161–1178, 1980.

[60] R. Calvo and S. D’Mello, “Affect detection: An interdisciplinary review of models, methods, and their applications,” *Affective Computing, IEEE Transactions on*, vol. 1, no. 1, pp. 18–37, Jan 2010.

[61] J. Cacioppo, L. Tassinary, and G. Berntson, *Handbook of Psychophysiology*. Cambridge University Press, 2007.

[62] R. Kocielnik, N. Sidorova, F. M. Maggi, M. Ouwerkerk, and J. H. Westerink, "Smart technologies for long-term stress monitoring at work," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on.* IEEE, 2013, pp. 53–58.

[63] C. Setz, B. Arnrich, J. Schumm, L. M. R., G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device." *IEEE Transactions on Information Technology in Biomedicine*, pp. 410–417, 2010.

[64] G. G. Berntson, *Psychophysiology.* Cambridge University Press, 1997, ch. Heart rate variability: Origins methods and interpretive caveats., pp. 623–648.

[65] Y. M. Chi, T.-P. Jung, and G. Cauwenberghs, "Dry-contact and noncontact biopotential electrodes: methodological review," *Biomedical Engineering, IEEE Reviews in*, vol. 3, pp. 106–119, 2010.

[66] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, vol. 31, no. 4, 2012.

[67] M. Brink, C. H. Müller, and C. Schierz, "Contact-free measurement of heart rate, respiration rate, and body movements during sleep," *Behavior Research Methods*, vol. 38, no. 3, pp. 511–521, 2006.

[68] P. Pelegris, K. Banitsas, T. Orbach, and K. Marias, "A novel method to detect heart beat rate using a mobile phone," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE.* IEEE, 2010, pp. 5488–5491.

[69] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz, "Heart rate variability standards of measurement, physiological interpretation, and clinical use," *European heart journal*, vol. 17, no. 3, pp. 354–381, 1996.

[70] K. Schaaff and M. T. Adam, "Measuring emotional arousal for online applications: Evaluation of ultra-short term heart rate variability

measures," in *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on.* IEEE, 2013, pp. 362–368.

[71] G. D. Clifford, F. Azuaje, and P. McSharry, "Ecg statistics, noise, artifacts, and missing data," *Advanced Methods and Tools for ECG Data Analysis*, pp. 55–99, 2006.

[72] A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti, "Cardiovascular neural regulation explored in the frequency domain." *Circulation*, vol. 84, no. 2, pp. 482–492, 1991.

[73] D. Eller-Berndl, *Herzratenvariabilität.* Verlag-Haus d. Ärzte, 2010.

[74] G. E. Billman, "The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance," *Frontiers in physiology*, vol. 4, 2013.

[75] M. Myrtek, *Heart and emotion: Ambulatory monitoring studies in everyday life.* Hogrefe & Huber Publishers, 2004.

[76] M. Kusserow, "Stress arousal monitoring in natural environments," Ph.D. dissertation, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 20236, 2012.

[77] C. Darwin, *The expression of the emotions in man and animals.* Oxford University Press, 1998.

[78] B. De Gelder, "Towards the neurobiology of emotional body language," *Nature Reviews Neuroscience*, vol. 7, no. 3, pp. 242–249, 2006.

[79] A. S. Pentland, *Honest signals.* MIT press, 2010.

[80] S. D'Mello, R. Picard, and A. Graesser, "Towards an affect-sensitive AutoTutor," *IEEE Intelligent Systems*, vol. 22, no. 4, pp. 53–61, 2007.

[81] S. Piana, A. Staglianò, A. Camurri, and F. Odone, "A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition," in *IDGEI International Workshop*, 2013.

[82] "Lumo back - quantify the invisible," Lumo BodyTech, Inc, [accessed 14-July-2014]. [Online]. Available: http://www.lumobodytech.com/lumoback/

[83] S. Mota and R. W. Picard, "Automated posture analysis for detecting learner's interest level," in *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, vol. 5. IEEE, 2003, pp. 49–49.

[84] B. Arnrich, C. Setz, R. La Marca, G. Troster, and U. Ehlert, "What does your chair know about your stress level?" *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 2, pp. 207–214, 2010.

[85] D. McDuff, A. Karlson, A. Kapoor, A. Roseway, and M. Czerwinski, "AffectAura: an intelligent system for emotional memory," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 849–858.

[86] D. Olguín Olguín, "Sociometric badges: wearable technology for measuring human behavior," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.

[87] A. Kleinsmith and N. Bianchi-Berthouze, "Affective body expression perception and recognition: A survey," *Affective Computing, IEEE Transactions on*, vol. 4, no. 1, pp. 15–33, 2013.

[88] P. Goodyear and S. Retalis, *Technology-enhanced learning*. Sense Publishers, 2010.

[89] D. Cernea, S. Mora, A. Perez, A. Ebert, A. Kerren, M. Divitini, D. G. de La Iglesia, and N. Otero, "Tangible and wearable user interfaces for supporting collaboration among emergency workers," in *Collaboration and Technology*. Springer, 2012, pp. 192–199.

[90] S. Mora and M. Divitini, "WATCHiT: Towards wearable data collection in crisis management," in *Work-in-progress at the Eight International Conference on Tangible, Embedded and Embodied Inter-action, TEI*, 2014.

[91] V. Rivera-Pelayo, J. Munk, V. Zacharias, and S. Braun, "Live interest meter: learning from quantified feedback in mass lectures," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. ACM, 2013, pp. 23–27.

[92] K. Zachos, N. Maiden, K. Pitts, S. Jones, I. Turner, M. Rose, K. Pudney, and J. MacManus, "A software app to support creativity in dementia care," in *Proceedings of the 9th ACM Conference on Creativity & Cognition*, ser. C&C '13. New York, NY, USA: ACM, 2013, pp. 124–133.

[93] M. Prilla, "User and group behavior in computer support for collaborative reflection in practice: An explorative data analysis," *COOP 2014-Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France)*, pp. 293–309, 2014.

[94] J. Gemmell, G. Bell, and R. Lueder, "MyLifeBits: A personal database for everything," *Communications of the ACM*, vol. 49, no. 1, pp. 88–95, January 2006.

[95] M. Blum, A. Pentland, and G. Troster, "Insense: Interest-based life logging," *IEEE MultiMedia*, vol. 13, no. 4, pp. 40–48, 2006.

[96] R. Fleck and G. Fitzpatrick, "Teachers' and tutors' social reflection around SenseCam images," *International Journal of Human-Computer Studies*, vol. 67, no. 12, pp. 1024–1036, 2009.

[97] A. J. Sellen and S. Whittaker, "Beyond total capture: a constructive critique of lifelogging," *Commun. ACM*, vol. 53, pp. 70–77, May 2010.

[98] I. Li, A. Dey, and J. Forlizzi, "A stage-based model of personal informatics systems," in *Proceedings of the 28th international conference on Human factors in computing systems*, ser. CHI '10. New York, NY, USA: ACM, 2010, pp. 557–566.

[99] S. T. Peesapati, V. Schwanda, J. Schultz, M. Lepage, S.-y. Jeong, and D. Cosley, "Pensieve: Supporting everyday reminiscence," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: ACM, 2010, pp. 2027–2036.

[100] S. Høyrup, "Reflection as a core process in organisational learning," *Journal of Workplace Learning*, vol. 16, no. 8, pp. 442–454, 2004.

[101] L. Mamykina, E. Mynatt, P. Davidson, and D. Greenblatt, "MAHI: Investigation of social scaffolding for reflective thinking in diabetes management," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: ACM, 2008, pp. 477–486.

[102] K. Schaaff, "Integration von physiologischem Feedback in Lernanwendungen unter Alltagsbedingungen," Ph.D. dissertation, Karlsruhe Institute of Technology, July 2014.

[103] K. Schaaff, "Enhancing mobile working memory training by using affective feedback," in *IADIS International Conference on Mobile Learning, 14-16 March 2013, Lisbon, Portugal*, 2013.

[104] J. Krumm and K. Hinckley, "The NearMe wireless proximity server," in *Ubicom 04*, 2004, pp. 283–300.

[105] M. B. Kjærgaard, G. Treu, P. Ruppel, and A. Küpper, "Efficient indoor proximity and separation detection for location fingerprinting," in *MOBILWARE '08*. ICST, 2007, pp. 1:1–1:8.

[106] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J. Pinton, and A. Vespignani, "Dynamics of person-to-person interactions from distributed RFID sensor networks," *PLOS ONE*, vol. 5, no. 7, p. e11596, 07 2010.

[107] A. Forys, K. Min, T. Schmid, W. Pettey, D. Toth, and M. Leecaster, "WRENMining: Large-scale data collection for human contact network research," in *Proceedings of First International Workshop on Sensing and Big Data Mining*. ACM, 2013, pp. 1–6.

[108] N. Banerjee, S. Agarwal, P. Bahl, R. Chandra, A. Wolman, and M. D. Corner, "Virtual compass: Relative positioning to sense mobile social interactions," in *Pervasive 04*, 2004.

[109] N. Eagle and A. Pentland, "Reality mining: sensing complex social systems," *Personal and ubiquitous computing*, vol. 10, no. 4, pp. 255–268, 2006.

[110] N. Eagle and A. Pentland, "Eigenbehaviors: Identifying structure in routine," *Behavioral Ecology and Sociobiology*, vol. 63:7, pp. 1057–1066, 2009.

[111] P. Bhagwat, B. Raman, and D. Sanghi, "Turning 802.11 inside-out," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 1, pp. 33–38, 2004.

[112] D. O. Olguin, B. N. Waber, T. Kim, A. Mohan, K. Ara, and A. Pentland, "Sensible organizations: Technology and methodology for automatically measuring organizational behavior," *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 39, no. 1, pp. 43–55, 2009.

[113] B. GmbH, "Open beacon," 2014, [accessed 15-October-2014]. [Online]. Available: http://www.openbeacon.org/

[114] A. Forys, J. Davies, A. Luong, K. Min, E. Lee, and T. Schmid, "Rapid deployable system for human contact network research," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*. ACM, 2012, pp. 383–384.

[115] D. O. Olguin and A. Pentland, "Sensor-based organisational design and engineering," *International Journal of Organisational Design and Engineering (IJODE)*, 2010.

[116] A. Gupta, M. Miettinen, M. Nagy, N. Asokan, and A. Wetzel, "Peersense: Who is near you?" in *PERCOM Workshops*, march 2012, pp. 516 –518.

[117] L. Isella, M. Romano, A. Barrat, C. Cattuto, V. Colizza, W. Van den Broeck, F. Gesualdo, E. Pandolfi, L. Ravà, C. Rizzo, and A. E. Tozzi, "Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors," *PLOS ONE*, vol. 6, no. 2, p. e17144, 02 2011.

[118] D. Knoke and S. Yang, *Social network analysis*. Sage, 2008, vol. 154.

[119] R. Fleck and G. Fitzpatrick, “Reflecting on reflection: framing a design landscape,” in *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*, ser. OZCHI ’10. New York, NY, USA: ACM, 2010, pp. 216–223.

[120] B. Fogg, *Persuasive technology Using Computers to Change What We Think and Do.* Morgan Kaufmann Publishers, 2003.

[121] J. J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H. B. Strub, “Fish’n’steps: Encouraging physical activity with an interactive computer game,” in *UbiComp 2006: Ubiquitous Computing.* Springer, 2006, pp. 261–278.

[122] J. Froehlich, T. Dillahunt, P. Klasnja, J. Mankoff, S. Consolvo, B. Harrison, and J. A. Landay, “UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2009, pp. 1043–1052.

[123] H. Oinas-Kukkonen and M. Harjumaa, “Persuasive systems design: Key issues, process model, and system features,” *Communications of the Association for Information Systems*, vol. 24, no. 1, p. 28, 2009.

[124] S. Consolvo, J. A. Landay, and D. W. McDonald, “Designing for behavior change in everyday life,” *Computer*, vol. 42, pp. 86–89, 2009.

[125] B. F. Skinner, *About behaviorism.* Knopf, 1974.

[126] E. K. Choe, N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz, “Understanding quantified-selfers’ practices in collecting and exploring personal data,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 1143–1152.

[127] I. Li, A. K. Dey, and J. Forlizzi, “Understanding my data, myself: Supporting self-reflection with ubicomp technologies,” in *Proceedings of the 13th International Conference on Ubiquitous Computing*, ser. UbiComp ’11. New York, NY, USA: ACM, 2011, pp. 405–414.

[128] M. Shaw, "The role of design spaces," *Software, IEEE*, vol. 29, no. 1, pp. 46–50, Jan 2012.

[129] D. Andrews, *The IRG Solution: Hierarchical Incompetence and How to Overcome It.* Souvenir Press, 1984.

[130] "RescueTime: Time management software to stay productive and happy in the modern workplace," RescueTime, [accessed 14-August-2014]. [Online]. Available: http://www.rescuetime.com

[131] D. V. Jaarsveld and W. R. Poster, "Call centers : emotional labor over the phone," *Emotional labor in the 21st century : diverse perspectives on the psychology of emotion regulation at work.*, 2013.

[132] J. Klerkx, K. Verbert, and E. Duval, "Enhancing learning with visualization techniques," in *Handbook of Research on Educational Communications and Technology*, J. M. Spector, M. D. Merrill, J. Elen, and M. J. Bishop, Eds. Springer New York, 2014, pp. 791–807.

[133] R. W. Robins and J. S. Beer, "Positive illusions about the self: short-term benefits and long-term costs." *Journal of personality and social psychology*, vol. 80, no. 2, p. 340, 2001.

[134] The Health and Safety Executive, "Statistics 2009/10," A National Statistics publication.

[135] M. Prince, M. Knapp, M. Guerchet, P. McCrone, M. Prina, A. Comas-Herrera, R. Wittenberg, B. Adelaja, B. Hu, D. King, A. Rehill, and D. Salimkumar, "Dementia UK second edition," Alzheimer's Society, 2014, [accessed 06-June-2014]. [Online]. Available: http://www.alzheimers.org.uk/dementiauk

[136] D. J. Brooker, *Person-centred Dementia Care: Making Services Better, Bradford Dementia Group Good Practice Guides.* Jessica Kingsley Publishers, 2007.

[137] "Personnel statistics report - A Survey of NCF Member Organisations," 2013.

[138] L. Cahill, R. J. Haier, J. Fallon, M. T. Alkire, C. Tang, D. Keator, J. Wu, and J. L. McGaugh, "Amygdala activity at encoding correlated with long-term, free recall of emotional information," *Proceedings of the National Academy of Sciences*, vol. 93, no. 15, pp. 8016–8021, 1996.

[139] G. Marcu, A. K. Dey, and S. Kiesler, "Designing for collaborative reflection," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2014 8th International Conference on*, 2014.

[140] A. Matic, A. Popleteev, S. Gabrielli, V. Osmani, and O. Mayora, "Happy or Moody? Why so? Monitoring Daily Routines at Work and Inferring Their Influence on Mood," *5th UbiHealth Workshop in conjunction with UBICOMP 2010 Conf., Copenhagen, Denmark*, 2010. [Online]. Available: http://www.create-net.org/ubint/ubihealth/papers/Matic_HappyOrMoody.pdf

[141] W. H. Rickles and J. L. Day, "Electrodermal activity in non-palmer skin sites," *Psychophysiology*, vol. 4, no. 4, pp. 421–435, 1968.

[142] "Unisens - a universal data format for multi sensor data," [accessed 21-October-2014]. [Online]. Available: http://www.unisens.org

[143] D. R. Millen, "Rapid ethnography: time deepening strategies for HCI field research," in *3rd conference on Designing interactive systems*, New York, 2000.

[144] D. M. Fetterman, *Ethnography Step by Step, London.* SAGE Publications Ltd, 1999.

[145] B. Jordan, "Ethnographic workplace studies and CSCW," *Human Factors in Information Technology*, vol. 12, pp. 17–42, 1996.

[146] P. Sanches, K. Höök, E. Vaara, C. Weymann, M. Bylund, P. Ferreira, N. Peira, and M. Sjölinder, "Mind the body!: designing a mobile stress management application encouraging personal reflection," in *Proceedings of the 8th ACM Conference on Designing Interactive Systems.* ACM, 2010, pp. 47–56.

[147] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface

(BCI) system," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.

[148] "Reports August 2014 US smartphone subscriber market share," comScore, Inc., [accessed 09-October-2014]. [Online]. Available: http://www.comscore.com/Insights/Market-Rankings/comScore-Reports-August-2014-US-Smartphone-Subscriber-Market-Share

[149] P. Paredes, D. Sun, and J. Canny, "Sensor-less sensing for affective computing and stress management technology," in *MindCare Workshop at the Pervasive Health 2013 Conference*, 2013.

[150] D. Sun, P. Paredes, and J. Canny, "Moustress: Detecting stress from mouse motion," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 61–70.

[151] J. Hernandez, P. Paredes, A. Roseway, and M. Czerwinski, "Under pressure: Sensing stress of computer users," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 51–60.

[152] "Android Wear," Google, [accessed 09-October-2014]. [Online]. Available: http://www.android.com/wear/

[153] Y.-S. Kuo, S. Verma, T. Schmid, and P. Dutta, "Hijacking power and bandwidth from the mobile phone's audio interface," in *Proceedings of the First ACM Symposium on Computing for Development*. ACM, 2010, p. 24.

[154] P. S. Hamilton, "Open source ECG analysis software documentation," Online, E.P. Limited, 2002, [accessed 20-October-2014]. [Online]. Available: http://www.eplimited.com

[155] "Mirror spaces framework," The MIRROR project, [accessed 21-October-2014]. [Online]. Available: https://github.com/MirrorIP/msf

[156] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[157] J. Gehring, “Android GraphView,” 2014, [accessed 21-October-2014]. [Online]. Available: http://android-graphview.org

[158] M. Prince, M. Prina, and M. Guerchet, “World alzheimer report 2013,” Alzheimer’s Disease International, 2013, [accessed 20-October-2014]. [Online]. Available: http://www.alz.co.uk/research/worldreport/

[159] N. Graham and J. Warner, *Understanding alzheimer’s disease and other dementias.* Family Doctor Publications Ltd, 2009.

[160] D. Brooker and C. Surr, “Dementia care mapping (DCM): initial validation of DCM 8 in UK field trial,” *International Journal of Geriatric Psychiatry*, vol. 21, no. 11, pp. 1018–1025, 2006.

[161] “Short observational framework for inspection 2,” [accessed 11-February-2013]. [Online]. Available: http://www.brad.ac.uk/health/dementia/dcm/sofi/

[162] D. Brooker and C. Surr, *See me, not just the dementia: Understanding peoples‘ experiences of living in a care home.*, CSCI, Ed. London: Commission for Social Care Inspection, 2008.

[163] “The butterfly scheme,” online, [accessed 06-June-2014]. [Online]. Available: http://butterflyscheme.org.uk/

[164] K. Furset and P. Hoffman, *High pulse drain impact on CR2032 coin cell battery capacity*, Nordic Semiconductor, [accessed 21-October-2014]. [Online]. Available: http://cdn.eetimes.com/ContentEETimes/Documents/Schweber/C0924/C0924post.pdf

[165] *MSP430™ SoC With RF Core*, Texas Instruments, [accessed 09-October-2014]. [Online]. Available: http://www.ti.com/lit/ds/symlink/cc430f5137.pdf

[166] *Low-Power Sub-GHz RF Transceiver*, Texas Instruments, [accessed 09-October-2014]. [Online]. Available: http://www.ti.com/lit/ds/symlink/cc1100e.pdf

[167] *Introduction to SimpliciTI*, Texas Instruments, [accessed 09-October-2014]. [Online]. Available: http://www.ti.com/lit/ml/swru130b/swru130b.pdf

[168] *CC430 Family User's Guide*, Texas Instruments, [accessed 21-October-2014]. [Online]. Available: http://www.ti.com/lit/ug/slau259d/slau259d.pdf

[169] *CC430 Family User's Guide*, Texas Instruments, [accessed 09-October-2014]. [Online]. Available: http://www.ti.com/lit/ug/slau259e/slau259e.pdf

[170] F. F. Reichheld, "The one number you need to grow," *Harvard business review*, vol. 81, no. 12, pp. 46–55, 2003.

[171] P. Klasnja, S. Consolvo, and W. Pratt, "How to evaluate technologies for health behavior change in HCI research," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2011, pp. 3063–3072.

[172] D. Berdichevsky and E. Neuenschwander, "Toward an ethics of persuasive technology," *Commun. ACM*, vol. 42, no. 5, pp. 51–58, May 1999.

[173] *Wireless connectivity guide*, Texas Instruments, [accessed 20-October-2014].

[174] *Bluetooth 4.1 features & technical descriptions*, Bluetooth Special Interest Group, [accessed 10-October-2014]. [Online]. Available: http://www.bluetooth.org/en-us/Documents/Bluetooth%204.1%20Technical%20Description.pdf

Supervised Student Research

[175] S. Baumeister, “Eine Tablet-Anwendung zur Förderung des Selbstmanagements in der Schlaganfallnachsorge,” Diplomarbeit, Karlsruher Institut für Technologie, 2012.

[176] T. Efimova and M. Zeier, “Emotionserkennung aus der Sprache,” Seminararbeit - Technologiegestütztes Lernen, Karlsruher Institut für Technologie, 2012.

[177] J. Eichler and L. Gribel, “Watch-based Learning by Reflection,” Seminararbeit - Technologiegestütztes Lernen, Hochschule Karlsruhe - Technik und Wirtschaft, 2012.

[178] Y. Hilgers, J. Schneider, and S. Walter, “Reflexion am Arbeitsplatz,” Seminararbeit - Technologiegestütztes Lernen, Karlsruher Institut für Technologie, 2011.

[179] L. Jornitz, “Anbindung von Sensoren an Android über die Audioschnittstelle,” Bachelor Thesis, Karlsruher Institut für Technologie, 2011.

[180] L. Krueßen, “Konzeption und Entwicklung einer Augmented-Reality-Plattform für Bedienungsanleitungen unter Android,” Diplomarbeit, Karlsruher Institut für Technologie, 2012.

[181] H. Medina, N. Plessing, and S. Stehmann, “Persuasive Technology From the lifestyle we have to the lifestyle we want,” Seminararbeit - Technologiegestütztes Lernen, Karlsruher Institut für Technologie, 2012.

[182] D. Pevoubou, “Konzeption und Entwicklung eines Psychophysiologischen Messsystems für Telekonferenzen,” Master Thesis, Universität Heidelberg, 2012.

[183] F. Riedle, S. Karrer, and N. Wirkuttis, “Augmented Reality im Umfeld von Ambient Assisted Living,” Seminararbeit - Ambient Assisted Living, Karlsruher Institut für Technologie, 2011.

[184] V. Schmidt-Bremme, A. Schmid, and A. Schlereth, “Lernen mit der Smart Watch,” Seminararbeit - Technologiegestütztes Lernen, Karlsruher Institut für Technologie, 2013.

[185] M. Sonnentag, “Low-Power Proximity Sensorik auf Basis einer programmierbaren Armbanduhr mit Funkmodul,” Bachelor Thesis, Hochschule Karlsruhe - Technik und Wirtschaft, 2011.

[186] A. Syndikus, “Analyse und Visualisierung der sozialen Netzwerke in der Pflege auf Basis von Sensordaten,” Bachelor Thesis, Karlsruher Institut für Technologie, 2013.

[187] S. Tauch, “Designing and Developing an Interactive Augmented Reality Editor for Digital Manuals,” Bachelor Thesis, Karlsruher Institut für Technologie, 2013.

[188] A. Übelhör, “Konzeption und Entwicklung eines mobilen Assistenten für die Facharztausbildung,” Diplomarbeit, Karlsruher Institut für Technologie, 2011.

[189] A. Weber, A. Miclaus, and O. Löwen, “Lernen in Alltagsumgebungen,” Seminararbeit - Technologiegestütztes Lernen, Karlsruher Institut für Technologie, 2013.

[190] P. S. Weber, “Konzeption und Entwicklung eines tragbaren Sensorsystems für die Erkennung der Körperhaltung,” Master Thesis, Karlsruher Institut für Technologie, 2013.

Publications

[191] A. Fessl, V. Rivera-Pelayo, L. Müller, V. Pammer, and S. Lindstaedt, "Motivation and user acceptance of using physiological data to support individual reflection," in *2nd International Workshop on Motivational and Affective Aspects in Technology Enhanced Learning*, 2011.

[192] R. Görlitz, L. Müller, and A. Rashid, "An individual patient-oriented stroke manager architecture," in *Proceedings of th eHealth 2012*, pp. 116–121.

[193] T. Holocher-Ertl, C. Kunzmann, L. Müller, V. Rivera-Pelayo, and A. P. Schmidt, "Motivational and affective aspects in technology enhanced learning: Topics, results and research route," in *EC-TEL 2013, Scaling up Learning for Sustained Impact.* Springer Berlin Heidelberg, 2013, pp. 460–465.

[194] N. Maiden, S. D'Souza, S. Jones, L. Müller, L. Pannese, K. Pitts, M. Prilla, K. Pudney, M. Rose, I. Turner *et al.*, "Computing technologies for reflective, creative care of people with dementia," *Communications of the ACM*, vol. 56, no. 11, pp. 60–67, 2013.

[195] S. Mora, V. Rivera-Pelayo, and L. Müller, "Supporting mood awareness in collaborative settings," in *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2011 7th International Conference on.* IEEE, 2011, pp. 268–277.

[196] L. Müller, "Pervasive monitoring to support reflective learning," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication.* ACM, 2013, pp. 349–354.

[197] L. Müller, I. Aslan, and L. Krüßen, "GuideMe: A mobile augmented reality system to display user manuals for home appliances," in *Advances in Computer Entertainment.* Springer International Publishing, 2013, pp. 152–167.

[198] L. Müller, M. Divitini, S. Mora, V. Rivera-Pelayo, and W. Stork, "Context becomes content: Sensor data for computer-supported reflective learning," *IEEE Transactions on Learning Technologies*, 2014.

[199] L. Müller, B. Krogstie, and A. Schmidt, "Towards capturing learning experiences," in *Context and Technology Enhanced Learning (ConTEL): Theory, methodology and design, ECTEL 2011, Palermo, Italy*, 2011.

[200] L. Müller, V. Rivera-Pelayo, and S. Heuer, "Persuasion and reflective learning: closing the feedback loop," in *Persuasive Technology. Design for Health and Safety.* Springer, 2012, pp. 133–144.

[201] L. Müller, V. Rivera-Pelayo, C. Kunzmann, and A. Schmidt, "From stress awareness to coping strategies of medical staff: supporting reflection on physiological data," in *Human Behavior Understanding.* Springer, 2011, pp. 93–103.

[202] L. Müller, M. Sonnentag, and S. Heuer, "Supporting reflection on dementia care using proximity sensors," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on.* IEEE, 2013, pp. 89–92.

[203] V. Rivera-Pelayo, V. Zacharias, L. Müller, and S. Braun, "Applying quantified self approaches to support reflective learning," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.* ACM, 2012, pp. 111–114.

[204] V. Rivera-Pelayo, V. Zacharias, L. Müller, and S. Braun, "A framework for applying quantified self approaches to support reflective learning," 2012, p. 123.

[205] K. Schaaff, L. Müller, M. Kirst, and S. Heuer, "Improving financial decision making using psychophysiological sensor data," in *Psychophysiology*, vol. 48. Wiley-Blackwell, MA USA, 2011, pp. 87–87.

[206] K. Schaaff, L. Müller, M. Kirst, and S. Heuer, "xAffect – a modular framework for online affect recognition and biofeedback applications," in *Proceedings of the 3rd international workshop on motivational and affective aspects of technology enhanced learning*, 2012.