Experiencer: An Open-Source Context-Sensitive Wearable Experience Sampling Tool

Alireza Khanshan[1](✉), Pieter Van Gorp[2], and Panos Markopoulos[1]

[1] Department of Industrial Design, Eindhoven Artificial Intelligence Systems Institute, Eindhoven University of Technology, 5612 Eindhoven, AZ, The Netherlands
{a.khanshan,p.markopoulos}@tue.nl
[2] Department of Industrial Engineering, Eindhoven University of Technology, 5612 Eindhoven, AZ, The Netherlands
p.m.e.v.gorp@tue.nl

Abstract. We introduce Experiencer, a newly developed Experience Sampling Method (ESM) software for commodity-level smartwatches. We designed this software mainly to address the compliance-related challenges, such as dropouts of study participants, that generations of ESM software solutions have faced. Dropouts are often caused by the inconvenient frequency and timing of the ESM prompts. This can partly be mitigated by utilizing physiological smartwatch sensors to learn which prompting moments are both convenient to the study participant and also relevant to the ESM study designer. Experiencer enables researchers to configure context-sensitive sampling protocols, providing access to raw sensor data, within the boundaries of European privacy legislation. In this paper, we describe the technical capabilities of our software, compare its features with the state-of-the-art, and showcase its application in studies that used Experiencer.

Keywords: Experience Sampling Method · Wearable ESM · mHealth · Ubiquitous Computing · Smartwatch application · Wearables · Software Framework

1 Introduction

Studying what people do, feel, and think during their daily routines is essential to understanding the dynamics of the human psyche [13]. Although the interest for such information originates in psychological research, it extends to a variety of other domains where humans are studied (e.g., healthcare [33], media studies [43], or education [6]). Actions and thoughts can be registered in diaries. Registering frequently and in a real-life context can result in entries with high ecological validity [41]. However, poor compliance to journaling protocols and human forgetfulness are well-documented challenges for the diary method [10]. Therefore, it is preferable to distribute self-report prompts throughout daily life rather

A. Tsanas and A. Triantafyllidis (Eds.): PH 2022, LNICST 488, pp. 315–331, 2023.
https://doi.org/10.1007/978-3-031-34586-9_21

than journaling retrospectively [3,32]. Such a process of collecting self-reported data about behaviors, thoughts, or feelings during the day-to-day activities of humans, is commonly called the Experience Sampling Method (ESM [32]). The same approach with an emphasis on psychological research is known as Ecological Momentary Assessment (EMA [48]). In the field of human-computer interaction, context sensing technologies were emerging two decades ago and led to the reintroduction of ESM as Contextual Aware Experience Sampling (CAES [24]) with the emphasis on utilizing context sensing to optimize the sampling procedures. For the sake of simplicity and consistency, we use the term ESM in the remainder of this paper to encompass the aforementioned concepts. Additionally, we introduce the term wESM (wearable ESM) to refer to the use of wearables (e.g., smartwatches) instead of or together with smartphones or other mediums to handle both collection of self-reported data as well as sensor data.

Considering that ESM aims to prompt during daily life, finding opportune moments that do not interfere with one's daily activities to deliver such self-report prompts (beeps) becomes essential. By doing so, the likelihood of disturbing study participants becomes lower, hence, compliance is potentially increased [29]. One common approach to detecting opportune moments is by longitudinal monitoring via physiological sensors to better perceive the momentary context of respondents.

In the context of wESM, interpretation of sensor data usually requires some knowledge about the context in which the sensor was used as well as the participant-specific subjective data. Thus, self-reports are critical to provide ground truth and complement the sensor data. Then again, study participants may skip self-reports, especially when studies last longer than a few days. How to minimize dropouts and maximize compliance requires further research, especially in the context of longitudinal studies. Typically, compliance is measured by calculating the dropout rate, response rate, response time, resolution time, volunteering rate, and the amount of presented information which oftentimes indicated the poor compliance of the participants [18,31,35,41,50,54].

Arguably, the choice of ESM device used for prompting and data entry, as well as the prompting schedule affects the participation experience and consequently impacts the extent to which the ESM process is perceived as seamless or in contrast, burdensome by study participants [20,23,29,34,49]. Recent advances in commodity-level smartwatches in terms of interactivity, connectivity, and embedded sensing technology, offer new opportunities for using them as wESM devices, which could help reduce the obtrusiveness of wESM signals and increase their availability as they are wrist-worn. Furthermore, they provide a quick-and-concise data entry interface, and their sensors pave the way to provide context-sensitive prompting regimes [46,47]. In the following, we introduce Experiencer [26], our open-source context-sensitive smartwatch-based wESM software that provides researchers with advanced data gathering features while striving to attain compliance of participants. Experiencer has been designed with the aim to enable researchers to flexibly configure their experiment protocol to potentially alleviate response fatigue and sustain sufficient response rates that are otherwise hampered by traditional scheduling regimes (e.g., random sampling) during ESM studies [12,18,55].

In the following, we first review research on ESM-related devices to derive requirements for wESM support. Subsequently, we introduce Experiencer by providing conceptual and technical information on client and server components and their interactions. We also demonstrate how Experiencer was already used in a variety of wESM studies. Afterward, we discuss limitations, weaknesses, and directions for future work. Finally, we summarize our contributions in the conclusion section and encourage the readers to utilize Experiencer for their own research.

2 Related Works

In this section, we initially summarize the history of digital ESM devices and their evolution over time. Then, we focus on the contemporary smartwatch-based wESM solutions and examine the state-of-the-art in the domain.

2.1 History of Digital ESM Devices

While the experience sampling *method* emerged already in the eighties [32], digital devices were gaining popularity around the turn of the century (e.g., Electronic Mood Device [22] by Hoeksma et al. or the Experience Sampling Program [3] by Barret et al.).

Due to the eventual prevalence of smartphones since the late 2000s,s, and the software development kits (SDK) supported by their operating systems (early on by PalmOS and Windows CE, and more recently by Android, and iOS), more advanced ESM tools have been created. The applications developed in the context of ESM have mainly sought to alleviate the complexity of configuring ESM protocols for researchers by offering custom configuration schemes (e.g., Momento [11], AndWellness [21], PsyMate [36], Tempest [4], [5], and formR [2]) and facilitating research-focused data collection (e.g., Funf [1], RADAR-base [42], and HOPES [53]). Meanwhile, fewer works have accounted for context awareness. By incorporating context-sensitive strategies the researchers can capture data at specific events (e.g., context-aware experience sampling tool [24], AWARE [17], and Paco [16]).

More recently, commodity-level wearables such as smartwatches are being utilized in the context of ESM as well. They offer SDKs on par with smartphones and their physiological sensors are more accurate and reliable since they are on the skin, rather than in the pocket. In addition, they are optimized to collect physiological (e.g., body activity) data more continuously.

2.2 Towards wESM Smartwatch Applications

In this section, we survey the state-of-the-art in the application domain for smartwatches utilized for ESM studies.

Intille et al. focused on the amount and length of interruption, and the difficulty of accessing the device of typical Ecological Momentary Assessment (EMA) delivered via smartphones [23]. They implemented μEMA as a smartwatch extension to smartphones that delivered prompts on the smartwatch as well as concise versions of ESM questions. A study was conducted where μEMA on a smartwatch was compared with EMA exclusively on a phone. Despite an $\approx$8 times increase in the number of interruptions, μEMA had a significantly higher compliance rate, completion rate, and first prompt response rate, and was perceived as less distracting. Although μEMA [23] suggested that a substantially higher prompting rate than EMA, may yield higher response rates and a lower participation burden, Ponnada et al. aimed to assess the validity of participant responses from μEMA self-reports [40]. It was concluded that for physical activity registrations, high-frequency μEMA self-reports were consistent with activities detected by a research-grade continuous sensor, even when prompting up to 72 times per day. This demonstrated that μEMA study participants were not carelessly answering prompts by randomly tapping on the smartwatch. Then again, the experiment by Ponnada et al. lasted only one week, so further development of smart prompting protocols may be needed to reduce participant burden and enhance compliance in longer studies.

Blaauw et al. presented Physiqual, a platform for researchers that gathers and integrates data from commercially available sensors and service providers into one unified format for use in ESM, and Quantified Self (QS) [9]. The Physiqual platform allows researchers to aggregate and integrate physiological sensor data with ESM. Although such a platform does not provide a dedicated wESM application for smartwatches, it facilitates the aggregation of such pre-existing data sources.

Kheirkhahan et al. developed a smartwatch-based framework for real-time and online assessment and mobility monitoring (ROAMM) [30]. The smartwatch application was used to collect and pre-process data. A server was used for storing and retrieving data, remote monitoring, data visualization, and summary statistics, and for other administrative purposes. Although the smartwatch app allowed configurable sensors and supports different types of studies, it is not openly available and does not support context-sensitive scheduling.

Hafiz et al. showed a strong correlation between the data gathered via their domain-specific smartwatch application and computer-based tests in a lab setting [19]. The aims of their study were to evaluate the Ubiquitous Cognitive Assessment Tool (UbiCAT), a smartwatch-based platform they developed to assess cognitive performance, to investigate its usability, and to understand participants' perceptions regarding the use of a smartwatch in cognitive assessment.

Park et al. developed a framework for collecting and analyzing physiological data using smartwatches in the wild and demonstrated its robustness away from controlled laboratory settings [38]. Their system sent random notifications during the day asking questions about subjective well-being. They concluded that methodological research needs to study how to interpret continuous physiological signals obtained through such platforms. Once such understanding is developed, sensing can potentially reduce the burden of self-reporting.

Collectively, these works demonstrate the feasibility, benefits, and pitfalls of using smartwatches in ESM studies. Unfortunately, however, the underlying software systems are not available for elaboration by other scholars.

3 Requirements for a Smartwatch-Based wESM Tool

In our analyses of the aforementioned state-of-the-art, we observed that none of the platforms was offered as a free, reliable, and continued service to other academics. Furthermore, we identified the lack of a thorough exploration of the design space concerning commodity-level smartwatches. In the following, we propose our software requirements for designing a smartwatch-based wESM tool that we realized by analyzing existing solutions, exploring the design space, and empirically during the development of Experiencer.

The flexible nature of ESM protocols demands guidelines that researchers can follow to set up their ESM studies. However, earlier considerations on tools to support ESM studies(e.g., [14], [44], and [39]) need to be reviewed to address the opportunities and challenges introduced by wearable devices and modern software technology. For example, the non-reactivity guideline proposed by Delespaul [14]: reactivity can be minimized by using small, reliable, and inexpensive devices that produce unpredictable prompts and can be fully employed within a range of environmental constraints [14]. Such concerns can now be addressed by using a commodity-level smartwatch that can run configurable wESM software.

By surveying the latest developments in this domain, literature study, running our own experiments each with at least 50 participants (the number of participants was determined based on power analysis, the number of available smartwatches, and the recruitment process [29,37]), and discussing the requirements with respective ESM researchers, we distilled a shortlist of features that wESM platforms should provide. We found four highly relevant multi-purpose wESM solutions that explicitly used the smartwatches as their main client or specifically developed to incorporate smartwatches as third-party devices in their ecosystem. We realized six categories of features to list the key similarities and differences between the aforementioned solutions: *Data collection and analysis*, *Scheduling*, *Data entry*, *Monitoring interface*, *Scalability*, and *Optimization*. Moreover, we added *Openness, availability, and security* as an additional category, to make clear that no solution so far could be used by other scholars continuously.

Table 1 provides a comparative overview of the software features of the aforementioned platforms in those seven categories. We define these features as follows:

Recording sensor data refers to the recording and storing of the data captured via physiological sensors such as an accelerometer and a heart rate monitor.

Table 1. Feature Comparison of recent wESM platforms

Category	Feature	Software				
		WellBeat [38]	*ROAMM* [30]	*μEMA* [23]	*Physiqual* [9]	*Experiencer* [26]
Data collection and analysis	Recording sensor data	✓	✓	✓	✓	✓
	Data analytics dashboard	✓	✓	✗	✓	✗
	Configurable sensors	✗	✓	✗	✗	✓
Scheduling	Context-sensitive	✗	✗	✗	✗	✓
	Temporal	✗	✓	✗	✗	✓
Data entry	Input widgets on the smartwatch	✓	✓	✓	✗	✓
	Configurable widgets	✗	✓	✗	✓	✓
Web interface	Data visualization	✗	✓	✓	✗	✗
	Administration dashboard	✗	✓	✗	✓	✗
Scalability	Remote device management	✗	✓	✗	✗	✓
Optimization	Event-based data collection	✗	✗	✗	✗	✓
	Custom data synchronization	✓	✓	✓	✗	✓
Openness, availability, and security	Reusability and availability	✗	✗	✗	✗	✓
	GDPR compliance	✗	✗	✗	?	✓

Data analytics dashboard enables the realization of statistical reports and/or discovery of patterns in the collected raw data after data is stored on the server. Such processing can ideally be customized by the study owner.

Configurable sensors indicates the possibility to enable the desired set of sensors, setting specific sampling frequencies, and setting the duration of the sensor data recording period.

Context-sensitive refers to the possibility of sending beeps at contextual events of interest rather than at random or interval-based times. Such events can refer to simpler conditions such as "send a prompt *when sedentary activity is detected*", or more complex ones such as "send a prompt *5 min after a running activity is ended*", or "send a prompt *as soon as the minimum heart rate during 10 straight minutes is higher than 100bmp*".

Temporal refers to the capability of defining wESM protocols, through configurable settings rather than hard-coded, in time-based manners that are signal-contingent (i.e., random) or interval-contingent (as defined in [8,14]).

Input widgets on the smartwatch turn the wearable into a data entry device rather than just a notification device (which would still require smartphone interactions). Typical widgets are *text inputs*, *radio buttons*, or *drop-downs*.

Configurable widgets facilitates the creation or customization of the look and feel of input widgets. This is usually supported by either HTML scripting or by parameterizing pre-built widget components.

Data visualization provides computer-generated representations of the data in form of charts (e.g., histogram, and scatter plot) in a dedicated dashboard.

Administration dashboard allows the creation and management of wESM protocols, monitoring the activity of participants, and accessing the collected data of the study via a graphical user interface (GUI).

Remote device management allows controlling and monitoring the wESM electronic device (e.g., smartwatch) remotely. Such a feature allows fast and easy scale-out by facilitating the configuration of numerous devices at once. The control over the device also helps with restraining specific out-of-the-box features of the device (e.g., disabling GPS to enhance battery life and ensure privacy, disallowing the study participants to install apps on their device, or preventing factory reset). Remote access can also help with updating the aforementioned constraints on the fly and seamlessly during the study according to researchers' requirements or participants' convenience.

Event-based data collection enables intermittent collection of data triggered by events of interest; e.g., recording heart rate data solely during answering a questionnaire. This eases the matching of sensory data with self-report periods and greatly reduces battery consumption compared to when data is recorded continuously.

Custom data synchronization supports the buffering of data on the wearable and smart synchronization with the server, potentially also balancing battery life with the information needs of the study owner. The wESM app may monitor the WiFi coverage, assess the Internet connection stability, and then transfer the data in controlled transactions to assure data persistence and consistency.

Reusability and availability is unfortunately seldom seen in the latest wESM solutions. Most of them have either become obsolete already (not maintained after a specific study), or they were never designed to be reusable (only for a specific use case, research question, or domain). The few solutions that are potentially reusable were never provided as a service to other scholars, Experiencer is open-source [27] and its back-end, GameBus, is open-access [51].

GDPR compliance The General Data Protection Regulation (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information from individuals who live in the European Union (EU) [15]. Compliance with the GDPR would lessen privacy concerns especially for EU users (be it researchers or participants). During our assessment, we could not find relevant or explicit information regarding the GDPR and privacy policy for most of the solutions. In the case of Physiqual, a platform that connects the different third-party tools, its conformity partially relies on such third-party tools (e.g., Google Fit, Fitbit, etc.). Thus, we put **?** in the table.

Note that none of the four tools support context-sensitive sampling. Nonetheless, such a feature can contribute greatly to the study participants (i.e., interruptions or prompts are received at more opportune moments resulting in less burden and fatigue), as well as the researchers (i.e., responses recorded at more opportune moments are less biased). Therefore, we did treat it as a requirement for Experiencer. The same holds for the event-based data collection feature.

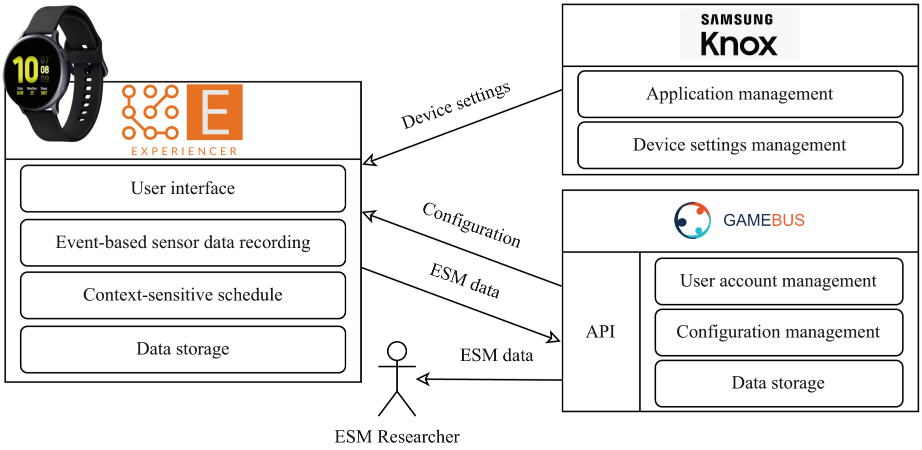


Fig. 1. Overview of system components and communications

4 Configuring wESM Protocols with Experiencer

Experiencer is a context-sensitive wESM tool that allows for recording of sensor data, configuring sensors, remote device management, event-based data collection, various sampling regimes, dynamic user interface (UI), and also optimizes device data storage and data transactions over the network while being open-access, available, and compliant with standard privacy measures. The back-end of Experiencer is built on top of GameBus, an open-access health data management platform that is offered non-commercially by academia [45]. GameBus, designed following the GDPR-oriented privacy and security measures, guarantees that all data is stored exclusively in Europe and provides to its users full control over their data. Experiencer also builds upon Knox, an industrial-strength device management system by Samsung. Specifically, Knox is used for the remote (re-)installation and (re-)configuration of Experiencer. Figure 1 visualizes this modular software architecture.

The behavior of Experiencer is defined by its *configuration* [27]. The current version of Experiencer allows setting the configuration through its API. Composing a configuration is the first step to conduct a study using Experiencer. That includes setting an inter-notification time value, specifying sensor(s) settings, defining contextual rules for context-sensitive sampling (if not, the fixed interval policy is adopted automatically), and defining a questionnaire. Table 2 overviews the capabilities proposed by our configurations. Each participant's account is linked to a specific smartwatch. Next, the participant-smartwatch pairs are linked to a configuration and a study (that is named by the researcher) (Fig. 2). During the linkage, the researcher can apply different configurations to the participants of the same study. By setting the configuration during the linkage procedure, the researcher can construct their treatment groups (e.g., by setting some participants to configuration A and some others to configuration

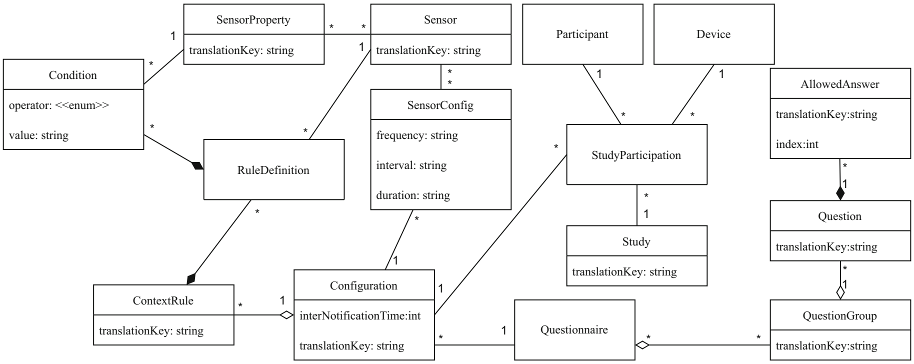


Fig. 2. Provisional UML class diagram of Participant, Device, Configuration, and Questionnaire. All translationKey attribute values should be unique.

B). Lastly, the final configuration (including the inter-notification time, sensor settings, questionnaire, etc.) is transferred from the server to the smartwatch over WiFi. After the successful transmission, the app is ready to use. In addition, the app periodically checks for configuration updates (e.g., the addition of a new question to the questionnaire, or a new value for the inter-notification time), so changes can be applied even after devices have been deployed to study participants.

Experiencer consumes a JSON-formatted configuration [27]. An essential part of this configuration is its questionnaire (Fig. 2). Experiencer is equipped with a parser that interprets the questionnaire dynamically and renders a layout based on the configurations. The current version of Experiencer supports questions that are chained in a sequence (e.g., a 2-question PANAS questionnaire starts with Fig. 3a followed by Fig. 3b). Additionally, if a researcher desires a specific and more complex UI for the questionnaire, they can program their own interface using HTML, CSS, and JavaScript.

Questionnaires in Experiencer are made up of question groups where each holds a set of one or more questions with their corresponding answers (Fig. 2). Question groups allow the researchers to categorize their questions. For example, a questionnaire can start with a number of *general* questions followed by a set of *personality*-related questions (*general* and *personality* are example categories). Such grouping aims at a more distinctive and informative user experience.

The key notion in creating a questionnaire is the *translation keys*. Translation keys are unique strings that serve both as identifiers and localization means. They are used to look up the GameBus-provided human-readable texts (translations) in different languages. Given such translations, the questionnaire object of the configuration can be transformed into a UI (similar to Fig. 3a and Fig. 3b). The generated UI is shown to the participants when they press the self-report button (Fig. 3c), or it appears automatically after a scheduled prompt. Once a participant responds to a question, their response is stored locally and then transferred to the server transactionally when a stable network connec-

Table 2. Configuration options of Experiencer

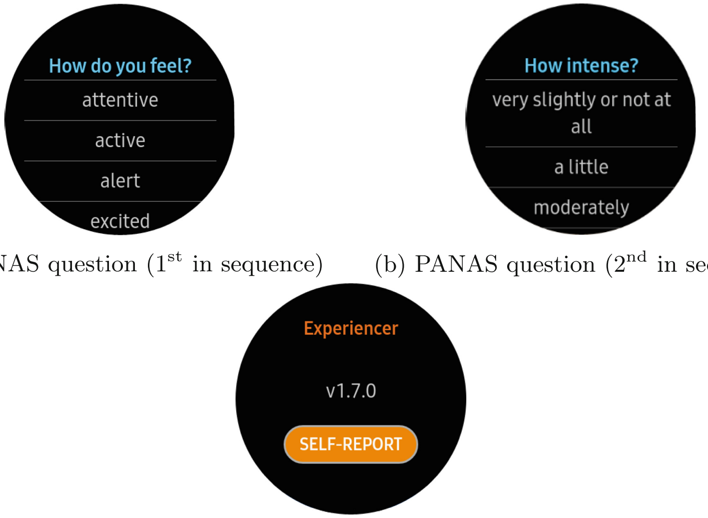
Configuration	Options	Description
Sampling policy	Context-sensitive sampling or fixed interval	Experiencer is programmed to read data from different sensor and can be configured to beep based on sensor data. For example, Experiencer can be configured to send beeps only when study participants engage in vigorous physical activities or vice versa [29]. The fixed interval policy provides the classic interval-based prompting regime when desired.
Inter-notification time	An integer value indicating inter-notification time	The inter-notification time, determines the time in-between each beep. The role of inter-notification time for the fixed interval policy is to determine the period between two beeps. For context-sensitive policies, the inter-notification time determines the cool-down period.
Unobtrusive sensing	Accelerometer, photoplethysmography, heart rate, peak-peak interval, body activity sensors	Body activity data is continuously monitored to ensure the accuracy of event-contingent policies. It also serves as a means to know if the smartwatch is being worn. The other sensors, if chosen by the researcher, are recorded during the period that a participant is filling an ESM form (e.g., a questionnaire), by default, for a maximum of 1 min. Continuous recording is also possible.
Compliance status	Timestamps related to beep received, read, and response submission times	To analyze compliance, the time when a beep is received, read, and submitted are by default recorded to facilitate the calculation of compliance-related indices such as response delay.
Questionnaire	Questions and answers as string literals	The set of ESM questions can be defined by the researcher. The list is then parsed within the app and represented to the study participant when a self-report procedure is started. The questionnaire in the current version is sequential rather than branched [14]

tion is detected. The GameBus back-end already pre-defines a variety of such questions. Still, researchers with ESM protocols involving custom questions can request the addition of such items [52].

5 Case Studies Run with Experiencer

We created Experiencer for 1) effective and accurate context sensitivity to help increase compliance, and 2) openness, availability, and security. Regarding the former, not only do we have run various wESM studies but also continuously plan to test different hypotheses regarding context percipiency. Our learned lessons from such studies help us fine-tune the context-sensitivity of Experiencer over time by continuously training models that can be leveraged for decision-support on expected compliance of protocols. For the latter, on the one hand, we allow interested researchers from various domains to set up their experiments with Experiencer. Moreover, we actively maintain and develop our software following the best practices and guidelines of software engineering and security [26].

To demonstrate the aforementioned capabilities and flexibility of our solution, below we describe the different wESM studies that utilized Experiencer along with their goal, and their outcome. The study-specific configuration files are publicly accessible via the project's GitHub examples repository [27].



(a) PANAS question (1st in sequence) (b) PANAS question (2nd in sequence)

(c) The main screen of Experiencer

Fig. 3. Screenshots of Experiencer application as it was configured for the SamenGezond 2020 and 2021 campaigns

The SamenGezond 2020 and 2021 Campaigns. During two health promotion campaigns [37] we used Experiencer, to assess the effects of physical activity upon experience sampling response rate on smartwatches. We adopted a context-sensitive schedule to prompt half of the participants at subtle and the other half at vigorous physical activity levels and observed a significant difference in response rates depending on such context [29].

GGz Centraal. To measure and predict the stress levels of a subgroup of GGz Centraal mental healthcare facility patients, Experiencer is being used since 2021 to collect valence and arousal data throughout the day. This ongoing study also helps assess the adaptability of Experiencer in targeting various cohorts. To meet the requirements of the target group, a custom user interface was created by exploiting the rotary capabilities of the smartwatches.

Persuasion Profiling. In another wESM study, Experiencer was used to capture student motivation, and to assess its influence on the response rate of students while applying persuasion profiling. Persuasion profiling involves tailoring notifications to users (the content of experience sampling prompts in this case) according to an individual's susceptibility to known social influence strategies [25]. On the other hand, the researcher's script to generate a custom data entry widget was incorporated.

Affective State. To complement our findings concerning opportune moments of interruption based on the context [29], in a collaborative wESM study to measure time-lagged associations between affective state, sleep, and several other lifestyle-related behaviors, Experiencer was set up to deliver experiential questions compliant with a context-sensitive schedule based on physical activity. Following the researcher's constraints, the app was configured to send beeps between habitual wake-up and sleep times.

6 Future Work and Limitations

We have sought to identify the requirements of researchers and study participants so far. Markedly, the addition of branching questionnaires where the transition of one question to the next is controlled by a rule-based system is crucial. Such a rule-based system takes into account the response to a question and/or perceived context to determine the subsequent question(s). Additionally, the context-sensing could be improved by including contextual information derived from device usage patterns or others alike related to participants' behavior to complement the sensory rules. Moreover, a smartwatch-based wESM software should include a wide range of user interface elements such as checkboxes, radio buttons, sliders, etc. to support a substantial subset of ESM studies (excluding those that are currently difficult to do because of the small smartwatch screen size and difficulty of text input). In addition, to simplify the interactions between the researcher and wESM software, interactive dashboards for the administration of wESM protocol configuration, as well as real-time monitoring and visualization of the data should be designed. Indeed such flexibilities, by design, do not guarantee higher compliance with respect to response rate and retention since researchers may apply sampling regimes that are deemed intrusive. Accordingly, the domain lacks fine-grained ESM-related data that would otherwise enable data-intensive approaches such as data-driven modeling and machine learning to help address compliance-related challenges in the ESM domain especially by learning response patterns [28,56]. Thus, more collective effort is required to share such data openly with other scholars.

Regardless, smartwatch-based wESM tools are affected by some limitations, such as their rather short battery life especially when sensor data is recorded for longer periods of time and with high frequencies. Although Experiencer provides enough flexibility to alter the period of the data collection and the sampling frequency of the sensors based on the researcher's requirements, there is an inevitable trade-off between the accuracy (and volume) of the collected sensor data and the battery life for wESMs.

7 Conclusions

Experience Sampling is a very popular research method that is used in multiple fields to study the experiences and thoughts of respondents over sustained periods of time, by repeatedly prompting them to respond to survey questions. Experience sampling presents several methodological challenges, like participant burden, low compliance, etc., which researchers have traditionally attempted to address through technological developments. While currently, most experience sampling studies rely on smartphones [7], the form factor and sensing capabilities of smartwatches and their growing prevalence open up new opportunities for researchers wishing to apply experience sampling studies. This paper reviewed the few attempts that have been made so far to support ESM studies with wearable devices and discussed the requirements for modern ESM tools that exploit the recent technological advancements in smartwatches. We described the state of the art in the emerging area of wESM, where wearables and more specifically smartwatches can be used as signaling and reporting devices in ESM studies. We created Experiencer; a context-sensitive wearable experience sampling application on commodity-level smartwatches. By perceiving context through physiological sensors, our software allows researchers to configure sampling regimes that potentially align with the opportune moments of interrupts. Thus, the likelihood of sustaining sufficient compliance is increased. Besides that, the input widgets are all on the smartwatch, accessible on their wrist, and are resolvable with just a few taps. Besides, context-sensitiveness helps with interrupting at more opportune moments, thus, the participants are potentially less likely to be disturbed. We expect that such capabilities result in less burden for the participants during the study period, and improved compliance and data quality. Future research could evaluate through case studies the extent to which these benefits can be delivered. Furthermore, we argued for a specific set of requirements that wESM platforms need to address, and have shown how Experiencer improves them compared to the state-of-the-art and emphasized how Experiencer advances software availability and context sensitivity. Moreover, we describe the technical aspects of our software, look into its configurable features and review how Experiencer was configured and used in different ESM studies. Lastly, we acknowledge the improvements required in this domain and point out the barriers caused by the device restrictions and constraints.

Acknowledgment. Special recognition to Jory Schoondermark and Barbara Montagne at GGz Centraal, as well as Karin Smolders and Lars Giling at the Eindhoven University of Technology who have contributed to the development of Experiencer through their input and feedback on early versions of Experiencer. Lastly, special appreciation goes to Samsung Nederland for their excellent support on the smartwatches we utilized for our research and development.

References

1. Aharony, N., Gardner, A., Sumter, C.: funf | open sensing framework (2021). http://www.funf.org/
2. Arslan, R.C., Walther, M.P., Tata, C.S.: formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. Behav. Res. Methods **52**(1), 376–387 (2020). https://doi.org/10.3758/s13428-019-01236-y
3. Barrett, L.F., Barrett, D.J.: An introduction to computerized experience sampling in psychology. Soc. Sci. Comput. Rev. **19**(2), 175–185 (2001). https://doi.org/10.1177/089443930101900204
4. Batalas, N., Markopoulos, P.: Introducing tempest, a modular platform for in situ data collection. In: Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, pp. 781–782 (2012)

5. Batalas, N., aan het Rot, M., Khan, V.J., Markopoulos, P.: Using tempest: End-user programming of web-based ecological momentary assessment protocols. Proceedings ACM Hum. Comput. Interact. **2**(EICS), 1–24 (2018)
6. Becker, E.S., Goetz, T., Morger, V., Ranellucci, J.: The importance of teachers' emotions and instructional behavior for their students' emotions-an experience sampling analysis. Teach. Teach. Educ. **43**, 15–26 (2014)
7. van Berkel, N., Ferreira, D., Kostakos, V.: The experience sampling method on mobile devices. ACM Comput. Surv. (CSUR) **50**(6), 93:1–93:40 (2017). https://doi.org/10.1145/3123988
8. van Berkel, N., Goncalves, J., Lovén, L., Ferreira, D., Hosio, S., Kostakos, V.: Effect of experience sampling schedules on response rate and recall accuracy of objective self-reports. Int. J. Hum Comput Stud. **125**, 118–128 (2019). https://doi.org/10.1016/j.ijhcs.2018.12.002
9. Blaauw, F.J., et al.: Let's get physiqual - an intuitive and generic method to combine sensor technology with ecological momentary assessments. J. Biomed. Inf. **63**, 141–149 (2016). https://doi.org/10.1016/j.jbi.2016.08.001
10. Bolger, N., Davis, A., Rafaeli, E.: Diary methods: capturing life as it is lived. Annu. Rev. Psychol. **54**(1), 579–616 (2003)
11. Carter, S., Mankoff, J., Heer, J.: Momento: support for situated ubicomp experimentation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 125–134. Association for Computing Machinery (2007). https://doi.org/10.1145/1240624.1240644
12. Collins, R.L., Kashdan, T.B., Gollnisch, G.: The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. Exp. Clin. Psychopharmacol. **11**(1), 73–78 (2003). https://doi.org/10.1037/1064-1297.11.1.73
13. Csikszentmihalyi, M., Larson, R.: Validity and reliability of the experience-sampling method. In: Flow and the Foundations of Positive Psychology, pp. 35–54. Springer, Dordrecht (2014). https://doi.org/10.1007/978-94-017-9088-8_3
14. Delespaul, P.A.E.G.: Technical note: devices and time-sampling procedures. In: Vries, M.W.d. (ed.) The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings, pp. 363–374. Cambridge University Press (1992). https://doi.org/10.1017/CBO9780511663246.033
15. European Parliament: General Data Protection Regulation (GDPR) (2016). https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04
16. Evans, B.: Paco-applying computational methods to scale qualitative methods, vol. 2016, no. 1, pp. 348–368 (2016). https://doi.org/10.1111/1559-8918.2016.01095
17. Ferreira, D., Kostakos, V., Dey, A.K.: AWARE: mobile context instrumentation framework. Front. ICT **2**, 6 (2015). https://doi.org/10.3389/fict.2015.00006
18. Fuller-Tyszkiewicz, M., Skouteris, H., Richardson, B., Blore, J., Holmes, M., Mills, J.: Does the burden of the experience sampling method undermine data quality in state body image research? Body Image **10**(4), 607–613 (2013). https://doi.org/10.1016/j.bodyim.2013.06.003
19. Hafiz, P., Bardram, J.E.: The ubiquitous cognitive assessment tool for smartwatches: design, implementation, and evaluation study. JMIR Mhealth Uhealth **8**(6), e17506 (2020). https://doi.org/10.2196/17506
20. Hernandez, J., McDuff, D., Infante, C., Maes, P., Quigley, K., Picard, R.: Wearable ESM: differences in the experience sampling method across wearable devices. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 195–205. MobileHCI 2016, Association for Computing Machinery (2016). https://doi.org/10.1145/2935334.2935340

21. Hicks, J., Ramanathan, N., Kim, D., Monibi, M., Selsky, J., Hansen, M., Estrin, D.: AndWellness: an open mobile system for activity and experience sampling. In: Wireless Health 2010, pp. 34–43. WH 2010, Association for Computing Machinery (2010). https://doi.org/10.1145/1921081.1921087
22. Hoeksma, J.B., Sep, S.M., Vester, F.C., Groot, P.F.C., Sijmons, R., De Vries, J.: The electronic mood device: design, construction, and application. Behav. Res. Methods Instrum. Comput. **32**(2), 322–326 (2000). https://doi.org/10.3758/BF03207801
23. Intille, S., Haynes, C., Maniar, D., Ponnada, A., Manjourides, J.: μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 1124–1128. UbiComp 2016, Association for Computing Machinery (2016). https://doi.org/10.1145/2971648.2971717
24. Intille, S.S., Rondoni, J., Kukla, C., Ancona, I., Bao, L.: A context-aware experience sampling tool. In: CHI 2003 Extended Abstracts on Human Factors in Computing Systems. pp. 972–973. CHI EA 2003, Association for Computing Machinery (2003). https://doi.org/10.1145/765891.766101
25. Kaptein, M., Markopoulos, P., de Ruyter, B., Aarts, E.: Personalizing persuasive technologies: explicit and implicit personalization using persuasion profiles. Int. J. Hum Comput Stud. **77**, 38–51 (2015)
26. Khanshan, A.: Experiencer, the experience sampling method software (2021). https://experiencer.eu/
27. Khanshan, A.: Experiencer ESM source code (2022). https://github.com/khnshn/Experiencer
28. Khanshan, A.: From simulation to reality and back again: a hybrid approach to estimate the compliance of ESM study participants to different ESM protocols. In: 14th ACM SIGCHI Symposium on Engineering Interactive Computing Systems Doctoral Consortium, EICS DC 2022. 21–24 June 06 2022 (2022). http://eics.acm.org/eics2022/submission_dc.html
29. Khanshan, A., Van Gorp, P., Nuijten, R., Markopoulos, P.: Assessing the influence of physical activity upon the experience sampling response rate on wrist-worn devices. Int. J. Environ. Res. Public Health **18**(20), 10593 (2021). https://doi.org/10.3390/ijerph182010593
30. Kheirkhahan, M., et al.: A smartwatch-based framework for real-time and online assessment and mobility monitoring. J. Biomed. Inf. **89**, 29–40 (2019). https://doi.org/10.1016/j.jbi.2018.11.003
31. Kini, S.: Please take my survey: compliance with smartphone-based EMA/ESM studies (2013). https://digitalcommons.dartmouth.edu/senior_theses/83/
32. Larson, R., Csikszentmihalyi, M.: The experience sampling method. New Dir. Methodol. Soc. Behav. Sci. **15**, 41–56 (1983)
33. Larson, R., Csikszentmihalyi, M.: The experience sampling method. In: Flow and the Foundations of Positive Psychology, pp. 21–34. Springer, Dordrecht (2014). https://doi.org/10.1007/978-94-017-9088-8_2
34. Manini, T.M., et al.: Perception of older adults toward smartwatch technology for assessing pain and related patient-reported outcomes: pilot study. JMIR Mhealth Uhealth **7**(3), e10044 (2019). https://doi.org/10.2196/10044
35. Morren, M., van Dulmen, S., Ouwerkerk, J., Bensing, J.: Compliance with momentary pain measurement using electronic diaries: a systematic review. Eur. J. Pain **13**(4), 354–365 (2009). https://doi.org/10.1016/j.ejpain.2008.05.010

36. Myin-Germeys, I., Birchwood, M., Kwapil, T.: From environment to therapy in psychosis: a real-world momentary assessment approach. Schizophrenia Bull. **37**(2), 244–247 (2011). https://doi.org/10.1093/schbul/sbq164
37. Nuijten, R., et al.: Health promotion through monetary incentives: evaluating the impact of different reinforcement schedules on engagement levels with a mHealth app. Electronics **10**(23), 2935 (2021)
38. Park, S., Constantinides, M., Aiello, L.M., Quercia, D., Van Gent, P.: WellBeat: a framework for tracking daily well-being using smartwatches. IEEE Internet Comput. **24**(5), 10–17 (2020). https://doi.org/10.1109/MIC.2020.3017867
39. Pejovic, V., Lathia, N., Mascolo, C., Musolesi, M.: Mobile-based experience sampling for behaviour research. In: Tkalčič, M., De De Carolis, B., de de Gemmis, M., Odić, A., Košir, A. (eds.) Emotions and Personality in Personalized Services. HIS, pp. 141–161. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-31413-6_8
40. Ponnada, A., Thapa-Chhetry, B., Manjourides, J., Intille, S.: Measuring criterion validity of microinteraction ecological momentary assessment (micro-EMA): Exploratory pilot study with physical activity measurement. JMIR Mhealth Uhealth **9**(3), e23391 (2021). https://doi.org/10.2196/23391
41. Ram, N., Brinberg, M., Pincus, A.L., Conroy, D.E.: The questionable ecological validity of ecological momentary assessment: considerations for design and analysis. Res. Hum. Dev. **14**(3), 253–270 (2017). https://doi.org/10.1080/15427609.2017.1340052
42. Ranjan, Y., et al.: RADAR-base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth **7**(8), e11734 (2019). https://doi.org/10.2196/11734
43. Redmiles, E.M., Bodford, J., Blackwell, L.: I just want to feel safe: a diary study of safety perceptions on social media. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 13, pp. 405–416 (2019)
44. Rough, D.J., Quigley, A.: End-user development of experience sampling smartphone apps-recommendations and requirements. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **4**(2), 1–19 (2020)
45. Shahrestani, A., Van Gorp, P., Le Blanc, P., Greidanus, F., de Groot, K., Leermakers, J.: Unified health gamification can significantly improve well-being in corporate environments. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4507–4511 (2017). https://doi.org/10.1109/EMBC.2017.8037858. ISSN: 1558-4615
46. Shin, D.H., Biocca, F.: Health experience model of personal informatics: the case of a quantified self. Comput. Hum. Behav. **69**, 62–74 (2017). https://doi.org/10.1016/j.chb.2016.12.019
47. Singh, G., Delamare, W., Irani, P.: D-SWIME: a design space for smartwatch interaction techniques supporting mobility and encumbrance. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–13. Association for Computing Machinery (2018). https://doi.org/10.1145/3173574.3174208
48. Stone, A.A., Shiffman, S.: Ecological momentary assessment (EMA) in behavorial medicine. Ann. Behav. Med. **16**(3), 199–202 (1994). https://doi.org/10.1093/abm/16.3.199
49. Timmermann, J., Heuten, W., Boll, S.: Input methods for the borg-RPE-scale on smartwatches. In: 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), pp. 80–83 (2015). https://doi.org/10.4108/icst.pervasivehealth.2015.259220. ISSN: 2153-1641

50. Trull, T.J., Ebner-Priemer, U.W.: Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. - PsycNET (2009). https://doi.org/10.1037/a0017653
51. Van Gorp, P.: Gamebus - social health games for the entire family (2015). http://www.gamebus.eu
52. Van Gorp, P., Surendranathan, A., Lesani, Y.: GameBus API guide (2021). http://devdocs.gamebus.eu
53. Wang, X., et al.: HOPES: an integrative digital phenotyping platform for data collection, monitoring, and machine learning. J. Med. Internet Res. **23**(3), e23984 (2021). https://doi.org/10.2196/23984
54. Wen, C.K.F., Schneider, S., Stone, A.A., Spruijt-Metz, D.: Compliance with mobile ecological momentary assessment protocols in children and adolescents: a systematic review and meta-analysis. J. Med. Internet Res. **19**(4), e6641 (2017). https://doi.org/10.2196/jmir.6641
55. Wheeler, L., Reis, H.T.: Self-recording of everyday life events: origins, types, and uses. J. Pers. **59**(3), 339–354 (1991). https://doi.org/10.1111/j.1467-6494.1991.tb00252.x
56. Zhang, C., Wang, S., Aarts, H., Dastani, M.: Using cognitive models to train warm start reinforcement learning agents for human-computer interactions. arXiv:2103.06160 (2021). arxiv.org/abs/2103.06160