2026/018

Confidential
☐ Yes
☑ No

Designing a Smartwatch-Based Interface for Wearable Experience Sampling in Health Data Collection

Tianqin Lu

Supervised by

Jun Hu

2026

EngD Designing Human-System Interaction

DESIGNING
HUMAN-SYSTEM
INTERACTION

TU/e
EINDHOVEN
UNIVERSITY OF
TECHNOLOGY

The design described in this thesis has been carried out in accordance with the TU/e Code of Scientific Conduct.

Thesis Examination Committee:

Scientific supervisor(s):	Dr. EngD MEng Jun Hu, TU/e
Company representative(s):	Dr. Rong-Hao Liang, TU/e
Independent members:	Dr. Pieter van Gorp, TU/e
	Prof. Dr. Aarnout Brombacher, TU/e
Chair:	Prof. Dr. Regina Bernhaupt, TU/e

A catalogue record is available from the Eindhoven University of Technology library.

ISBN:

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Jun Hu. Over the past few years, from my master's studies to this EngD journey, working under your guidance has been an invaluable experience. Thank you for generously sharing your knowledge, expertise, experience, and for supporting not only the completion of this thesis but also my professional and personal development.

I would also like to acknowledge the STRAP (Self TRAcking for Prevention and diagnosis of heart disease) consortium for funding this EngD project. Their support made this research possible and provided the resources necessary to carry out the work presented in this thesis.

My sincere thanks go to Dr. Rong-Hao Liang, the case owner, for his expertise and guidance during the development of the prototype. Your insights on user interface design and rapid prototyping were instrumental in shaping the system and ensuring its usability. I also thank the research group at Maastricht University for their collaboration and valuable input throughout the project.

I am grateful to the members of my examination committee for their time, review, and feedback, which significantly improved the quality of this thesis.

My thanks go to the GameBus team for their support and willingness to share their expertise throughout this project. In particular, I would like to thank Olga Glazunova for her help with everything from watch setup to platform integration. Special thanks to all participants involved in the usability studies, whose input and feedback contributed to the outcomes of this project.

I would like to acknowledge my EngD colleagues. It has been a pleasure to work alongside you and to share this creative and supportive environment.

Finally, I would like to thank my family, partner, and friends for their support and encouragement throughout my study.

EXECUTIVE SUMMARY

This EngD project presents the design, implementation, and evaluation of a smartwatch-based experience sampling (ESM) interface to support health data collection. The work was conducted in the context of a genito-pelvic pain (GPP) study at Maastricht University, which required frequent self-reporting alongside physiological sensing in daily life.

The project resulted in a user-friendly smartwatch-based ESM interface. It was implemented as a custom front-end in *Experiencer* (the smartwatch ESM client application) and integrated with the *GameBus* backend for study configuration and secure data management. The resulting prototype enables low-effort, microinteraction-based self-reporting and implements the required ESM protocol logic. Passive sensor data collection is supported as well, and data capture and transfer were verified end-to-end. The design was deployed on the Samsung Galaxy Watch7 running Wear OS.

Usability and user experience were assessed in a controlled, single-session laboratory study with healthy participants ($N = 12$). To evaluate the interface without exposing participants to sensitive clinical content, the original GPP items were replaced with an English proxy questionnaire that preserved the interaction structure of the clinical version (response formats and branching logic). Participants were able to complete the self-report tasks efficiently and without prior training, and the system achieved a high usability rating. The evaluation also revealed usability challenges: backward navigation was not always discoverable for first-time users, and longer labels reduced readability on the small circular display. The results also suggested that high usability alone does not guarantee willingness to engage in frequent ESM; perceived relevance and protocol design strongly influence anticipated sustained participation.

Overall, this thesis demonstrates that smartwatch-based ESM can be supported through careful front-end interaction design. It recommends prioritizing self-explanatory, linear interaction flows; making response commitment and correction explicitly discoverable; treating text length as a primary constraint for smartwatch deployment; maintaining a visually neutral design while providing functional feedback (e.g., subtle haptics); and validating the system further through longitudinal, in-the-wild evaluation with target user groups.

In summary, this thesis contributes a functional smartwatch-based experience sampling prototype, with an end-to-end verified data pipeline. It also contributes usability evidence and practical design implications for deploying experience sampling applications on circular smartwatches, supporting future wearable ESM studies and multi-dimensional health data collection.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

This thesis presents the design, implementation, and evaluation of a smartwatch-based user interface (UI) for experience sampling that supports brief, low-effort self-reporting through wearable microinteractions. The work focuses on front-end interaction design for smartwatches and integrates with a backend research platform, GameBus[1], for study configuration and secure data management.

This chapter outlines the background and relevance of the project, provides a problem analysis, and defines the goals and scope of the work. It also presents the structure of the remaining chapters.

1.1 Background

Recent health monitoring research increasingly complements episodic, clinic-based assessment with continuous, in-the-wild measurement, because many health-relevant processes fluctuate across daily contexts and are difficult to capture reliably in laboratory settings (Butler et al., 2025; Löchner et al., 2024; Roos & Slavich, 2023). This shift emphasizes the need for user-friendly, ecologically valid instruments that enable continuous assessment of subjective experiences and related physiological data in real-life contexts (Rashid & Nemati, 2024; Shen et al., 2025). One widely used methodology for such in-situ data collection is the Experience Sampling Method (ESM). ESM is a structured diary technique in which individuals are prompted throughout the day to report on their thoughts, feelings, and behaviors (Myin-Germeys et al., 2009).

With the growing availability of consumer wearables, smartwatches have become an effective platform for real-world health assessment. They support continuous passive sensing of physiological and behavioral signals (e.g., heart rate and movement) while enabling short, low-burden self-reports through microinteractions (Intille et al., 2016; Volsa et al., 2022). These affordances make smartwatches well-suited to studies that require repeated measurements across the day.

This thesis focuses on designing a user-friendly smartwatch-based interface that supports microinteraction-based ESM. To ground and evaluate the design, we sit-

---

[1] https://blog.gamebus.eu/

uate it within a concrete and demanding application context: a ZonMw-funded research project on genito-pelvic pain (GPP), conducted by a research team from the Faculty of Psychology and Neuroscience at Maastricht University. This project requires high-frequency, real-time self-reports in combination with physiological data collection.

GPP, also known as penetration disorder, is a prevalent and complex condition involving ongoing or repeated difficulties with vaginal penetration, pain during intercourse, anticipatory fear or anxiety, and involuntary contraction of the pelvic floor muscles (Black & Grant, 2014). It is associated with significant psychological distress, interpersonal challenges, and a diminished quality of life (Bergeron et al., 2015), and remains largely underdiagnosed due to its complex etiology (Zarski et al., 2025). Another contributing factor is the lack of comprehensive, real-time assessment instruments capable of capturing multiple psychophysiological signals in daily life, e.g., combining heart rate, accelerometry, and self-reports (May et al., 2018; Seidman et al., 2020).

While the GPP project provided a concrete and demanding real-world context to ground and evaluate the UI and its associated system, the resulting smartwatch-based ESM solution is not specific to GPP and is intended to support broader longitudinal health tracking. This thesis contributes by designing, implementing, and evaluating the user experience of a smartwatch-based ESM system built on Experiencer. This work was conducted as part of the STRAP (Self-Tracking for Prevention and Diagnosis of Heart Disease) consortium, supporting its goal of developing low-cost, low-burden data-collection methods that enable continuous monitoring and support preventive healthcare.

1.2 Problem analysis

Current clinical approaches for assessing health conditions, such as retrospective questionnaires and structured interviews, are limited by recall bias, loss of situational context, and inability to capture fluctuations over time. These limitations can lead to incomplete or inaccurate representations of individuals' lived experiences (Haase, 2023; Previtali et al., 2022), often resulting in overly generalized approaches to diagnosis and treatment. To better understand underlying mechanisms in various health contexts, more sophisticated and temporally sensitive assessment methods are required.

ESM offers an ecologically valid and detailed alternative by capturing real-time self-reports in naturalistic settings. Compared to retrospective reporting, ESM provides more accurate, context-sensitive insights into dynamic psychological and behavioral processes (Ellison et al., 2020; Fillingim et al., 2016). However, effectively deploying ESM in everyday life, particularly at high sampling frequencies, places substantial demands on both the technology and the interaction design.

As health research increasingly adopts comprehensive models that emphasize interactions among psychological, physiological, and contextual factors, assessment toolkits must evolve accordingly. In particular, there is a growing need for systems capable of simultaneously capturing subjective experiences and physiological responses in both experimental and everyday settings. Wearable technologies, particularly smartwatches, are well-suited to this purpose. They enable continuous physiological sensing through embedded sensors alongside in-the-moment self-reports, supporting more process-oriented and ecologically valid analyses of health-related dynamics in daily life (Arias et al., 2023; Avila et al., 2021).

However, many existing ESM solutions are not designed around the interaction constraints of smartwatches. Small screens, limited input bandwidth, and interrupt-driven use make the user interface a key determinant of whether high-frequency ESM is feasible in practice. If responding to prompts is slow, error-prone, or disruptive, compliance and data quality are likely to suffer. The GPP study illustrates this challenge: it requires frequent, low-burden self-reporting alongside physiological data collection, which in turn demands a smartwatch-based ESM solution that is both usable and robust across repeated daily use.

1.3 Project goals and scope

The goal of this EngD project was to deliver a functional smartwatch-based experience sampling prototype that supports in-situ self-reporting in everyday life. The prototype was implemented on Samsung smartwatches running Wear OS (Google's Android-based smartwatch operating system) using *Experiencer*, an open-source wearable ESM application (Khanshan et al., 2023), and integrated with *GameBus* as the backend for study configuration and secure data management.

While the delivered outcome is a working end-to-end setup (including study configuration, questionnaire delivery, data transmission, and data export), the primary focus of this thesis is the design and evaluation of the smartwatch front-end: a new UI for Experiencer that enables brief, low-effort microinteractions on circular smartwatch displays. The interface was designed to support high-frequency prompting and to reduce interaction burden through a simple, glanceable interaction flow.

By enabling real-time, in-situ reporting, the prototype aims to support users in reporting symptoms, emotions, and daily context that may contribute to or alleviate their experiences. In the broader research context, this capability can support longitudinal, multi-dimensional data collection that is useful for developing personalized models in GPP research and in other health domains that require high-frequency self-reporting combined with physiological sensing.

The scope of this thesis is the design, implementation, and evaluation of the smartwatch-based experience sampling UI and its deployment within the Experiencer–GameBus ecosystem. While informed by a use case involving women with GPP and their partners, the evaluation reported in this thesis is limited to usability testing with healthy participants using ethically appropriate proxy questionnaire items.

This development process could lead to the following outcomes:

- A user-friendly smartwatch interface integrated with the *Experiencer* application for collecting participants' self-reports.
- A fully configured ESM protocol and sensor setup suitable for deployment in field studies.
- Usability insights gathered from healthy participants to inform further refinements.
- Discussion of design implications for subsequent studies.

The following are out of scope:

- Clinical trials involving participants with specific health conditions, such as GPP, as this project is limited to evaluation with healthy participants.
- Hardware development or modifications to Experiencer or GameBus, beyond adapting existing features for this project's requirements (see Subsection 4.4 for details on Experiencer and GameBus).
- Validation of sensor data collection regarding accuracy, reliability, or clinical interpretation.

1.4 Outline

The remainder of this thesis is structured as follows. Chapter II describes the project management, including planning, risks, and stakeholder engagement. Chapter III reviews the related work on ESM for health data collection and designing wearable interfaces for ESM. Chapter IV presents the design and development phases, including requirements gathering, initial prototype creation, iterative refinements, and the implementation of the final smartwatch UI. Chapter V details the evaluation, covering the participant recruitment, data collection, and results, followed by a discussion of the findings. Finally, Chapter VI concludes the thesis and reflects on its contributions.

Chapter 2

PROJECT MANAGEMENT

This chapter describes how the project was structured and managed to support the timely delivery of a functional smartwatch-based experience sampling prototype. Project management focused on defining clear phases, milestones, and deliverables, and on maintaining alignment with key stakeholders throughout requirements clarification, implementation, and evaluation.

2.1 Process Overview and Timeline

This project was planned and executed as a design and engineering process. To ensure coherent progress and alignment with stakeholders, the work was structured into a sequence of phases spanning initial exploration, design and development, evaluation, and reflection. The design & development phase included an internal workflow of task clarification, conceptual design, implementation, and technical verification, which is documented in Section 4.1. Table 2.1 summarizes the phases, milestones, and deliverables.

The project started with a **start-up phase**, in which the design challenge was defined, and the goals and scope of the project were established. During this phase, initial stakeholder alignment was achieved, and a project plan was formulated to guide subsequent activities.

This was followed by an **analysis phase**, which focused on establishing the research and design context. A literature review was conducted to position the project within existing work on experience sampling, wearable interaction, and health monitoring. This phase identified knowledge gaps and informed the framing of design requirements.

The project then moved into a **design and development phase**. In this phase, interface ideas and concepts were explored and a smartwatch-based experience sampling solution was designed and implemented. This phase encompassed requirements clarification, conceptual design, and technical realization of a functional prototype.

Next, an **evaluation phase** was conducted to assess the usability and user experience of the implemented prototype. This included the design and execution of a user study, as well as data collection and analysis to evaluate interaction quality.

Finally, the project concluded with a **reflection and reporting phase**. In this phase, the design process, implementation outcomes, and evaluation results were consolidated and documented in this thesis. Reflections on limitations and design implications for future work were also formulated.

Table 2.1: Project phases with milestones and deliverables used for planning and monitoring progress.

Phase	Milestones (progress checkpoints)	Key Deliverables (reviewable outputs)
Start-up	Project goals and scope agreed; initial stakeholder alignment established.	Problem statement and project objectives; stakeholder map.
Analysis	Relevant prior work reviewed; design constraints and implications for wearable ESM clarified.	Related work with initial design considerations and constraints.
Design & Development	Requirements formed; interaction concept selected; working prototype implemented and technically verified end-to-end.	Requirements specification; concept and decision matrix; implemented smartwatch prototype; technical verification results.
Evaluation	Usability study completed; findings consolidated into actionable interface observations.	Study protocol and instruments; evaluation results and discussion.
Reflection & Reporting	Thesis finalized; limitations and design implications identified.	Final thesis document; summarized design implications and recommendations.

The detailed project timeline (including task durations and dependencies) is provided in Appendix A.

2.2 Risk Management

This project involved uncertainties arising at different stages of the design and development process. These included delayed availability of the full ESM protocol during design, interaction constraints during interface design, platform dependencies during implementation, and limitations related to evaluation scope. Table 2.2 summarizes the main risks and the mitigation strategies applied.

Protocol and Design Uncertainty (R1 & R2): A primary risk during the early design phase was the unavailability of the full ESM protocol. To avoid delaying progress, initial interface ideas were developed using representative generic ESM

Table 2.2: Key risks identified in the project and the mitigation strategies.

ID	Risk Description	Mitigation Strategy
R1	Late availability of the full ESM protocol during early design exploration due to dependency on external study stakeholders (timing outside the project's direct control).	Early interface concepts were developed using representative placeholder questions and revisited once the final protocol became available.
R2	Questionnaire length and wording exceeding smartwatch display constraints.	Reviewed with domain experts, and shortened questionnaire items where necessary without altering meaning.
R3	Dependence on existing Experiencer–GameBus functionality and hardware support.	Design and implementation were constrained to existing platform capabilities, without modifying core components.
R4	Evaluation conducted with healthy participants using proxy items rather than the clinical target population.	The evaluation scope was limited to usability, and findings were interpreted accordingly.

questions. Once the complete questionnaire became available, the designs were reviewed and adapted to align with the final protocol.

A related risk was that the lengthy clinical questionnaire items would be incompatible with the physical constraints of a circular smartwatch display (e.g., text clipping or excessive scrolling). To address this risk, long questionnaire items were discussed with the Maastricht University research team and shortened where necessary without altering their intended meaning.

Technical Dependencies (R3): To manage risks related to the Experiencer–GameBus ecosystem, the system was designed to build upon existing platform functionality rather than modifying core components. Technical feasibility was continuously verified through documentation and direct support from the GameBus team to prevent integration failures.

Evaluation Scope (R4): A limitation of the project relates to the scope of the usability evaluation. Due to accessibility constraints, evaluation was conducted with healthy participants using proxy questionnaire items, rather than with the target clinical population and full ESM protocol. This constraint was accepted as part of the project scope. To mitigate potential misinterpretation of data, findings were interpreted as usability insights into the UI, rather than as evidence of clinical validity or long-term deployment performance.

2.3 Stakeholder Engagement

Stakeholder engagement was structured according to the different roles and levels of influence identified in the stakeholder map (Figure 2.1). Engagement activities primarily focused on task clarification, iterative design feedback, and ensuring technical feasibility within the existing Experiencer–GameBus ecosystem.

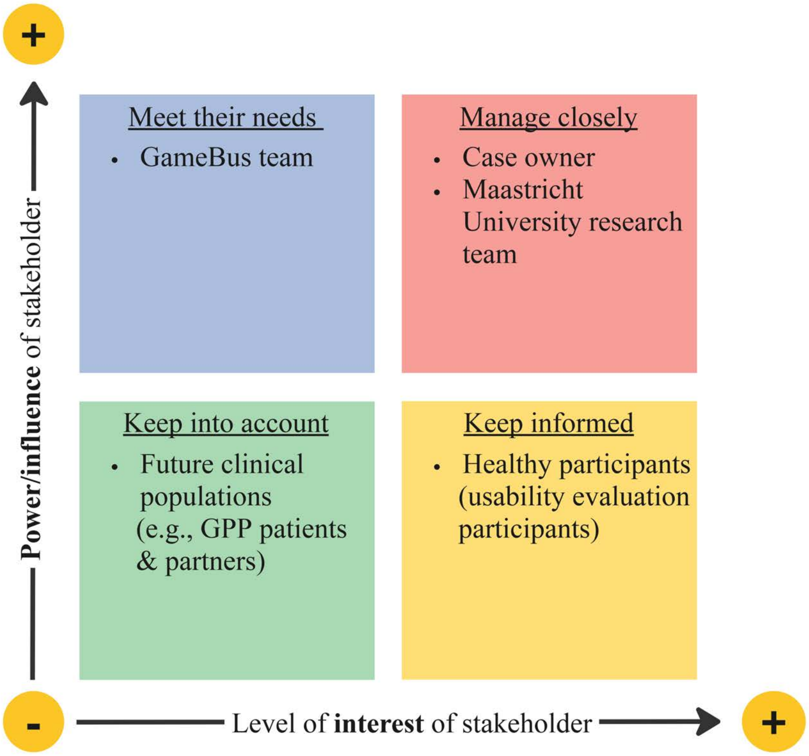


Figure 2.1: Stakeholder map, showing stakeholders positioned according to their level of influence on project decisions and their level of interest in the project outcomes.

The Maastricht University research team, including the case owner, was engaged throughout the project and constituted the primary stakeholders with high influence and high interest. Their involvement was essential in defining the design challenge, project goals, and study constraints. The team provided the complete ESM protocol for the use case, which formed the basis for the functional and technical requirements described in Section 4.2. During the concept and prototyping phases, iterative meetings were held to review interface concepts, discuss adaptations of questionnaire items for presentation on a smartwatch display, and assess the suitability of interaction designs for experience sampling. The case owner also contributed domain expertise in wearable UI design and rapid prototyping, providing targeted feedback on interaction concepts and design trade-offs that influenced key design decisions.

The GameBus team, including developers and researchers associated with both the GameBus platform and the Experiencer application, was engaged primarily in a supportive role. Interaction focused on device supply, platform usage, training, and clarification of existing functionality rather than joint design activities. The team provided technical support related to hardware deployment, study configuration in GameBus Campaigns, recently added platform features such as branching logic, and data extraction. In parallel, the author provided feedback to the GameBus team regarding limitations and issues encountered during implementation and desired functionality. This bidirectional exchange informed implementation decisions while staying within the boundaries of the existing platform and without modifying core GameBus or Experiencer components.

Healthy participants were involved exclusively during the usability evaluation phase. Their engagement was limited to interacting with the implemented smartwatch interface using ethically appropriate proxy questionnaire items. Participants provided feedback through structured usability testing. As these participants did not influence system requirements, design scope, or architectural decisions, they are positioned as stakeholders with high interest but low influence. Future clinical populations, such as individuals with GPP and their partners, were not directly involved in the design or evaluation phases of this project. However, their needs were considered indirectly through the underlying use case and ESM protocol provided by the Maastricht University research team.

Chapter 3

RELATED WORK

This chapter presents the prior research relevant to the design and development of a smartwatch-based ESM application for longitudinal health monitoring. The related work spans three domains: (1) the limitations of retrospective measurement and in-situ health assessment, (2) the role of ESM in health data collection with a particular focus on wearable ESM, and (3) interface design considerations for ESM on smartwatch platforms.

3.1 Limitations of Retrospective Measures and In-situ Health Assessment

Prior work on various health conditions, such as chronic pain disorders, highlights the complex and multidimensional nature that often fluctuates in intensity and is influenced by contextual factors, such as activity, emotional state, and interpersonal dynamics (Bergeron et al., 2015). These characteristics pose a challenge for traditional health assessment methods, which typically rely on retrospective questionnaires or clinical interviews administered at infrequent intervals.

Retrospective assessments rely heavily on patients' memory over extended periods, making them prone to recall bias (Lewandowski et al., 2009; Previtali et al., 2022). Research in chronic pain populations has shown that individuals tend to recall peak pain episodes more vividly than typical or average experiences, often inflating retrospective self-reports compared to real-time data (Previtali et al., 2022). As a result, retrospective assessments may not fully capture the complexity of such conditions, failing to reflect the day-to-day experience.

To address these limitations, in-situ data collection approaches, like daily diaries and ESM have been widely adopted. These methods collect repeated self-reports in participants' natural environments, enabling more accurate tracking of symptoms and their contextual influences (Fillingim et al., 2016; Schneider et al., 2021). For example, longitudinal studies of urologic chronic pelvic pain syndrome have shown fluctuating symptom cycles that include periods of remission and flare-ups. To capture such temporal dynamics, Sutcliffe et al. (2019) argue for the use of in-situ methods. In research on conditions like GPP, daily diary studies have been used to track symptoms, function, and psychological well-being over time, revealing complex and individualized interactions between symptoms and daily life (Bergeron et al., 2021). While daily diaries offer more detailed information than retrospective assessments, ESM goes a step further by collecting real-time responses through repeated prompts during the day (Horstmann, 2020), enabling more fine-grained and context-sensitive data collection. Building on this body of work, the present thesis implements a smartwatch-based ESM system to support frequent, in-situ reporting in everyday contexts.

3.2 ESM for Health Data Collection

ESM, also referred to as Ecological Momentary Assessment (EMA), involves repeated sampling of individuals' experiences in real time and in their natural environments (Myin-Germeys et al., 2009; Shiffman et al., 2008). In health research, ESM is valued for its ability to reduce retrospective bias and capture how symptoms, behaviors, and contextual factors unfold over daily life (Shiffman et al., 2008; Trull & Ebner-Priemer, 2009; Verhagen et al., 2016). These properties make ESM particularly useful for studying biopsychosocial health conditions with high within-person variability, such as chronic pain and affective disorders (Castilla et al., 2012; Stone & Broderick, 2007).

Wearable ESM

Technological advancements have extended the potential of ESM to wearable platforms, giving rise to what is often referred to as wearable ESM (Hernandez et al., 2016). Wearable devices are worn or attached to the body, such as smartwatches. Smartwatches combine interactive displays with computational functionality, enabling dynamic interactions such as notifications and survey prompts. These features make them well-suited for self-reporting in ESM studies (Volsa et al., 2022). In contrast to other wearables, such as one-button trackers or sensor-only devices like pedometers and actigraphy monitors, smartwatches provide both input and output capabilities. Unlike smartphones, which may be stored away or require deliberate retrieval, smartwatches are worn on the wrist and can be accessed quickly, potentially reducing interaction effort and disruption (Volsa et al., 2022).

In addition to active self-reporting, smartwatches support passive sensing of physiological and contextual signals, such as movement and heart rate, which can complement subjective reports (Khanshan et al., 2023; Volsa et al., 2022). This combination is particularly relevant for longitudinal health monitoring, where contextual information can help interpret symptom patterns. At the same time, smartwatch-based

ESM must account for device constraints, including small screens and limited input options, which place strong demands on interface design.

Data Quality and Participant Engagement in ESM

High-quality data and sustained participant engagement are critical to the success of ESM studies. Previous work has shown that ESM protocol characteristics, such as sampling frequency, duration of the study, and length of each questionnaire, can significantly influence data quality and introduce bias (Carpenter et al., 2016; Reiter et al., 2025). If a protocol is too burdensome, it may lead to participant fatigue, careless responding, or dropout, thereby undermining the scientific value of the data.

As a result, many studies have sought to optimize ESM protocol parameters to balance data richness with participant burden. Examples include determining the optimal timing for prompting participants (Intille et al., 2003; Khanshan et al., 2023) or reducing the number and length of questions (Li et al., 2024). The interface through which ESM prompts are delivered, including its design, layout, interactivity, and responsiveness, also affects participant experience and engagement. A poorly designed UI can increase the perceived burden of participation, reduce motivation to respond thoughtfully, and even lead to partial or skipped entries (Hentati et al., 2021; Torous et al., 2018). In contrast, a user-friendly and context-aware application can integrate ESM more smoothly into daily routines, reduce cognitive load, and make the process feel intuitive and less intrusive. This is especially important when ESM is applied in sensitive health domains where both privacy and ease of use are paramount. In response, this thesis focuses on the design of a smartwatch-based ESM UI that prioritizes usability and low interaction effort, with the ultimate goal of supporting longitudinal data collection in everyday life.

3.3 Designing Wearable Interfaces for ESM

Designing wearable interfaces for experience sampling in health contexts presents unique challenges and opportunities that differ from mobile or desktop platforms. Smartwatches, in particular, offer constant physical proximity and timely access for in-situ data capture, but their small screens and constrained input modalities (e.g., tap, swipe, voice) demand carefully optimized UIs. Prior work in mobile health (mHealth) has shown that interface complexity, navigation effort, and interaction time directly impact user engagement and adherence in longitudinal studies (Deniz-Garcia et al., 2023). These concerns are magnified on wearables, where interactions are typically brief, interruptive, and glance-based (Markopoulos et al., 2023; Neshati et al., 2019; Patel et al., 2024).

To reduce cognitive and interaction burden, smartwatch-based ESM applications must emphasize immediacy, simplicity, and minimal disruption. This often involves favoring fixed-response formats over free-text input and leveraging microinteractions that require only a single tap or swipe. For example, Intille et al. (2016) introduced µEMA, a microinteraction-based EMA method that allowed users to respond to prompts with a single tap. Despite delivering prompts up to eight times more frequently than traditional smartphone-based EMA, µEMA achieved higher compliance and was perceived as less distracting. Similarly, Volsa et al. (2024) found that users of wrist-worn one-button devices, employing a simple Likert-style interface, reported more events and demonstrated greater adherence than those using smartphones, highlighting the efficacy of low-effort input designs in real-world data collection.

Further insights come from comparative evaluations of input techniques. Markopoulos et al. (2023) conducted a detailed study comparing touch-based input methods, such as tapping, bezel rotation, and swiping, for ESM on smartwatches. Their findings showed that tap-based and short-list selection methods enabled the fastest and most accurate interactions, especially during mobile use (e.g., walking), while complex gestures like swiping or bezel rotation led to higher error rates and longer task durations.

Overall, this body of work highlights a set of recurring principles: ESM applications on smartwatches benefit from simplified interaction flows, minimal cognitive load, and short-duration input methods tailored for glanceable use. These design considerations are especially critical when targeting sensitive health conditions, where the ability to quickly self-report during moments of discomfort, stress, or disruption can determine the quality and completeness of the data collected.

Chapter 4

DESIGN AND IMPLEMENTATION

This chapter describes the design and development process for the smartwatch-based experience sampling UI. Building on the project goals and requirements introduced earlier, it outlines how requirements were translated into interface concepts, how design ideas were explored and evaluated, and how these were developed into mockups and refined before implementation. The chapter then presents the implementation of the finalized prototype on Samsung smartwatches using Experiencer and its integration with GameBus for ESM protocol configuration and data management.

4.1 Design Process

The work followed four phases:

1. **Task clarification:** This phase involved collecting information about the requirements the UI must meet, under ESM protocol and deployment constraints.
2. **Conceptual design:** Exploration of alternative interface designs through low-fidelity prototypes and mockups, supported by a decision matrix. Each concept was evaluated against predefined criteria, and the design that best met these requirements was selected for implementation.
3. **Implementation:** Development of the selected design as a working prototype on Wear OS-based Samsung smartwatches, integrating Experiencer for survey delivery and GameBus for data management.
4. **Evaluation:** Assessment of usability and user experience with healthy participants.

This structure helped ensure that design decisions were grounded in both methodological requirements from ESM research and practical constraints of smartwatch interaction, while allowing iterative refinement based on stakeholder feedback and emerging insights during the project.

4.2 Task Clarification

The task clarification phase focused on identifying the requirements that the smartwatch-based ESM UI and its surrounding system and configuration needed to meet, as well as the constraints that would shape the design space. While the primary focus of this project was the smartwatch interface and interaction flow, the UI had to operate within a broader ESM setup that included questionnaire logic, study configuration, and data integration. These contextual factors informed the design and implementation choices described in this chapter.

Requirements

The requirements for the smartwatch-based ESM UI were defined together with the case owner. They were based on the assessment protocol developed by the researchers from Maastricht University (e.g., item content, length, and branching logic) and on UI/UX design considerations specific to smartwatch interfaces. In line with insights from prior work on wearable ESM (Intille et al., 2016; Markopoulos et al., 2023; Volsa et al., 2024), the requirements emphasized low-effort and glanceable interactions on the watch. Although the primary design focus was the UI, system-level requirements were included to enable the full ESM protocol deployment. Together, these inputs led to a set of requirements grouped into three categories:

- **Functional requirements:** The system shall:
  - F1 Deliver health assessment questionnaires on a circular smartwatch display.
  - F2 Support branching logic to present follow-up items only when relevant.
  - F3 Integrate with GameBus for secure data storage and management.
  - F4 Provide multiple questionnaire versions, as specified by Maastricht University: (i) Dutch version for patients, (ii) English version for patients, (iii) Dutch version for partners, (iv) English version for partners.
  - F5 Collect passive sensor data during the study (e.g., heart rate and step count) alongside self-reports, when supported by the device and platform.
- **Usability requirements:** The system shall:

U1 Support fast, low-effort microinteractions by presenting one questionnaire item per screen and enabling direct selection through simple touch input.

U2 Avoid pre-selected defaults and keep all response options visible without scrolling where possible. If scrolling is required, remaining options shall be clearly discoverable.

U3 Provide confirmation of selections to reduce accidental input and support confident responding.

U4 Ensure readability and accessibility on a circular smartwatch display by using clear typography, high contrast, and touch target sizes appropriate for small screens.

- **Technical requirements:** The system shall:

T1 Ensure compatibility with Samsung Wear OS devices.

T2 Build upon the open-source Experiencer application and GameBus Campaigns[1] for study configuration and survey delivery, while providing a custom smartwatch UI tailored to the requirements of this project.

T3 Guarantee GDPR-compliant data handling through the GameBus infrastructure.

Constraints

Key constraints included:

- **Platform and device limitations:** The capabilities of the smartwatch hardware, Wear OS, Experiencer, and the GameBus platform constrained certain design choices. For example, if the smartwatch does not support specific sensors, corresponding data cannot be collected. Similarly, if GameBus does not provide functionality for certain scheduling or branching logic, these features cannot be implemented in the prototype.
- **Device characteristics:** Small, circular smartwatch displays impose strict limits on screen real estate, text readability, and touch target sizes, which must be accommodated in the interface design.

---

[1] https://campaigns.gamebus.eu/

- **Privacy and regulatory constraints:** All data collection must comply with the General Data Protection Regulation (GDPR), requiring secure storage, controlled access, and informed consent procedures.
- **Project scope constraints:** The project focused exclusively on software development. Hardware selection was fixed (Samsung Galaxy Watch7), and usability testing was conducted using non-sensitive proxy questions with a healthy participant sample rather than the full assessment protocol from the Maastricht University use case.

The relative importance of these constraints was considered alongside requirements to define the design space for the subsequent conceptual design and implementation. Moreover, since this project focuses on UI design, considerations such as material costs or manufacturing constraints were not included.

4.3 Conceptual Design

The conceptual design step focused on generating and refining UI ideas for the smartwatch-based ESM interface, based on the aforementioned requirements and constraints.

Interface Ideation

To explore potential interaction approaches, several UI designs were created in Figma. Three representative example questions were used as placeholders:

- “How is your mood?”
- “How stressed are you?”
- “How is your pain?”

Each question was paired with five response options on a Likert-type scale, providing a realistic structure for prototyping layouts and interaction styles. Rather than converging on a single solution, this stage prioritized diversity. We considered the distinction between visual analogue (continuous) and numerical (discrete) response formats, which is especially relevant for pain assessment (Lund et al., 2005).

Although the subsequent evaluation was scoped to healthy participants, accessibility considerations informed several design decisions, in line with usability requirement

U4. High visual contrast was applied across all concepts, using light elements on a dark background (Google, 2025a). This supports legibility in varying lighting conditions and aligns with Wear OS design guidelines for glanceable interaction (Google, 2025b). In addition, circular layouts were explored to maximize available touch target sizes by distributing interaction elements along the perimeter of the display, mitigating input issues on small touchscreens.

To limit cognitive load, the interface presents one question per screen. Multi-item layouts and scrolling were avoided where possible. This structure aims to support rapid comprehension and minimize reading effort during microinteractions.

Figure 4.1 shows a selection of these early concepts.

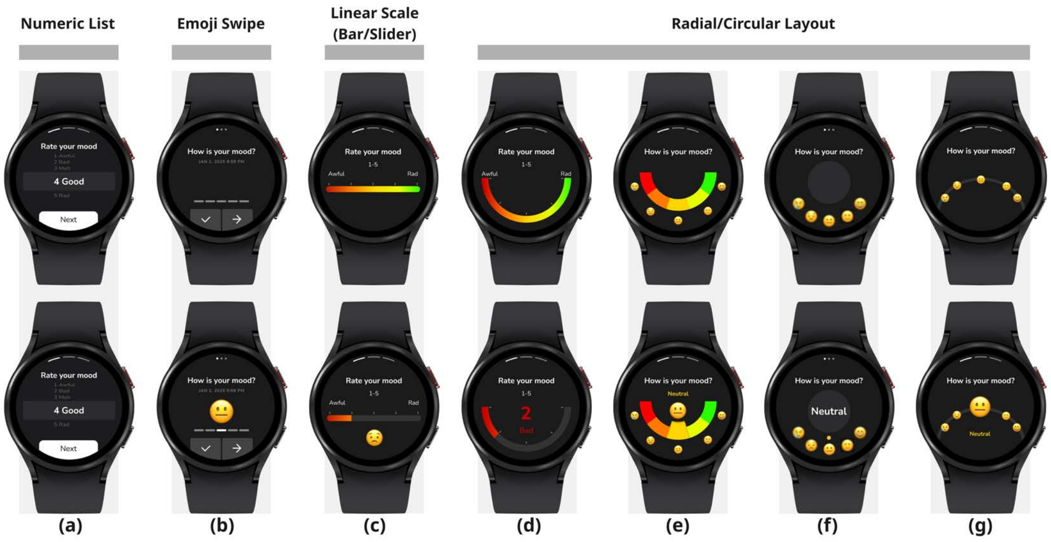


Figure 4.1: Example smartwatch interface concepts for the question “How is your mood?” designed in Figma. The designs illustrate different representation styles, including (a) numeric list selection, (b) single emoji swipe, (c) linear scale (bar/slider),(d) numeric radial gauges, (e) hybrid radial gauges with emojis, and (f–g) emoji-based circular layouts. All concepts include a progress indicator (dots or bar) at the top of the screen. Each top–bottom pair represents a variation of the same design concept.

The numeric list selection design (a) presents a vertical list of five labels ranging from negative to positive (e.g., Awful, Bad, Meh, Good, Rad). Users select an answer by tapping the corresponding number and label, followed by pressing a “Next” button to confirm their choice. This design provides a straightforward, text-based interaction familiar from traditional survey lists.

In the single emoji swipe confirmation design (b), the response is represented by a single emoji displayed at the center of the screen. Users can swipe left or right to navigate between emojis and confirm their choice using a checkmark button. A right arrow allows them to move to the next question.

The linear scale (bar/slider) design (c) uses a horizontal slider spanning from 1 to 5, with a continuous color gradient running from red (negative) to green (positive). Textual anchors are shown at each end of the scale (e.g., “Awful” to “Rad”). Users indicate their answers by dragging along the bar, and when a value is selected, a corresponding emoji appears above the slider to provide additional affective feedback. Double-tap advances to the next question.

Moving to circular formats, the numeric radial gauges design (d) presents a semi-circular arc segmented into five levels, color-coded from red to green. Users select a position on the arc, and the selected value is shown numerically (e.g., “2 Bad”) at the center of the display. Once a value is highlighted, users confirm their response by tapping the central area of the screen. This approach leverages the circular watch face to create a compact yet information-rich layout.

Building on this, the hybrid radial gauges with emojis design (e) integrates emojis into the same radial layout. Each segment of the arc is associated with an emoji, and selecting a segment displays both the numeric rating and the chosen emoji with its textual label (e.g., “Neutral”). This design combines the precision of numeric ratings with the expressiveness of emoji-based feedback. Similar to (d), users confirm their choice by tapping the central selected emoji.

Finally, the emoji-based circular layouts (f and g) remove numeric elements entirely and place only emojis along a circular arc around the screen. In (f), the emojis are arranged in a tight arc, and when one is selected, it becomes visually sharp while the other emojis are blurred, and its corresponding textual label appears at the center of the display. In (g), the emojis are positioned along a wider arc, and the selected emoji appears enlarged and centered with its label underneath (e.g., “Neutral”). In both (f) and (g), users confirm their response by tapping the central area of the screen after making a selection.

To identify the most suitable design, the decision matrix primarily addressed usability-related criteria (U1–U3) while also considering functional requirement F1 (questionnaire delivery on a circular smartwatch). Most other functional (F2, F3) and technical (T1–T3) requirements were not yet applicable at this stage because the concepts were low-fidelity and did not include system-level behavior (e.g., back-end integration, branching logic, data handling). The decision matrix in Table 4.1 was therefore used as a structured comparison tool to support early convergence rather than as a statistically valid evaluation. Scores were assigned based on design inspection against the key requirements, complemented by feedback from review moments with the case owner.

Table 4.1: Decision matrix comparing early interface concepts against key requirements. Ratings: 1 = very poor, 5 = excellent. Maximum possible total score = 40. (List = numeric list selection, shown as subfigure a in Figure 4.1; Emoji Swipe = single emoji swipe confirmation, subfigure b; Bar/Slider = linear scale with color gradient and emoji feedback, subfigure c; Radial Numeric = numeric radial gauge, subfigure d; Radial Emoji = hybrid radial gauge with emojis, subfigure e; Emoji Circle 1 & 2 = emoji-based circular layouts, subfigures f and g).

Criterion	**List**	**Emoji Swipe**	**Bar/Slider**	**Radial Numeric**	**Radial Emoji**	**Emoji Circle 1 & 2**
Readability	5	3	4	4	3	4
Simplicity	5	4	4	3	2	2
Circular Fit	2	3	1	5	5	5
Questionnaire Fit	4	3	3	5	4	4
Feasibility	4	5	3	4	3	3
Clarity	5	2	3	2	3	3
Avoid Default Bias	1	1	5	5	5	5
Interactivity	2	2	3	4	5	5
Total	28	23	26	32	30	31

The evaluation indicated that the Radial Numeric and Emoji Circle 1 & 2 designs achieved the highest overall scores. While other concepts (e.g., List and Bar/Slider) performed well in terms of readability and simplicity, they were less suited to the circular smartwatch interface and interactive selection. Although the Emoji Circle designs scored highly on usability criteria such as interactivity and avoidance of default bias, they were ultimately deemed less suitable for clinical deployment. Their high visual expressiveness is effective for single-item feedback, but when applied to longer questionnaires, as required in the Maastricht University use case, it could increase cognitive load and cause confusion. Based on these considerations, the Radial Numeric layout was selected for further refinement.

Design Trade-offs

The selection process highlighted that designing for a circular smartwatch requires balancing competing requirements. Several critical trade-offs were identified during the comparison of concepts; resolving these tensions was necessary to refine the chosen Radial Numeric interaction. These trade-offs and resulting design decisions are summarized in Table 4.2.

Table 4.2: Key design trade-offs identified during conceptual design and their implications for interface refinement.

Trade-off	**Design Tension**	**Implication for Design Refinement**
Circular fit vs. list interaction	Text-based list layouts are familiar and readable but do not map well to circular smartwatch displays.	A radial layout was preferred to leverage the circular form factor and support reliable touch interaction.
Visual expressiveness vs. clarity	Emoji-based designs provide engaging feedback but may increase cognitive load and ambiguity when repeated across longer questionnaires.	Emoji-only layouts were deprioritized; numeric scales were retained to support consistent interpretation.
Avoiding default bias vs. response speed	Pre-selected defaults can reduce interaction time but risk biasing self-reports.	No response option is pre-selected; users explicitly select and confirm before advancing.
Forward-only flow vs. back navigation	A persistent back button improves error recovery but competes for limited screen space and reduces room for response options and readable text.	Forward progression was prioritized; backward navigation relies on standard system gestures rather than on-screen controls.
Questionnaire fidelity vs. readability	Validated clinical items are often lengthy, whereas smartwatch displays require short, glanceable text to remain readable.	Item wording should be adapted in consultation with researchers to preserve meaning while improving readability on the watch.

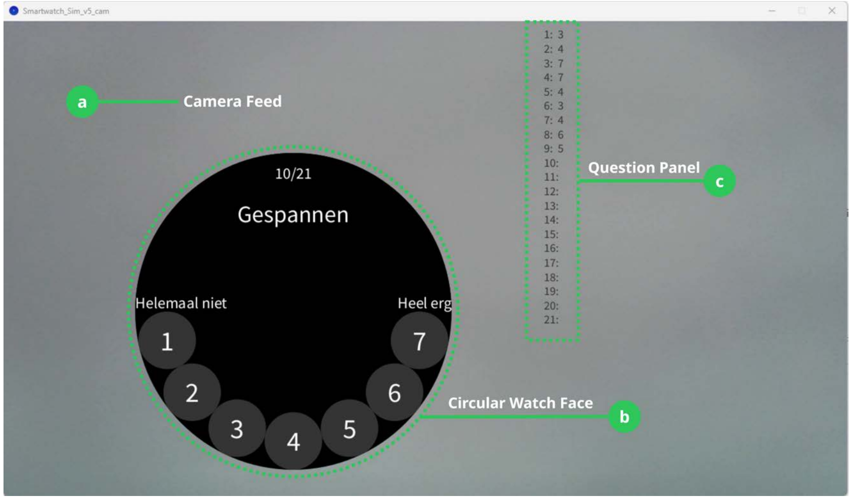
Low-Fidelity Prototyping

Following the selection of a preferred interaction idea and the design trade-off analysis, a low-fidelity prototype was developed using Processing. This stage focused on translating the chosen design into an interactive representation that could be used to examine layout, interaction flow, and readability on a circular smartwatch display.

At this stage, the complete Maastricht University GPP questionnaire was available, targeting two groups: patients and their partners. Each group had two language versions (Dutch and English), resulting in four versions in total. To resolve the *Questionnaire fidelity vs. readability* trade-off (Table 4.2), the research team first discussed shortening longer items where necessary without altering their meaning. These adapted questions were then integrated into the Processing mockup to simulate the questionnaire flow on the smartwatch.

To improve clarity and interpretability, labels were added to the first and last response options in the circular layout, clearly indicating the scale extremes (e.g., "Not at all" and "Very much"). All response options remained in a circular arrangement, with the center of the circle functioning as the confirmation area for the selected choice. This adjustment helped participants quickly understand the scale and reduced the risk of misinterpretation.

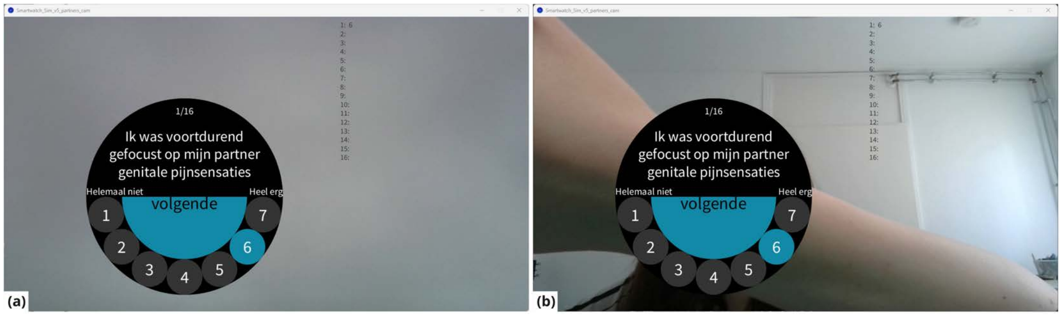
The prototype was integrated with a live camera feed using the `Capture` library. This function simulates on-wrist use by displaying the interface over a live camera feed. Likert-scale response options were displayed in a circular layout, reflecting the intended smartwatch interface. The prototype captured user selections in real time. Example screenshots of the interactive mockups are provided in Figure 4.2.



```
1: 3
2: 4
3: 7
4: 7
5: 4
6: 3
7: 4
8: 6
9: 5
10:
11:
12:
13:
14:
15:
16:
17:
18:
19:
20:
21:
```

Figure 4.2: Low-fidelity Processing-based smartwatch prototype for the GPP study (example: Dutch version for GPP patients). The interface is divided into three main areas: (a) background showing a live camera feed simulating the participant wearing the watch, (b) left area displaying the circular watch face with interactive response options, and (c) right area listing the current question number, total questions, and participant-selected responses.

Consistent with the emphasis on visual simplicity and forward progression established in earlier design stages, backward navigation was not represented as a persis-



Ik was voortdurend gefocust op mijn partner genitale pijnsensaties

Figure 4.3: Additional examples. The interface is shown with two different questionnaires: (a) the Dutch version for GPP partners, and (b) with the participant's arm visible in the live camera feed background.

tent on-screen control. The prototype maintained a forward-only visual flow, with one question presented per screen and display space reserved for response options and text readability. As noted in the design trade-offs, introducing an explicit back button would have required either reducing touch target sizes or increasing visual density. Instead, backward navigation was decided to rely on standard smartwatch interactions, namely the hardware side button or swipe gesture.

During this stage, a review meeting was conducted with the Maastricht University research team to review the low-fidelity mockups. The team confirmed that the interface layout, interaction flow, and text presentation were appropriate for the target user group and aligned with the study goals. With this consensus established and no major structural changes required, the design was confirmed for implementation on the target smartwatch device.

4.4 Implementation

This section describes how the functional, usability, and technical requirements identified earlier were realized in the implemented prototype. The implementation focused on translating the specified requirements into a working smartwatch-based application with the newly designed UI. In total, the application was deployed on eight smartwatches for the research project at Maastricht University, configured for the four questionnaire variants (Dutch/English × GPP patient/partner). Each smartwatch was linked to a dedicated participant account. For the usability evaluation in Chapter 5 reported in this thesis, a separate single-session setup with English proxy items was used.

System Architecture

Figure 4.4 provides an overview of the system architecture: (1) the wearable client running on the Samsung Galaxy Watch7, (2) the backend infrastructure provided by the GameBus platform, and (3) the researcher-facing tools used for study configuration and data extraction.

On the wearable side, the Experiencer application handles user interaction, ESM prompting, questionnaire rendering, and the collection of self-report and passive sensor data. The smartwatch communicates with the GameBus backend via the GameBus API to retrieve study configurations and to upload collected data. On the researcher side, GameBus Campaigns is used to define and deploy the ESM protocol, while a separate data extraction pipeline supports the retrieval of raw study data for analysis. The next section further details the roles of Experiencer and GameBus.

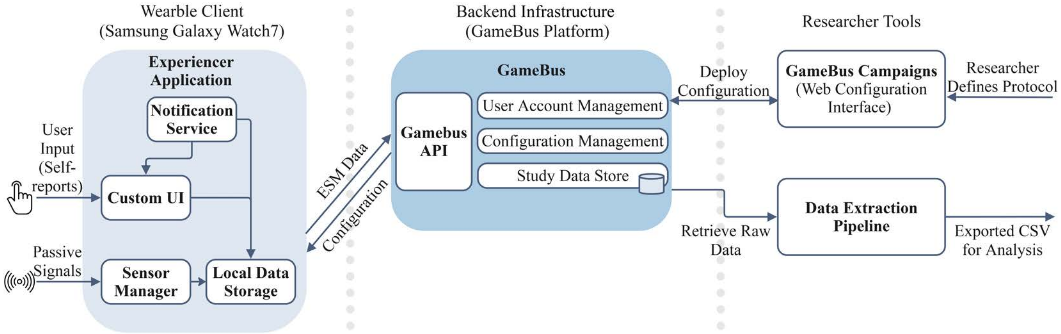


Figure 4.4: System architecture of the smartwatch-based ESM setup. The Experiencer application runs on a Samsung Galaxy Watch7 as a wearable client, providing the custom questionnaire UI, prompt handling, and sensor data collection. Study configuration is retrieved from the GameBus backend via the GameBus API, while self-report and sensor data are uploaded for secure storage. Researchers can define and deploy study protocols using GameBus Campaigns, and collected data are retrieved through a data extraction pipeline and exported for analysis.

Experiencer and GameBus

The smartwatch interface was implemented as a custom UI of *Experiencer*, an open-source and GDPR-compliant ESM application (Khanshan et al., 2023). Experiencer functions as a *client-side* application running on the smartwatch: it delivers ESM prompts to participants, renders questionnaire screens, collects self-reports and sensor data on the device, and communicates with a backend infrastructure (GameBus) for study configuration and data storage. Earlier versions of Experiencer were developed on the Tizen OS Web API and deployed on Samsung Galaxy Watch Active 2 devices (Khanshan et al., 2021). In this project, we used the latest release of Experiencer (version v3.1.1-prod), which supports Wear OS-based Samsung smartwatches and was deployed on the Samsung Galaxy Watch7, satisfying technical requirement T1.

Experiencer relies on *GameBus* as its backend platform. GameBus provides the server-side infrastructure for secure, GDPR-compliant study management and data handling (Shahrestani et al., 2017). In practice, Experiencer communicates with GameBus through the GameBus API for authentication, study configuration retrieval, and data upload. This integration fulfills functional requirement F3 by ensuring that data collected on the smartwatch are stored and managed through the GameBus infrastructure.

To run Experiencer using GameBus, two main components are available:

- **GameBus Studio**: the participant-facing interface used to deliver health promotion campaigns and allow end users to review their submitted data. This component was not used by participants in the present project.
- **GameBus Campaigns**: the researcher-facing interface used by organizers to configure ESM, define survey content, timing, and duration, specify sensor settings, and manage participants.

In this project, all study configuration was performed in GameBus Campaigns, including specifying the inter-notification interval, configuring sensor settings for passive data collection, and defining contextual rules for context-sensitive sampling. For this project, these settings were configured in accordance with the GPP ESM protocol design. The configuration also defines the questionnaire to be delivered to participants. In this project, we configured four questionnaire versions to meet requirement F4 (Dutch/English for patients and partners). Additionally, in its latest release, GameBus added support for branching logic in questionnaires (e.g., *if–then* conditions). This functionality was essential for implementing the pain-related questionnaire developed by Maastricht University and thus fulfilling functional requirement F2.

The contribution of this project lies in redesigning and implementing a smartwatch ESM UI on top of the Experiencer–GameBus infrastructure, adapting the system to circular Wear OS displays and wearable ESM usability requirements.

Front-End Interface

After finalizing the interface design, the next step was to implement the questionnaire interface on Wear OS-based Samsung smartwatches. The implementation was carried out in *Android Studio* using Kotlin, following Wear OS development guidelines.

Questionnaire items and the sampling schedule provided by Maastricht University were configured in GameBus Campaigns using JSON files, with branching logic applied to allow conditional navigation between questions, displaying follow-up items only when relevant.

The implementation focused on translating the selected interaction design into a robust and usable smartwatch interface. Response options were arranged in a semi-circular layout, and a select–then–confirm interaction was implemented using a central semi-circular 'NEXT/VOLGENDE' button, allowing users to verify their selection before proceeding.

Additional refinements included adjustments to response option size, touch target dimensions, button placement, and visual contrast to support reliable touch input and improve readability on the circular smartwatch display.

Figure 4.5: Implemented smartwatch-based ESM interface on a Samsung Galaxy Watch7 (Wear OS), illustrating the questionnaire interaction flow: (a) start screen for initiating a self-report, (b) presentation of a single questionnaire item with circular Likert-scale response options, (c) selected response with central 'NEXT' confirmation button, (d) completion screen following submission, and (e) push-up notification prompting the user to initiate a self-report.

Data Extraction

To facilitate the analysis of participant responses, the *GameBus Data Analyzer*[2] was employed. For this study, the author adapted the original code to allow all entries to be collected and stored. In addition, a Python script was developed to automate the

---

[2] https://github.com/oaglazunova/gamebus-data-analyzer (accessed August 18, 2025)

extraction, transformation, and export of the data into commonly used formats, i.e., CSV.

This step confirmed that the implemented setup supported reliable data retrieval as required for evaluation, building on the secure data handling provided by the GameBus infrastructure (F3, T3).

User Manual and Training

A user manual was prepared for the Maastricht University research team to support the use of the smartwatch-based ESM application. The manual provides instructions for configuring the Experiencer application, initializing devices and linking them to participant accounts via GameBus Campaigns, managing daily device use, and extracting study data. A training session was provided to ensure that the research team could operate the devices.

For complete instructions, see Appendix B.

4.5 Technical Evaluation

A technical check was conducted to confirm that the ESM setup functioned as intended. The objective was to ensure that responses collected on the smartwatch could be transmitted through Experiencer, stored on the GameBus backend, and subsequently extracted for analysis.

A two-day technical evaluation was conducted by the author, during which scheduled ESM prompts were delivered to the Samsung Galaxy Watch7 running the Experiencer application with the designed interface installed. Test responses were submitted, and the data flow throughout the system was tracked. No transmission failures or data loss were observed during the evaluation period. All responses, together with accelerometer data and heart rate measurements, were successfully retrieved from GameBus in raw format and exported to CSV using a custom Python script. An excerpt of the exported response table (with item text removed and identifiers masked) is provided in Appendix C. For privacy and ethical reasons, the full question items and any sensor logs are not included in this thesis.

Chapter 5

EVALUATION

This chapter describes the evaluation conducted to assess the usability and user experience of the smartwatch-based ESM application's UI. The study protocol was reviewed and approved by the Ethics Review Board of Eindhoven University of Technology.

5.1 Proxy Questionnaire Design and Rationale

The focus of this evaluation was on usability and user experience, that is, how effectively participants could operate the interface, navigate between screens, and respond to prompts on the smartwatch. The questionnaire developed for GPP research at Maastricht University was not used for evaluation because these items were considered highly sensitive (e.g., intensity of genital pain, sexual arousal, relationship conflict). Assessing these clinical items falls outside the scope of the present work and will be carried out separately by Maastricht University as part of their ongoing clinical research.

To enable a realistic, ethically appropriate evaluation, the sensitive clinical items were replaced with non-sensitive, daily experience proxy questions presented in English. These proxy items were primarily adapted from the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988), a validated and widely used instrument for measuring momentary affective states. Using PANAS-based items ensured that the proxy questions were scientifically grounded, concise, and suitable for a general participant population. The primary objective was to test the smartwatch application's UI while preserving the interaction fidelity of the original design. Accordingly, the proxy questionnaire replicated the key interaction structures of the clinical version:

- a 7-point Likert scale for rating momentary states,
- a 3-choice categorical response for contextual items, and
- conditional branching logic between certain questions.

Figure 5.1: Examples of proxy questionnaire items. (a–b) PANAS-based items using a 7-point Likert scale; (c–e) adapted items in which partner-related wording from the original questionnaire was replaced with friend-related wording.

Although the content was substituted, the interaction flow and structure were kept consistent across all 24 items (including response format and step sequence) so that the usability findings reflect the interface design. The complete list of the proxy questions, their corresponding UI scale, and rationale is provided in Appendix D. The original questionnaire is omitted due to the confidentiality of the Maastricht University research at this stage.

5.2 Participants

12 participants (age range 25–36 years, M = 29.1, SD = 3.0; 6 female, 6 male) were recruited for the study and are referred to as P1–P12. Of them, eight participants (P3, P4, P5, P6, P7, P8, P11, and P12) had prior experience participating in a study involving self-reporting experiences or health data using a digital device (e.g., smartphone, smartwatch, or tablet). Each participant was provided with a smartwatch (Samsung Galaxy Watch7) that had Experiencer pre-installed. The devices were set up before each session. Participation was voluntary, and informed consent was obtained from all participants prior to data collection.

5.3 Procedure

The user study was conducted in a single session for each participant, with a total duration of approximately 30 minutes. The study followed a structured sequence:

1. **Introduction and consent (5 minutes):** Participants were welcomed, and the purpose of the study was explained. Written informed consent was obtained, and participants then completed a short demographics questionnaire (age, gender, smartwatch experience).

2. **Main task: ESM prompt completion (10 minutes):** Participants then received the smartwatch and began the task immediately. No formal training or

demonstration was provided; participants were told they could ask questions at any time if they needed assistance. They then completed one full proxy questionnaire on the smartwatch. After confirming the instructions, the researcher left the room to avoid observer influence. Participants completed the task independently, with assistance provided only upon explicit request.

3. **Post-study questionnaire (5 minutes):** Participants completed the System Usability Scale (SUS) to provide standardized subjective assessments of usability and user experience.

4. **Semi-structured interview (10 minutes):** To gain a deeper understanding of participants' experiences using the application, semi-structured interviews were conducted to gather qualitative feedback on participants' experience with the interface. The interview focused on participants' subjective experience with the interface, including first impressions, ease of navigation, clarity of questions, interaction flow, readability, perceived burden, and suggestions for improving the interface. The full interview guide is provided in Appendix F.

5.4 Measurements

Usability was assessed using the System Usability Scale (SUS) (Brooke, 1996), a ten-item questionnaire (items Q1–Q10) rated on a five-point Likert scale from strongly disagree to strongly agree (Lewis, 2018). SUS provides a quantitative measure of overall system usability. Responses were converted to a 0–100 scale for each participant. Higher SUS scores indicate better perceived usability. According to standard benchmarks (Bangor et al., 2008; Lewis & Sauro, 2018): a score of 68 corresponds to average usability relative to a broad sample of products. Scores above 68 generally reflect above-average usability; scores substantially higher (e.g., in the upper 70s or 80s) correspond to 'good' or 'excellent' usability, while scores below average suggest usability problems or room for improvement. The complete list of SUS items used in this study is provided in Appendix E.

5.5 Data Analysis

All data were de-identified and pseudonymized prior to analysis to ensure participant confidentiality. SUS responses were summarized using descriptive statistics (e.g., mean, standard deviation, and score distribution). Qualitative interview notes were transcribed and analyzed thematically following an iterative coding approach.

5.6 Results

This section presents the results of the study. We first report the quantitative results, followed by the qualitative findings.

System Usability Scale (SUS)

The application achieved a mean SUS score of 83.1 ($SD = 11.2$). According to the adjective SUS rating scale, this score corresponds to an *excellent* usability rating and is well above the established benchmark of 68 (Bangor et al., 2008; Lewis & Sauro, 2018). 83.3% ($N = 10$) of participants rated the system with a SUS score of 80 or higher, while 16.7% ($N = 2$) scored below the benchmark of 68. Figure 5.2 presents individual SUS scores per participant, and Figure 5.3 illustrates the distribution of responses for each SUS item (Q1–Q10).

Overall, participants evaluated the system as highly user-friendly, particularly in terms of learnability and ease of use. All 12 participants agreed or strongly agreed that the system was easy to use (Q3) and that they could learn to use it very quickly (Q7). All participants also disagreed that they would require technical support to use the system (Q4) and disagreed that they needed to learn many things before getting started (Q10). This consistency indicates that the onboarding process and interaction flow were intuitive and required minimal prior instruction. Responses related to complexity and function integration showed some variation. Regarding perceived complexity (Q2), three participants selected the neutral option, while the remaining nine disagreed that the system was unnecessarily complex. Most of the participants felt that the functions of the system were well integrated (Q5), although five provided neutral responses.

Item Q1 (“I think that I would like to use this system frequently”) showed the widest spread of responses: six participants selected the neutral option, three disagreed, and three agreed.

Qualitative Results

Interface is clear, intuitive, and easy to operate. Participants consistently described the interface as clear, easy to understand, and intuitive. Descriptions included “self-explanatory” (P1, P7), “clear” (P3, P6), “easy to understand” (P3, P4), “easy to use” (P5, P10), “simple” (P1, P2, P10), “straightforward” (P7, P9, P12), and “intuitive” (P1, P2, P9). The linear workflow and fast interaction design contributed to this perception. For example, P1 noted that “it’s very self-explanatory... you press the [answer], and you press next,” and P10 emphasized the speed to finish the

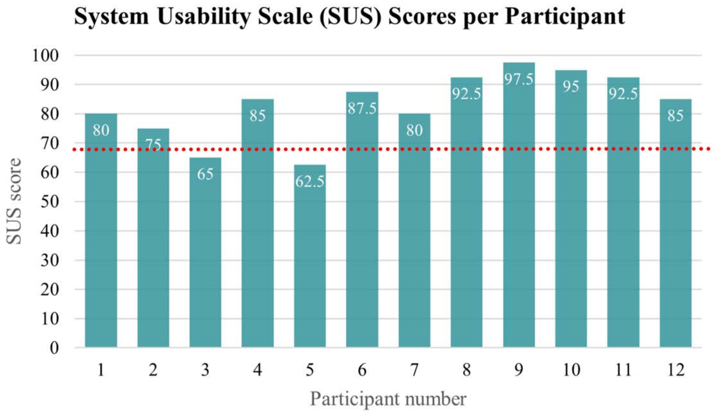


Figure 5.2: System Usability Scale score. The dotted line indicates the benchmark score of 68.

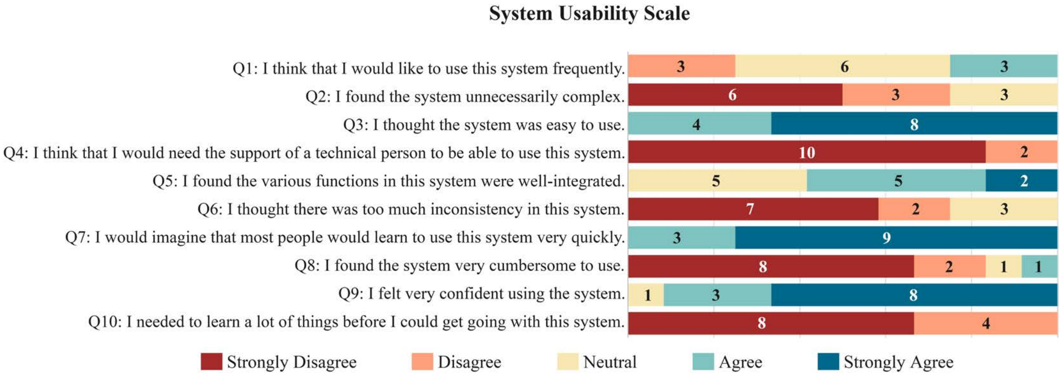


Figure 5.3: Participant responses to each SUS item (Q1–Q10). Each segment of the stacked bars includes the raw number of participants ($N = 12$) selecting that option.

questionnaire felt very “fast.” P12 described the first impression as “simplistic and straightforward” and “procedural,” noting that the one-question-at-a-time flow made it “easy to focus”. Overall, participants felt they did not need prior instruction or training to operate the interface, which aligns with the ratings given for SUS items Q3 (easy to use) and Q4 (need for technical support).

The select-then-confirm interaction, where users select an option and then press a distinct ‘NEXT’ button, was perceived as a necessary safety mechanism, helping to prevent input errors. P8 explained that the flashing ‘NEXT’ button “helped to indicate that I need to press it... maybe if I miss a click, I can still change my answer”. P11 appreciated being able to verify selections before confirming, describing it as “safe to explore.” Participants also observed that the layout facilitated rapid decision-making, allowing direct selection of options without confusion (P6).

Circular layout makes efficient use of limited screen space. Participants found that the interface handled the constraints of the smartwatch form factor effectively. P5 praised the circular layout, noting that it was "coherent with the ring-shaped screen" and "makes use of every inch of the space." P11 described the positioning of the Likert scale as "making a lot of sense" in terms of spatial mapping.

Despite the small touch targets, all participants reported confidence in selecting options. P2 explained, "I try to see if I am going to tap something wrong, but every time it recognized my choice immediately." P8 noted that their fingers occasionally occluded labels, but the confirmation step mitigated potential errors. P11 noted that while they personally "had no problem" selecting options due to having "small kind of fingertips," users with larger fingers might find it more challenging.

Readability constraints on small smartwatch displays. Five participants (out of 12) reported that long labels appeared cramped or split awkwardly across lines (see Figure 5.4), making them harder to parse. P6 mentioned the word being split across two lines, stating that it "took me a while to recognize what that means." P12 noted that for multi-line descriptions, "you have to assemble the word in your head a little bit," requiring "additional effort as a non-native English speaker", which interrupted the interaction flow. Additionally, P11 raised inclusivity concerns, highlighting that while text size was adequate for a healthy population, fragmented text and small labels could be difficult for older populations or those with visual impairments.

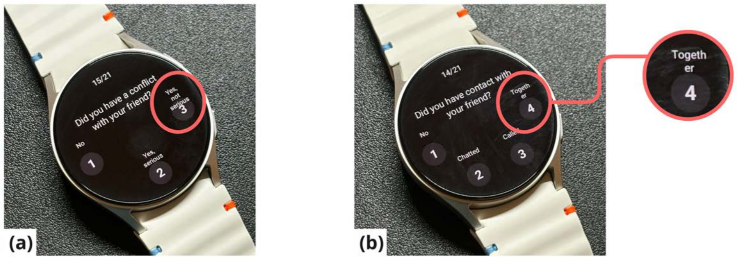


Figure 5.4: Examples of readability issues on the UI: (a) Multi-line labels and (b) split words increase interpretive effort during interaction (highlighted in red).

Some users were uncertain about how to navigate backward.

Three participants (P2, P9, P10) reported confusion about how to go back and revise a previous answer. Notably, these participants also reported no prior experience with digital self-report studies (e.g., on a smartphone, smartwatch, or tablet), whereas most other participants either did not encounter this issue or did not attempt to navigate backward. The interface follows a forward-only visual flow, with backward navigation accessible via a swipe gesture or the hardware button. P9 suggested adding a visual indicator or arrow to clarify this feature for first-time users. P12 mentioned that they “did not know” the back gesture existed and assumed that once the ’Next’ button appeared, they “cannot change [the] answer anymore.”

Participants anticipate effort when imagining repeated use. Although the single-session interaction was consistently described as “quick” and “manageable” (P1, P3, P4, P12), participants expressed concerns when asked to imagine completing the full questionnaires multiple times per day. P1 highlighted this distinction: selecting answers “is very quick,” but repeated completions throughout the day “feel a lot.” P3 anticipated that responding five or six times daily would be “too much,” whereas P9 suggested that “two or three times at most” would be acceptable, and the application should provide effective reminders. These reflections are hypothetical and represent perceived rather than experienced burden. This aligns with SUS Q1 responses, where most participants remained neutral regarding whether they would use the system frequently. Some participants noted a higher willingness to engage in repeated use if it served a clear clinical purpose (P4, P11) or supported self-tracking (P1).

Suggestions for visual design and feedback. Participants generally found the interface functional, but several noted that the visual design felt overly “basic” (P8) or “boring” (P1). P10 expressed a desire for something “more beautiful, more exciting to my eyes.” Participants also recommended richer feedback to enhance the interaction experience. P3 and P7 recommended adding haptic vibrations or sound to confirm interactions or remind users to complete surveys. P3 also suggested using color coding (e.g., red to green) to make the Likert scales more intuitive. P12 proposed gamification elements, such as a circular progress indicator, to indicate completion in a visually rewarding way.

5.7 Discussion

This evaluation assessed the usability and user experience of a smartwatch-based ESM interface, with a focus on whether participants could navigate the interaction flow, respond to prompts, and operate the questionnaire clearly and confidently on a small touch display. Overall, the results suggest that the interface enables efficient microinteractions, reflected in the high SUS score ($M = 83.1$) and supportive qualitative feedback. At the same time, several findings (e.g., neutral/negative responses on SUS Q1, readability issues, and uncertainty about backward navigation) indicate friction points that could reduce compliance in real-world ESM use.

Microinteraction Usability: Fast, Learnable, and Low-effort

Participants were able to complete the full prompt sequence independently and without training, and they described the interaction as clear, intuitive, and straightforward. This finding aligns with prior work on wearable ESM, which argues that smartwatch-based reporting is feasible when prompts are answered through brief, glanceable microinteractions with constrained response sets (Intille et al., 2016). In our study, the one-question-per-screen structure and linear progression supported a clear and predictable interaction flow, enabling participants to quickly understand how to proceed through the questionnaire. This is reflected in unanimous agreement on SUS items related to ease of use and learnability (Q3, Q7), and unanimous disagreement on items suggesting a need for technical support or extensive learning (Q4, Q10).

Importantly, the positive usability outcomes were achieved despite typical smartwatch constraints (small display, limited space, touch input). This suggests that for ESM interfaces, interaction simplicity and a consistent microinteraction pattern can compensate for some hardware limitations, supporting the core requirement of low interaction burden during in-situ responding.

Usability is not the Same as Sustained Engagement

Despite high usability ratings, half of the participants ($N = 6$) reported neutral willingness to use the system frequently (SUS Q1). Interview responses indicate that this was primarily about *perceived value and context* rather than interaction difficulty: participants were healthy, the evaluation was a single-time data-collection exercise, and completing repeated prompts lacked an immediate personal benefit.

This distinction mirrors prior ESM research showing that compliance and engagement depend not only on interface quality but also on protocol-level factors such as perceived purpose, personal relevance, prompting frequency, participant motivation, and external incentives (Scollon et al., 2003; Wrzus & Neubauer, 2023). Several participants in the present study similarly indicated that their willingness to engage more frequently would increase if the system supported a clear clinical goal or a meaningful self-monitoring purpose. These findings indicate the importance of evaluating wearable interface ESM systems within realistic use contexts, where both interaction burden and motivational factors affect sustained participation.

Form-factor-aware Layout and Error Prevention

Participants responded positively to the circular Likert layout, explicitly describing it as coherent with the watch's round display and easy to interpret at a glance. This aligns with established principles of natural mapping and spatial similarity: when interface structure mirrors the relevant physical or spatial layout, interactions can be faster to learn and support more efficient responding (Norman, 2013). Despite the limited screen real estate of smartwatches, participants reported that the circular layout facilitated rapid interpretation and confident selection of response options.

However, smartwatch platforms differ substantially in display geometry. While the Samsung devices used in this study feature a circular screen, other widely used devices (e.g., Apple Watch) use rectangular displays. A circular layout can still be shown on a rectangular screen, but the geometry affects interaction in practice: it changes usable space, option spacing, touch-target size, and the likelihood of finger occlusion near edges. These factors can influence speed and error rates. Therefore, wearable ESM interfaces should be validated on the target form factor and adapted when needed to maintain readability and reliable input.

Additionally, the select-then-confirm mechanism was repeatedly described as a safety feature that reduced fear of mis-taps and supported confident input. This is particularly relevant for smartwatch ESM, where accidental selections are more likely due to small touch targets and finger occlusion. The present findings therefore suggest that combining (1) a form-factor-aware arrangement of options with (2) an explicit confirmation step can maintain speed while improving perceived reliability of input on small screens.

Forward Navigation is Clear, Backward Navigation is Not

While the forward interaction was perceived as self-explanatory, three participants (all without prior experience with digital self-report studies) mentioned they were unsure how to go back or revise an answer after pressing 'NEXT'. This uncertainty was directly related to the design choice not to include an on-screen back button in order to preserve screen space and visual simplicity on the circular smartwatch display. Backward navigation instead relied on standard smartwatch interactions, such as swipe gestures or a hardware back button, which were not immediately discoverable to novice users.

Participants who raised this issue reported no prior experience with digital self-report studies, suggesting that implicit smartwatch operations are less discoverable for first-time users. Therefore, providing first-use guidance or a minimal visual cue on navigation for first-time users is important, as it could help clarify how to revisit previous questions without compromising the intended simplicity and efficiency of the interaction flow.

Readability and Inclusivity on Small Displays

A key limitation observed in the interaction was text rendering: five participants reported that longer labels appeared split or cramped, increasing interpretive effort. Although not all participants were affected, this finding highlights a known constraint of smartwatch interfaces: microinteractions are well suited to short responses and minimal text (Intille et al., 2016), and readability quickly degrades as textual density increases. In our study, fragmented or broken words were reported to disrupt interaction flow even within a short, single-session evaluation, and participants explicitly noted that non-native speakers and older or visually impaired users might be disproportionately affected. For wearable ESM systems, this suggests that questions and label length and typography are not cosmetic concerns but usability-relevant constraints. Possible mitigations include abbreviated on-device labels (with clearer explanations provided during onboarding or training) or alternative phrasing that preserves meaning while reducing character count. These adjustments are particularly important when the target population includes clinical users or older adults, where readability barriers can directly impact data quality through skipped items or incorrect responses.

Aesthetics and Feedback: Balancing Neutrality with Engagement

Several participants described the interface as visually “basic” or “boring” and suggested richer feedback or more expressive visual elements. While such feedback reflects users’ desire for a more engaging experience, prior research in survey methodology cautions that visual presentation and formatting can influence how questions are interpreted and answered, potentially introducing response bias or reducing data quality (Stern et al., 2007). In addition, simple and minimalist interface designs are known to support processing fluency (Choi & Lee, 2012), which is particularly important for performing brief self-report tasks on small screens with limited display space and interaction constraints. Therefore, ESM systems, particularly in clinical assessment contexts, should favor neutral, simple, and nondistracting interface designs to minimize unintended influences on self-reported data, while carefully considering how limited feedback may be introduced without compromising data quality.

At the same time, participants' comments indicate that a lack of visual richness should not be conflated with a lack of interaction feedback. In particular, *micro-level feedback*, such as haptic or auditory confirmation, was perceived as potentially beneficial for improving clarity and responsiveness during interaction. Although the Experiencer application supports vibration-based cues when prompt notifications are delivered, these feedback mechanisms were not experienced by participants in the present evaluation.

Prior Human-Computer Interaction (HCI) research has shown that haptic feedback is effective when users' visual attention is limited, enabling confirmation and guidance without requiring continuous visual focus (Wang et al., 2022; Whitmore et al., 2024). This is particularly relevant for smartwatch-based ESM, where interactions often occur in interruptive or attention-constrained contexts. Feedback delivered through non-visual channels can therefore enhance interaction quality while preserving a restrained visual presentation. A reasonable design implication is therefore to distinguish between *decorative styling*, which may be intentionally minimized in clinical ESM to preserve neutrality, and *functional feedback*, which can support learnability, accessibility, and prompt compliance. Subtle haptic confirmation after selection or minimal progress indicators may strengthen the interaction experience while remaining appropriate for clinical and research-oriented ESM applications.

Implications for Wearable ESM User Interface Design

Based on the usability outcomes and qualitative findings, several design implications can be derived for smartwatch-based ESM systems intended for repeated, in-situ use in health research contexts.

- **Prioritize interaction patterns that require no training.** Participants were able to complete the ESM task without prior instruction and consistently described the interface as self-explanatory. This indicates that smartwatch-based ESM interfaces should be designed such that correct use can be inferred naturally from the interface itself. For researchers deploying ESM in real-world contexts, this is critical: participants cannot be assumed to remember instructions when prompts occur intermittently throughout the day. Designers should therefore favor simple, linear interaction flows and avoid reliance on onboarding tutorials or written guidance.

- **Explicitly inform users that a response has been made.** Participants valued being able to confirm a selected response when moving on, which increased clarity during input on the small touchscreen. However, uncertainty about when a response was finalized raised concerns about accidental input. In high-frequency ESM, where responses are brief and interaction errors are likely, this uncertainty can lead to hesitation or uncorrected mistakes. Designers should therefore clearly indicate when a response is committed and provide a discoverable way (or brief first-use instruction) to revise answers.

  Participants described the select-then-NEXT interaction as supporting careful input, as it allowed them to review a selection before moving on. However, some participants were unsure whether responses could be revised after advancing. In ESM, this uncertainty can result in uncorrected mistakes or slower interaction, especially on small touchscreens where mis-taps are common. Wearable ESM interfaces should therefore make the point of response commitment and the possibility of correction explicit, rather than relying on implicit navigation conventions.

- **Treat text length as a primary design constraint, not a cosmetic issue.** Readability problems emerged when labels exceeded the spatial limits of the smartwatch display, leading to split words and increased interpretive effort. These issues were reported even in a short, one-time interaction. For researchers designing ESM questionnaires, this implies that question wording must be adapted to the constraints of the device, rather than transferred directly from smartphone or paper-based instruments. Designers should assume that longer or complex phrasing will negatively affect both usability and response quality on smartwatches.

- **Separate interface usability from engagement and compliance.** Despite high usability scores, participants were largely neutral about frequent use. Interview data indicate that this was driven by perceived relevance and context, not interaction difficulty. For researchers, this highlights that usability evaluations alone are insufficient to predict long-term ESM compliance. Engagement depends on study design factors such as perceived value, prompting frequency, and personal relevance. Usability testing should therefore be complemented with longitudinal or in-the-wild evaluations when sustained participation is a goal.

- **Design for visual simplicity, but support interaction feedback.** Participants described the interface as visually minimal and, in some cases, unengaging, yet did not report confusion during interaction. This suggests that a neutral visual design is appropriate for clinical or research-oriented ESM, where reducing bias and distraction is important. However, participants explicitly requested clearer feedback (e.g., haptic confirmation), indicating that visual simplicity should not come at the expense of interaction clarity. Designers should distinguish between decorative elements, which may be unnecessary, and functional feedback, which supports confidence and accessibility.

Limitations and Future Work

Several limitations should be acknowledged when interpreting the findings of this study. A primary limitation is that the evaluation examined a single interaction in a controlled laboratory setting. Participants completed a one-time self-report using non-sensitive proxy questions rather than the original clinical items. As such, the findings reflect first-time use and immediate interaction quality, but do not capture how users might adapt to the interface, experience fatigue, or engage with the system over repeated use in everyday contexts.

The study was also conducted with healthy adult participants. While this allowed us to evaluate general usability, this limits how far the findings can be generalized to target populations who may use ESM in clinical or long-term health contexts. During the interviews, several participants pointed out that the small watch face and text presentation could pose challenges for users with reduced visual acuity. These accessibility-related concerns were beyond the scope of the present study, but should be considered in future iterations of the interface.

In addition, some practical aspects of smartwatch-based ESM could not be assessed within the scope of this study. In particular, smartwatch battery life, an important factor for intensive, multi-day ESM protocols, was not evaluated, as participants interacted with the system only briefly. More generally, real-world use of wearable devices is influenced by everyday routines, charging behavior, and occasional non-wear, which can affect data completeness and reliability (Tang et al., 2018). These factors are difficult to examine in laboratory studies and were outside the scope of the present work.

Importantly, this study highlights that high interface usability does not necessarily translate into sustained engagement. Despite high usability ratings, participants expressed neutral willingness to use the system frequently, primarily due to the lack of personal relevance and the artificial, one-time study context. This suggests that disengagement in long-term ESM use cannot be attributed to interface design alone, but may also result from protocol-level factors such as questionnaire content, prompting frequency, timing, and perceived purpose.

Future work should therefore involve longitudinal, in-the-wild studies with target user groups. A multi-week field deployment (e.g., 2–4 weeks) using the full questionnaire content and realistic prompting schedules would allow evaluation of sustained engagement under naturalistic conditions. Such a study should combine interaction logs (e.g., response rates, latency, missed prompts), wear-time indicators, and periodic self-reports of perceived burden and usefulness, complemented by qualitative feedback through interviews or diaries. Analysis would focus on engagement patterns over time and on identifying when and why disengagement occurs. Such analysis should consider contextual and protocol-related factors (e.g., time of day, questionnaire length, or perceived relevance of items) that may influence sustained participation. This approach would make it possible to distinguish interface-related issues from protocol-level and motivational factors, and to identify conditions under which smartwatch-based ESM remains acceptable or becomes burdensome over time.

Chapter 6

CONCLUSION

Smartwatch-based experience sampling offers clear potential for longitudinal health research. However, the small display and limited input options (e.g., small touch targets and minimal text entry) can increase interaction burden and raise the risk of input errors, which can affect response quality. This thesis addressed this challenge by designing, implementing, and evaluating a smartwatch ESM interface that supports fast, glanceable, low-effort self-reporting. The prototype was implemented as a custom front-end in Experiencer for delivering ESM prompts and collecting self-reports (and sensor data when available), with GameBus as the backend for study configuration and secure data management. The work was conducted in the context of a Maastricht University research project on GPP. Although the project delivered a functional end-to-end pipeline, the main contribution of this thesis lies in the front-end interaction design and its usability evaluation.

A usability study with healthy participants ($N = 12$), using a proxy questionnaire, indicated that the interface supports efficient and intuitive interaction on a smartwatch without prior training. Participants reported confident responding and valued interaction choices intended to reduce mis-taps. At the same time, the evaluation identified practical design considerations: backward navigation was not always discoverable, and longer labels reduced readability on small displays. As the study assessed first-time use in a laboratory setting, the results demonstrate short-term usability rather than sustained engagement. Long-term participation in experience sampling is shaped not only by interface usability but also by protocol-level factors such as questionnaire content, prompting frequency, and perceived relevance, which require further investigation in real-world settings.

In summary, this thesis delivers (1) a functional smartwatch-based ESM prototype implemented in Experiencer and integrated with GameBus, and (2) empirical usability evidence for the proposed front-end interaction flow on circular Wear OS smartwatches. The results suggest that the interface enables low-effort microinteractions, while also identifying concrete improvements needed before deployment. Beyond the implementation, the thesis provides practical design recommendations for wearable ESM interfaces and offers a basis for future field evaluations, supporting longitudinal, multi-dimensional data collection in health research.

BIBLIOGRAPHY

Arias, O., Groehler, J., Wolff, M., & Choi, S. D. (2023). Assessment of musculoskeletal pain and physical demands using a wearable smartwatch heart monitor among precast concrete construction workers: A field case study. *Applied Sciences*, *13*(4). https://doi.org/10.3390/app13042347

Avila, F. R., McLeod, C. J., Huayllani, M. T., Boczar, D., Giardi, D., Bruce, C. J., Carter, R. E., & Forte, A. J. (2021). Wearable electronic devices for chronic pain intensity assessment: A systematic review [Publisher Copyright: © 2021 World Institute of Pain]. *Pain Practice*, *21*(8), 955–965. https://doi.org/10.1111/papr.13047

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Bergeron, S., Corsini-Munt, S., Aerts, L., Rancourt, K., & Rosen, N. O. (2015). Female sexual pain disorders: A review of the literature on etiology and treatment. *Current Sexual Health Reports*, *7*(3), 159–169. https://doi.org/10.1007/s11930-015-0053-y

Bergeron, S., Pâquet, M., Steben, M., & Rosen, N. O. (2021). Perceived partner responsiveness is associated with sexual well-being in couples with genito-pelvic pain. *Journal of Family Psychology*, *35*(5), 628. https://doi.org/10.1037/fam0000829

Black, D. W., & Grant, J. E. (2014). *Dsm-5® guidebook: The essential companion to the diagnostic and statistical manual of mental disorders*. American Psychiatric Publishing.

Brooke, J. (1996). SUS—a quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Butler, P. M., Yang, J., Brown, R., et al. (2025). Smartwatch- and smartphone-based remote assessment of brain health and detection of mild cognitive impairment. *Nature Medicine*, *31*, 829–839. https://doi.org/10.1038/s41591-024-03475-9

Carpenter, R. W., Wycoff, A. M., & Trull, T. J. (2016). Ambulatory assessment: New adventures in characterizing dynamic processes [PMID: 26887808]. *Assessment*, *23*(4), 414–424. https://doi.org/10.1177/1073191116632341

Castilla, D., Botella, C., Garcia-Palacios, A., Farfallini, L., & Miralles, N. (2012). Experience-sampling methodology with a mobile device in fibromyalgia. *International Journal of Telemedicine and Applications*. https://doi.org/10.1155/2012/162673

Choi, J. H., & Lee, H.-J. (2012). Facets of simplicity for the smartphone interface: A structural model. *International Journal of Human-Computer Studies*, *70*(2), 129–142. https://doi.org/10.1016/j.ijhcs.2011.09.002

Deniz-Garcia, A., Fabelo, H., Rodriguez-Almeida, A. J., Zamora-Zamorano, G., Castro-Fernandez, M., Alberiche Ruano, M. D. P., Solvoll, T., Granja, C., Schopf, T. R., Callico, G. M., Soguero-Ruiz, C., Wägner, A. M., & Consortium, W. (2023). Quality, usability, and effectiveness of mhealth apps and the role of artificial intelligence: Current scenario and challenges. *Journal of Medical Internet Research*, *25*, e44030. https://doi.org/10.2196/44030

Ellison, W. D., Trahan, A. C., Pinzon, J. C., Gillespie, M. E., Simmons, L. M., & King, K. Y. (2020). For whom, and for what, is experience sampling more accurate than retrospective report? *Personality and Individual Differences*, *163*, 110071. https://doi.org/10.1016/j.paid.2020.110071

Fillingim, R. B., Loeser, J. D., Baron, R., & Edwards, R. R. (2016). Assessment of chronic pain: Domains, methods, and mechanisms. *The Journal of Pain*, *17*(9), T10–T20. https://doi.org/10.1016/j.jpain.2015.08.010

Google. (2025a). *Color | wear os*. Google. Retrieved October 2, 2025, from https://developer.android.com/design/ui/wear/guides/m2-5/styles/color

Google. (2025b). *Design principles | wear os*. Google. Retrieved September 20, 2025, from https://developer.android.com/design/ui/wear/guides/m2-5/foundations/design-principles

Haase, I. (2023). Accuracy of retrospective pain measurement in patients with chronic pain. *Medicine International*, *3*. https://doi.org/10.3892/mi.2023.95

Hentati, A., Forsell, E., Ljótsson, B., Kaldo, V., Lindefors, N., & Kraepelien, M. (2021). The effect of user interface on treatment engagement in a self-guided digital problem-solving intervention: A randomized controlled trial. *Internet Interventions*, *26*, 100448. https://doi.org/10.1016/j.invent.2021.100448

Hernandez, J., McDuff, D., Infante, C., Maes, P., Quigley, K., & Picard, R. (2016). Wearable ESM: Differences in the experience sampling method across wearable devices. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 195–205. https://doi.org/10.1145/2935334.2935340

Horstmann, K. (2020, April). Experience sampling and daily diary studies: Basic concepts, designs, and challenges. Academic Press. https://doi.org/10.1016/b978-0-12-813995-0.00030-3

Intille, S., Haynes, C., Maniar, D., Ponnada, A., & Manjourides, J. (2016). μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1124–1128. https://doi.org/10.1145/2971648.2971717

Intille, S., Rondoni, J., Kukla, C., Iacono, I., & Bao, L. (2003). A context-aware experience sampling tool. *Proceedings of the Conference on Human Factors and Computing Systems: Extended Abstracts*. https://doi.org/10.1145/765891.766101

Khanshan, A., Van Gorp, P., & Markopoulos, P. (2023). Experiencer: An open-source context-sensitive wearable experience sampling tool. In A. Tsanas & A. Triantafyllidis (Eds.), *Pervasive computing technologies for healthcare* (pp. 315–331). Springer Nature Switzerland.

Khanshan, A., Van Gorp, P., Nuijten, R., & Markopoulos, P. (2021). Assessing the influence of physical activity upon the experience sampling response rate on wrist-worn devices. *International Journal of Environmental Research and Public Health*, *18*, 10593. https://doi.org/10.3390/ijerph182010593

Lewandowski, A. S., Palermo, T. M., Kirchner, H. L., & Drotar, D. (2009). Comparing diary and retrospective reports of pain and activity restriction in children and adolescents with chronic pain conditions. *The Clinical Journal of Pain*, *25*(4), 299–306. https://doi.org/10.1097/AJP.0b013e31819a6cfe

Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human–Computer Interaction*, *34*(7), 577–590. https://doi.org/10.1080/10447318.2018.1455307

Lewis, J. R., & Sauro, J. (2018). Item benchmarks for the system usability scale. *J. Usability Studies*, *13*(3), 158–167.

Li, J., Ponnada, A., Wang, W.-L., Dunton, G., & Intille, S. (2024). Ask less, learn more: Adapting ecological momentary assessment survey length by modeling question-answer information gain. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, *8*(4). https://doi.org/10.1145/3699735

Löchner, J., Moshe, I., Schiepek, G., Schuller, B. W., Schoedel, R., Rodgar, M., Wac, K., & Seizer, L. (2024). Tracking health-related quality of life dynamics: Advances in ambulatory assessment methods. *Brain Behavior and Immunity Integrative*, *6*, 100061. https://doi.org/10.1016/j.bbii.2024.100061

Lund, I., Lundeberg, T., Sandberg, L., Budh, C. N., Kowalski, J., & Svensson, E. (2005). Lack of interchangeability between visual analogue and verbal rating pain scales: A cross sectional description of pain etiology groups. *BMC Medical Research Methodology*, *5*(1). https://doi.org/10.1186/1471-2288-5-31

Markopoulos, P., Khanshan, A., Bormans, S., Tisza, G., Kang, L., & Van Gorp, P. (2023). Comparative evaluation of touch-based input techniques for experience sampling on smartwatches. *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*, 74–86. https://doi.org/10.1145/3626705.3627768

May, M., Junghaenel, D. U., Ono, M., Stone, A. A., & Schneider, S. (2018). Ecological momentary assessment methodology in chronic pain research: A systematic review. *The Journal of Pain*, *19*(7), 699–716. https://doi.org/10.1016/j.jpain.2018.01.006

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine*, *39*(9), 1533–1547. https://doi.org/10.1017/S0033291708004947

Neshati, A., Sakamoto, Y., Leboe-McGowan, L. C., Leboe-McGowan, J., Serrano, M., & Irani, P. (2019). G-sparks: Glanceable sparklines on smartwatches. *Proceedings of the 45th Graphics Interface Conference*. https://doi.org/10.20380/GI2019.23

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic Books.

Patel, A. M., Baxter, W., & Porat, T. (2024). Toward guidelines for designing holistic integrated information visualizations for time-critical contexts: Systematic review. *Journal of Medical Internet Research*, *26*, e58088. https://doi.org/10.2196/58088

Previtali, D., Boffa, A., Martino, A. D., Deabate, L., Delcogliano, M., & Filardo, G. (2022). Recall bias affects pain assessment in knee osteoarthritis: A pilot study [PMID: 36345999]. *CARTILAGE*, *13*(4), 50–58. https://doi.org/10.1177/19476035221118417

Rashid, S., & Nemati, A. (2024). Human-centered iot-based health monitoring in the healthcare 5.0 era: Literature descriptive analysis and future research guidelines. *Discover Internet of Things*, *4*(26). https://doi.org/10.1007/s43926-024-00082-5

Reiter, T., Sakel, S., Scharbert, J., ter Horst, J., van Zalk, M., Back, M., Bühner, M., & Schoedel, R. (2025). Side effects of experience-sampling protocols: A systematic analysis of how they affect data quality, data quantity, and bias in study results. *Advances in Methods and Practices in Psychological Science*, *8*(3), 25152459251347274. https://doi.org/10.1177/25152459251347274

Roos, L. G., & Slavich, G. M. (2023). Wearable technologies for health research: Opportunities, limitations, and practical and conceptual considerations. *Brain, Behavior, and Immunity*, *113*, 444–452. https://doi.org/10.1016/j.bbi.2023.08.008

Schneider, S., Junghaenel, D. U., Broderick, J. E., Ono, M., May, M., & Stone, A. A. (2021). Ii. indices of pain intensity derived from ecological momentary assessments and their relationships with patient functioning: An individual patient data meta-analysis. *The Journal of Pain*, *22*(4), 371–385. https://doi.org/10.1016/j.jpain.2020.10.002

Scollon, C., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, *4*, 5–34. https://doi.org/10.1023/A:1023605205115

Seidman, L. C., Temme, C. R., Zeltzer, L. K., Rapkin, A. J., Naliboff, B. D., & Payne, L. A. (2020). Ecological momentary assessment of non-menstrual pelvic pain: Potential pathways of central sensitization in adolescents and young adults with and without primary dysmenorrhea. *Journal of Pain Research*, *13*, 3447–3456. https://doi.org/10.2147/JPR.S283363

Shahrestani, A., Van Gorp, P., Le Blanc, P., Greidanus, F., de Groot, K., & Leermakers, J. (2017). Unified health gamification can significantly improve well-being in corporate environments. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4507–4511. https://doi.org/10.1109/EMBC.2017.8037858

Shen, S., Qi, W., Zeng, J., Li, S., Liu, X., Zhu, X., Dong, C., Wang, B., Shi, Y., Yao, J., Wang, B., Lou, X., Gu, S., Li, P., Wang, J., Jiang, G., & Cao, S. (2025). Passive sensing for mental health monitoring using machine learning with wearables and smartphones: Scoping review. *Journal of Medical Internet Research*, *27*, e77066. https://doi.org/10.2196/77066

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annu. Rev. Clin. Psychol.*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Stern, M., Dillman, D., & Smyth, J. (2007). Visual design, order effects, and respondent characteristics in a self-administered survey. *Survey Research Methods*, *1*, 121–138. https://doi.org/10.18148/srm/2007.v1i3.600

Stone, A., & Broderick, J. (2007). Real-time data collection for pain: Appraisal and current status. *Pain medicine (Malden, Mass.)*, *8 Suppl 3*, S85–93. https://doi.org/10.1111/j.1526-4637.2007.00372.x

Sutcliffe, S., Gallop, R., Henry Lai, H. H., Andriole, G. L., Bradley, C. S., Chelimsky, G., Chelimsky, T., Quentin Clemens, J., Colditz, G. A., Erickson, B., et al. (2019). A longitudinal analysis of urological chronic pelvic pain syndrome flares in the multidisciplinary approach to the study of chronic pelvic pain (mapp) research network. *BJU international*, *124*(3), 522–531. https://doi.org/10.1111/bju.14783

Tang, L. M., Meyer, J., Epstein, D. A., Bragg, K., Engelen, L., Bauman, A., & Kay, J. (2018). Defining adherence: Making sense of physical activity tracker data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, *2*(1). https://doi.org/10.1145/3191769

Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. *Evidence Based Mental Health*, *21*(3), 116–119. https://doi.org/10.1136/eb-2018-102891

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*. https://doi.org/10.1037/a0017653

Verhagen, S., Hasmi, L., Drukker, M., van Os, J., & Delespaul, P. (2016). Use of the experience sampling method in the context of clinical trials. *Evidence-Based Mental Health*, *19*(3), 86–89. https://doi.org/10.1136/ebmental-2016-102418

Volsa, S., Lewetz, D., Mlakic, V., Bertagnoli, C., Hochstöger, S., Rechl, M., Sertic, H., Batinic, B., & Stieger, S. (2024). Development of an open-source solution to facilitate the use of one-button wearables in experience sampling designs. *Behavior Research Methods*, *56*(6), 5876–5899. https://doi.org/10.3758/s13428-023-02322-y

Volsa, S., Batinic, B., & Stieger, S. (2022). Self-reports in the field using smartwatches: An open-source firmware solution. *Sensors*, *22*(5). https://doi.org/10.3390/s22051980

Wang, G., Ren, G., Hong, X., Peng, X., Li, W., & O'Neill, E. (2022). Freehand gestural selection with haptic feedback in wearable optical see-through augmented reality. *Information*, *13*(12), 566. https://doi.org/10.3390/info13120566

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037//0022-3514.54.6.1063

Whitmore, N. W., Chan, S., Zhang, J., Chwalek, P., Chin, S., & Maes, P. (2024). Improving attention using wearables via haptic and multimodal rhythmic stimuli. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3613904.3642256

Wrzus, C., & Neubauer, A. B. (2023). Ecological momentary assessment: A meta-analysis on designs, samples, and compliance across research fields [PMID: 35016567]. *Assessment*, *30*(3), 825–846. https://doi.org/10.1177/10731911211067538

Zarski, A.-C., Baumeister, H., & Kählke, F. (2025). DSM-5 genito-pelvic pain/penetration disorder: Prevalence, comorbidities, and associated factors in university students. *International Journal of Clinical and Health Psychology*, *25*, 100529. https://doi.org/10.1016/j.ijchp.2024.100529

Appendix A

GANTT CHARTS

This appendix provides the detailed project planning timeline used to structure and monitor the EngD project. The Gantt charts illustrate the sequencing of activities across different phases. In the following figures, the left-hand panel lists all planned activities and milestones with their full names, whereas the timeline on the right visualizes their temporal sequencing.

Note: During execution, the schedule required adjustments primarily due to availability constraints rather than changes in scope, deliverables, or risk events. The overall sequencing and dependencies remained unchanged, but the timing of later phases shifted accordingly.

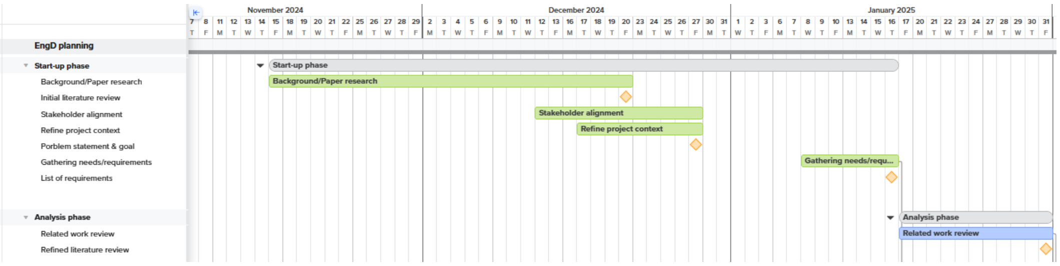


Figure A.1: Gantt chart for the start-up and analysis phases.

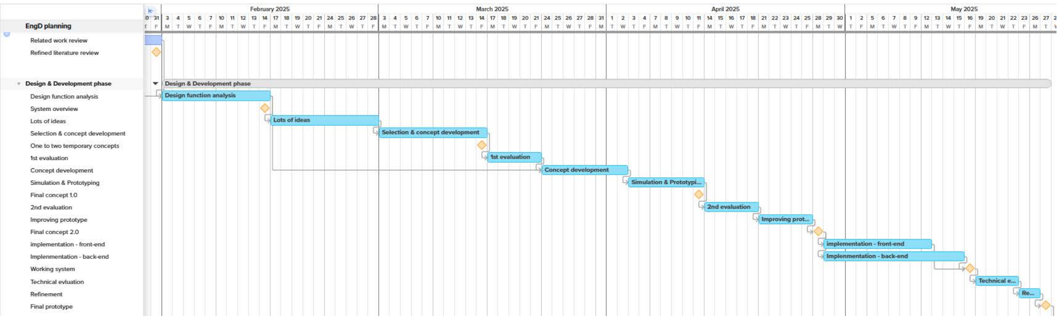


Figure A.2: Gantt chart for the design & development phase.

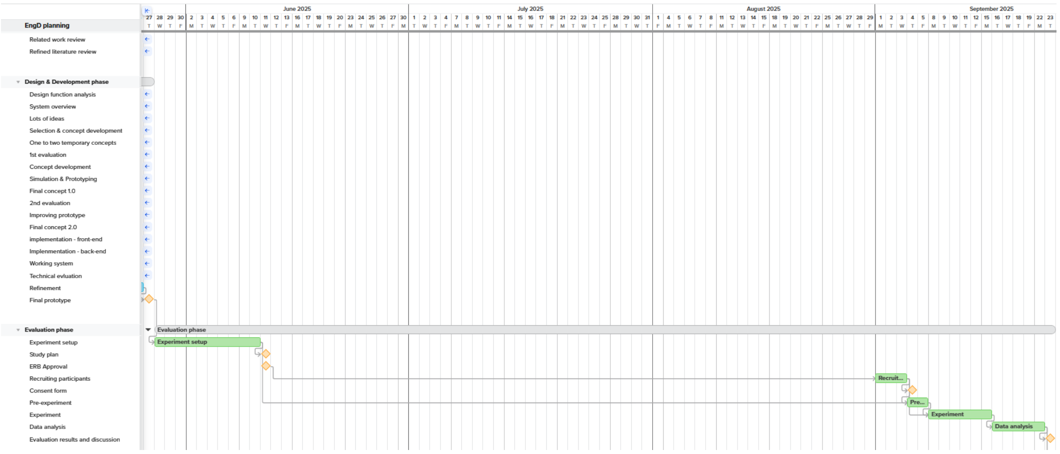


Figure A.3: Gantt chart for the (part of) evaluation phase.

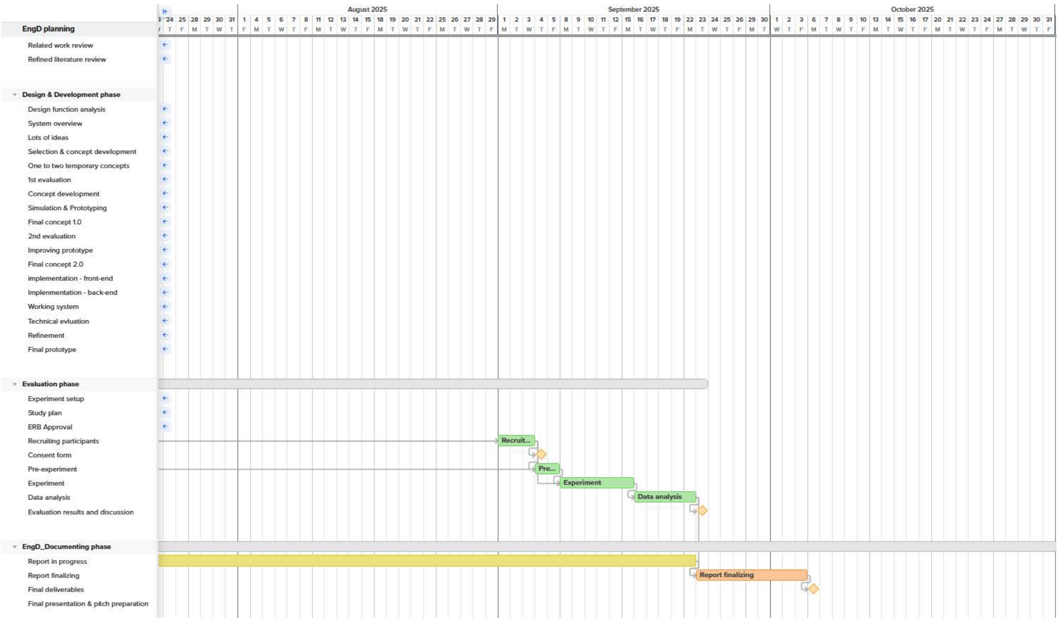


Figure A.4: Gantt chart for the (part of) evaluation and reflection & reporting phases.

Appendix B

USER MANUAL

Note: User account credentials have been removed from this manual for privacy and ethical considerations.

Power On & Basic Settings

1. Hold the side button until the Samsung logo appears.
2. Swipe down and navigate to:

   `Settings > General > Date & time`

   Ensure the **Time zone** is correct.
3. Swipe down from the top of the watch face to access the **Quick Panel** and select the **Always On** icon (represented by a watch symbol).

Connect to Wi-Fi

1. Go to:

   `Settings > Connections > Wi-Fi`
2. Select your network, enter the password, and ensure the status shows **Connected**.
3. If the watch does not connect:
   - Double-check the password
   - Wait a moment and try again

Note: The watch cannot reliably connect to public Wi-Fi and may take longer than expected to connect to home or office Wi-Fi. This behavior is normal.

Experiencer App Initialization

Initialization codes for the GPP project are already set up. If re-initialization is needed, see Appendix for the `wearable_code` (has been intentionally removed for privacy and ethical considerations).

Launch and Grant Permissions

1. Tap **Experiencer** to open the app.
2. Tap **Allow** on all permission dialogs:
    - Location
    - Sensors
    - Notifications

Initialize with User Code

1. Tap **Initialize watch** and select **Enter code**.
2. Enter the unique code from Appendix I (`wearable_code`).
3. Tap **Submit** to bind the watch to the participant's GameBus account.

Share this code with the participant in case they need to re-initialize the watch.

Daily Management and Data Upload

Every evening, participants should:

- Place the watch on its charger
- Keep Wi-Fi enabled to allow data upload

ESM Protocol Updates

For any changes to the ESM protocol (e.g., survey content, start/end time, duration), please contact Tianqin. Once the update is confirmed, follow these steps:

1. Tap **Experiencer** eight times to open the **Check debug info** panel.
2. Follow the figures provided to apply the update.
3. Tap **Experiencer** eight times again to close the panel.

Appendix C

TECHNICAL TEST RESULTS (EXCERPT)

Note: The ESM questionnaire is unpublished research material owned by Maastricht University and is therefore not included. Sensor logs are omitted due to privacy and data protection considerations.

Answer	Index	Item label	Timestamp	Session
5	1	ITEM_01	2025-09-02 14:51:33	SESSION_A
5	2	ITEM_02	2025-09-02 14:51:33	SESSION_A
5	3	ITEM_03	2025-09-02 14:51:33	SESSION_A
7	4	ITEM_04	2025-09-02 14:51:33	SESSION_A
7	5	ITEM_05	2025-09-02 14:51:33	SESSION_A
7	6	ITEM_06	2025-09-02 14:51:33	SESSION_A
5	7	ITEM_07	2025-09-02 14:51:33	SESSION_A
5	8	ITEM_08	2025-09-02 14:51:33	SESSION_A
Called	9	ITEM_09	2025-09-02 14:51:33	SESSION_A
Yes, serious	10	ITEM_10	2025-09-02 14:51:33	SESSION_A
7	11	ITEM_11	2025-09-02 14:51:33	SESSION_A
4	12	ITEM_12	2025-09-02 14:51:33	SESSION_A
5	13	ITEM_13	2025-09-02 14:51:33	SESSION_A
5	14	ITEM_14	2025-09-02 14:51:33	SESSION_A
5	15	ITEM_15	2025-09-02 14:51:33	SESSION_A
5	16	ITEM_16	2025-09-02 14:51:33	SESSION_A
5	1	ITEM_01	2025-08-27 17:00:14	SESSION_B
5	2	ITEM_02	2025-08-27 17:00:14	SESSION_B
4	3	ITEM_03	2025-08-27 17:00:14	SESSION_B
4	4	ITEM_04	2025-08-27 17:00:14	SESSION_B
4	5	ITEM_05	2025-08-27 17:00:14	SESSION_B
6	6	ITEM_06	2025-08-27 17:00:14	SESSION_B
4	7	ITEM_07	2025-08-27 17:00:14	SESSION_B
6	8	ITEM_08	2025-08-27 17:00:14	SESSION_B
Called	9	ITEM_09	2025-08-27 17:00:14	SESSION_B
Yes, serious	10	ITEM_10	2025-08-27 17:00:14	SESSION_B
3	11	ITEM_11	2025-08-27 17:00:14	SESSION_B
3	12	ITEM_12	2025-08-27 17:00:14	SESSION_B
3	13	ITEM_13	2025-08-27 17:00:14	SESSION_B
4	14	ITEM_14	2025-08-27 17:00:14	SESSION_B
5	15	ITEM_15	2025-08-27 17:00:14	SESSION_B
6	16	ITEM_16	2025-08-27 17:00:14	SESSION_B

Table C.1: Edited excerpt of ESM responses exported from GameBus during technical evaluation. Identifiers are masked and item texts are omitted.

Appendix D

PROXY QUESTIONNAIRE

No.	Proxy Question	Proxy UI Scale Type	Affect	Original UI Scale Type	Validated Scale / Rationale
1	Did you engage in physical activity?	Categorical (3 options)	Functional	Categorical (3 options)	Tests original three-choice UI and conditional logic
2	Rate your physical discomfort's impact on activities.	Likert (1–7)	Functional	Likert (1–7)	Proxy for conditional branching question
3	Rate your physical well-being.	Likert (1–7)	Functional	Likert (1–7)	Proxy for multi-option branching question
4	I felt distressed.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
5	I felt upset.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
6	I felt jittery.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
7	I felt scared.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
8	I felt ashamed.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
9	I felt guilty.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
10	I felt irritable.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
11	I felt enthusiastic.	Likert (1–7)	Positive	Likert (1–7)	PANAS (Positive Affect)
12	I felt hostile.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)
13	I felt proud.	Likert (1–7)	Positive	Likert (1–7)	PANAS (Positive Affect)
14	Did you have contact with your friend?	Categorical (4 options)	Functional	Categorical (4 options)	Tests four-option categorical UI and contact-based logic
15	Did you have a conflict with your friend?	Categorical (3 options)	Functional	Categorical (3 options)	Proxy for conditional branching question
16	How satisfied were you with the resolution?	Likert (1–7)	Functional	Likert (1–7)	Direct test of scale functionality
17	Felt supported by my friend.	Likert (1–7)	Functional	Likert (1–7)	Direct test of scale functionality
18	Felt satisfied with my friendship.	Likert (1–7)	Functional	Likert (1–7)	Direct test of scale functionality
19	I felt excited.	Likert (1–7)	Positive	Likert (1–7)	PANAS (Positive Affect)
20	I felt active.	Likert (1–7)	Positive	Likert (1–7)	PANAS (Positive Affect)
21	I felt nervous.	Likert (1–7)	Negative	Likert (1–7)	PANAS (Negative Affect)

Table D.1: Proxy questionnaire used in the usability evaluation. The proxy items were derived from the English women's version of the original questionnaire, which is not included due to confidentiality requirements.

Appendix E

SYSTEM USABILITY SCALE (SUS)

Participants indicated their level of agreement with the following statements on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*):

1. **Q1.** I think that I would like to use this system frequently.
2. **Q2.** I found the system unnecessarily complex.
3. **Q3.** I thought the system was easy to use.
4. **Q4.** I think that I would need the support of a technical person to be able to use this system.
5. **Q5.** I found the various functions in this system were well integrated.
6. **Q6.** I thought there was too much inconsistency in this system.
7. **Q7.** I would imagine that most people would learn to use this system very quickly.
8. **Q8.** I found the system very cumbersome to use.
9. **Q9.** I felt very confident using the system.
10. **Q10.** I needed to learn a lot of things before I could get going with this system.

Appendix F

SEMI-STRUCTURED INTERVIEW GUIDE

The following interview guide was used to structure the semi-structured interviews.

First Impressions and Overall Experience

- What was your first impression of the smartwatch interface when you started using it?
- How did you feel about the overall experience of answering questions on the watch? Did it feel intuitive, or did you need time to figure it out?

Ease of Navigation and Interaction

- How easy or difficult was it to understand what to do when a new question appeared?
- How easy or difficult was it to move between questions?
- Were the response options (e.g., buttons) easy to select and confirm?
- Did you feel confident that your responses were recorded correctly? Why (not)?

Readability and Presentation

- Was the text size, layout, and contrast comfortable to read on the smartwatch screen?

Burden, Satisfaction, and Suggestions

- Did answering the questions feel quick and manageable, or too long for a smartwatch interaction?
- If you imagine receiving these questions several times a day, how comfortable would you feel responding regularly?
- What was your favorite aspect of using the app?
- What was your least favorite aspect?

- If you could change one thing about the design, what would it be?

Additional Feedback

- Do you have any final comments or suggestions about the app's user interface?