

# **Improving the CARWatch Framework for Objective Cortisol Awakening Response Assessment**

## **Master's Thesis in Computer Science**

submitted  
by

Robert Schröter

born 02.10.1998 in Berlin

Written at

Machine Learning and Data Analytics Lab  
Department Artificial Intelligence in Biomedical Engineering  
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

in Cooperation with

Chair of Health Psychology, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Advisors: Robert Richer M. Sc, Luca Abel M. Sc., Prof. Dr. Bjoern Eskofier

Started: 01.01.2024

Finished: 02.07.2024



Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Richtlinien des Lehrstuhls für Bachelor- und Masterarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Erlangen, den 2. Juli 2024



## Übersicht

In dieser Arbeit wird die Weiterentwicklung des CARWatch-Frameworks beschrieben. CAR-Watch wurde entwickelt, um Studien zur Untersuchung der Cortisolawachreaktion (CAR) zu unterstützen, die eine genaue und objektive Bewertung der Aufwach- und Speichelprobenzeiten erfordern. CARWatch ist ein Open-Source-Framework, das aus einer Webapp, einer Android-App und einem Python-package besteht. Zu den Erweiterungen gehörten Funktionen zur Aufzeichnung zusätzlicher Daten, die für CAR-Bewertungen von Nutzen sein können, sowie Verbesserungen hinsichtlich der Anpassbarkeit, Flexibilität und Benutzeroberfläche der Anwendungen. Das Framework wurde anhand von Daten der Trierer Alltagsstressstudie evaluiert, einer von der Universität Trier durchgeführten Studie, die den Zusammenhang zwischen dem Stressniveau im Alltag und verschiedenen individuellen Variablen untersucht. Die für die Evaluation verwendete Studienphase umfasste 30 Teilnehmende, die angewiesen wurden, 6 Speichelproben pro Tag über 4 Tage zu nehmen, einschließlich 2 Arbeitstagen und 2 Nicht-Arbeitstagen. Die Android-App wurde als objektives Kontrollinstrument für die Probenahmezeiten verwendet. Es wurden Daten über das Probenahmeverhalten der Teilnehmenden sowie die CAR-Merkmale,  $AUC_G$  und  $AUC_I$  analysiert. Die erhobenen Daten zeigten ein konsistentes Probenahmeverhalten mit insgesamt geringen Probenahmeverzögerungen. Anders als in der bisherigen Forschung wurden, außer für berufstätige Teilnehmer, keine nennenswerten Unterschiede bei den Probenahmezeiten und den CAR-Merkmalen zwischen Arbeitstagen und Nicht-Arbeitstagen festgestellt. Dies könnte darauf hindeuten, dass die Auswirkung der Tagesform auf den CAR vom Beschäftigungsstatus des Einzelnen abhängt. Um diese Annahme zu beweisen, sind jedoch weitere Untersuchungen erforderlich.

## Abstract

This thesis describes the further development of the CARWatch framework. CARWatch is designed to support studies investigating the cortisol awakening response (CAR), which requires an exact and objective assessment of awakening and saliva sampling times. CARWatch is an open-source framework that consists of a webapp, an Android app, and a Python package. Enhancements included functionality to record additional data that can be beneficial for CAR assessments, and improvements regarding the customizability, flexibility, and user interface of the applications. The framework was evaluated using data from the Trier Everyday Stress Study, a study conducted by the University of Trier that examines the relationship between stress levels in everyday life and various individual variables. The study phase used for evaluation included 30 participants who were instructed to take 6 saliva samples per day over 4 days, including 2 workdays and 2 non-workdays. The Android app was used as an objective verification tool for sampling times. Data was analyzed on participants' sampling behavior as well as the CAR features  $AUC_G$  and  $AUC_I$ . The collected data indicated consistent sampling behavior with overall low sampling delays. In contrast to existing research, no noticeable differences in sampling times and CAR features between workdays and non-workdays were found except for employed participants. This might indicate that the effect of the day type on the CAR depends on the employment status of the individual. However, further investigation is needed to prove this assumption.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Related Work</b>	<b>5</b>
2.1	CAR Assessment . . . . .	5
2.2	Objective Verification Methods . . . . .	9
<b>3</b>	<b>Methods</b>	<b>13</b>
3.1	Framework overview . . . . .	13
3.2	Webapp . . . . .	15
3.3	App . . . . .	20
3.3.1	User Interface . . . . .	20
3.3.2	Features . . . . .	23
3.4	Evaluation . . . . .	30
3.4.1	Study design . . . . .	31
3.4.2	Data analysis . . . . .	31
<b>4</b>	<b>Results</b>	<b>35</b>
4.1	Data cleaning . . . . .	35
4.2	Analysis of CAR features . . . . .	36
4.2.1	CAR throughout the different reporting strategies . . . . .	36
4.2.2	CAR on workdays compared to CAR on non-workdays . . . . .	36
4.3	Analysis of sampling times . . . . .	37
4.3.1	Sampling times throughout the different reporting strategies . . . . .	37
4.3.2	Sampling delays on workdays compared to sampling delays on non-workdays	39
<b>5</b>	<b>Discussion</b>	<b>45</b>

<b>6 Conclusion &amp; Outlook</b>	<b>49</b>
<b>List of Figures</b>	<b>51</b>
<b>List of Tables</b>	<b>53</b>
<b>Bibliography</b>	<b>55</b>
<b>A Acronyms</b>	<b>65</b>



# Chapter 1

## Introduction

Experiencing stress is a regular part of our daily life and although stress is a healthy reaction of our body that can increase our reactivity and performance, long-term stress is known to have a negative influence on our health. Psychological stress is even associated with various diseases, such as depression, cardiovascular disease (CVD), and HIV/AIDS [Coh07]. The human body's response to a stressor is governed by two main stress pathways, the autonomic nervous system (ANS) and the hypothalamic-pituitary-adrenocortical (HPA) axis. The ANS triggers various protective mechanisms, such as increasing the heart rate and blood pressure. The activation of the HPA axis results in a release of glucocorticoid hormones that influence the distribution of the body's energy resources depending on the needs. One of those hormones is cortisol, which makes it an objective and well-established stress marker [Ulr09].

The cortisol concentration in blood and saliva usually follows a diurnal pattern. Cortisol levels sharply increase during the first 30 to 45 min after awakening. Then they decline steadily until reaching their nadir around midnight. In the second half of the night, they slowly rise again [Ada09]. The rapid increase in cortisol levels of 50 to 156 % after awakening [Clo04] is also known as cortisol awakening response (CAR). The CAR was first systematically described by Pruessner et al. [Pru97] and indicates the basal HPA axis activity [Chi09; Kud10]. CAR research quickly gained popularity, with over 700 publications related to CAR according to Scopus, due to its relation to several psychosocial factors like exhaustion, burnout, or different types of stress [Chi09; Clo04; Fri09; Kud10]. The CAR is typically assessed by measuring the cortisol levels in saliva samples upon awakening and in fixed intervals of 10 to 15 min throughout the first 30 to 60 min after awakening [Sta16; Sta22]. Most CAR studies use two parameters to analyze the CAR. First, the overall amount of cortisol released during the observed period, which is determined by computing the integrated area under the curve ( $AUC_G$ ), and second, the increase of cortisol from the level

recorded on waking ( $AUC_1$ ) [Chi09; Pru03]. Due to the sharp rise of cortisol levels within a short time period during the CAR, delayed sampling can have a large impact on those metrics. Precise sampling is therefore essential for CAR assessments but often hard to achieve since most studies are conducted ambulatory with unsupervised self-collection of saliva samples [Sta16; Sta22]. Researchers need to ensure that both awakening time and sampling times are recorded correctly and objectively to prevent incorrect conclusions caused by the participants.

The gold standard to measure sleep objectively is polysomnography (PSG), obtaining recordings from several channels, such as electroencephalography (EEG), electrooculography (EOG), electromyography (EMG), and electrocardiography (ECG). Afterwards, the recordings are analyzed to assess sleep parameters. Although already used in CAR research [Oku10], PSG is costly, labor-intensive, disruptive to the participants' normal sleep routines, and not applicable for ambulatory measurements [Sta16]. Thus, it is not suitable for large-scale field studies. More cost-effective methods for determining the wake-up time are wrist actigraphy [Lic06] or forced awakening [Sta16]. Recent studies also showed that consumer electronic wearables, such as fitness trackers or smartwatches, can be a suitable alternative to assess sleep parameters [Hag19; Chi22].

Regarding objective sampling time verification, the best practice is to use electronic monitoring devices, such as screw up bottles or boxes that record the times of bottle openings [Sta16]. To record the sampling times, the sampling devices are stored inside the bottle/box and the participants are instructed to open them before sampling so that the opening time stamps can be used as a measure for the sampling time. The main disadvantage hereby are costs and scalability as one box is needed for each sample and participant. A cheaper method is to use smartphones since most people already have access to a smartphone. Previous work has shown that smartphones can be utilized in CAR assessments as they provide multiple beneficial built-in functionalities. Smartphones were integrated into studies in different ways, e.g., by prompting the participant to take a selfie each time a sample was taken [Zhu19], presenting a unique code that the participants had to record on the label of the sampling device used [Bed19; Pow15; Pow12], or sending text messages to the participant via a customized software, which they had to answer within a short time window after taking the sample [Rod18]. While these methods offer a low-cost solution, they can lead to issues, such as privacy violations and inaccurate awakening times due to the reliance on participants' self-reporting their wake-up times. Furthermore, the software used in the last two approaches is not publicly available, which would incur additional costs for future studies.

Although available, most studies (68.5 %) still lack objective assessment methods [Sta22]. Existing methods are often not suited for studies with a limited budget as they require cost-intensive technology or software and cannot be integrated in studies that take place in the participants'

domestic settings. To solve this methodological issue, CARWatch was developed as a low-cost tool that enables an objective assessment of CAR sampling times [Ric23]. CARWatch is an open-source project designed for ambulatory studies, which utilizes smartphones and their in-built features, such as alarms and cameras to remind the participant of taking samples and to confirm the sampling by scanning a barcode attached to the sampling device. In the scope of this thesis, the CARWatch framework was enhanced by integrating access to external sleep recording devices, improving the user experience by providing an in-built introduction to the CARWatch app and a redesigned user interface, and increasing customizability to make CARWatch accessible to a broader range of researchers. The app was evaluated based on data from the Trier Everyday stress study (TriASS)<sup>1</sup> (originally: “Trierer Alltagsstress-Studie”), which was collected using the CARWatch framework.

---

<sup>1</sup><https://www.uni-trier.de/universitaet/fachbereiche-faecher/fachbereich-i/faecher-und-institute/psychologie/professuren/biologische-klinische-psychologie/forschung-v1/trierer-alltagsstress-studie>



# Chapter 2

## Related Work

The CAR can be described as a rapid increase in cortisol concentration during the first hour after awakening. It serves as a reliable biological marker for adrenocortical activity [Pru97] and can give insights into multiple psychosocial factors [Kud03a; Chi09].

The following Sections will explain the typical CAR assessment process, identify factors that can confound the results, and describe methods that minimize their impact to prevent erroneous conclusions.

### 2.1 CAR Assessment

CAR assessment requires determination of the cortisol levels throughout the first 30 to 60 min after awakening [Clo04; Clo10; Kud10; Eld14]. This is usually done by collecting either blood or saliva samples upon awakening and in intervals of 10 to 15 min thereafter for a total of up to 5 samples. Depending on the study settings, the sampling process is repeated over the course of up to 5 days [Sta16]. A large proportion of CAR studies take place in the participants' domestic settings and rely on salivary cortisol samples that are easily collectible and minimally invasive to minimize costs and maintain the participants' daily routines [Sta16]. To guarantee the stability of the cortisol concentration, participants are usually instructed to place the samples in their home freezer and return them to the laboratory where they are analyzed as soon as possible. Concerning the measurement of cortisol concentration, antibody-based immunoassays are usually used [Sta16; Cho22]. In an ambulatory study, awakening and sampling times are often recorded by the participants themselves, typically using pen-and-paper-based protocols, which increases the error-proneness, especially since the temporal accuracy of reported awakening and sampling times has a large impact on the CAR curve derived from the measurements.

### *Delay between awakening and first sample*

Kupper et al. were the first to describe awakening time related sampling inaccuracies [Kup05]. They conducted a study with more than 700 participants of which 77 showed a negative CAR. Additional ECG and body movement recordings of 59 of those participants revealed a time difference between reported awakening and measured awakening of 42 min (range between 10 min and 135 min) for 80 % of them which caused the inaccuracies.

To show the impact of inaccurate awakening times, Smyth et al. conducted a study with 50 healthy female participants [Smy13]. Each participant had to collect one sample (S1) immediately after awakening and three follow up samples (S2, S3, and S4) in intervals of 15 min thereafter. The samples were collected over 4 days. The awakening time was determined by self-report as well as actigraphy, a motion-based method to measure sleep (for further explanation see Section 2.2). On average, the self-reported awakening time was around 4 min later than the awakening time estimated by actigraphy. To show the actual impact of delayed awakening reports, the cortisol measurements of adherent days (delay < 5 min) were compared to the measurements of non-adherent days (delay of 10 to 15 min). The results showed that CAR estimates were significantly higher on non-adherent days compared to adherent days with cortisol levels peaking at S3 for non-adherent days and S4 for adherent days. According to Smyth et al., this can be explained by a short time lag of 10 to 15 min between awakening and cortisol rise (see Figure 2.1) which would not have been detected if the self-reported awakening times were used [Smy13]. As depicted in Figure 2.2, increasing delays also have an increasing impact on the  $AUC_I$  [Sta16].

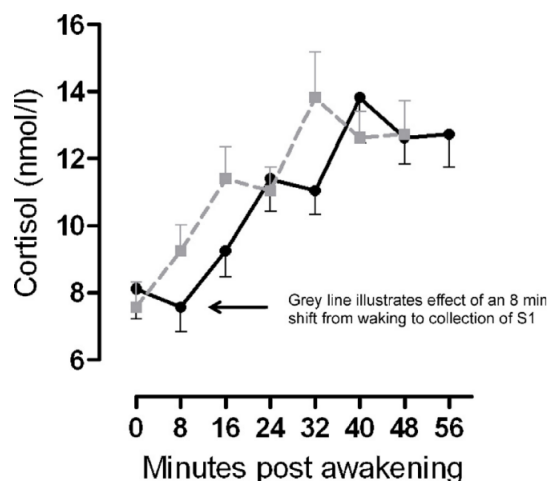


Figure 2.1: Illustration of the impact of delayed awakening time reporting on the estimated CAR profile [Smy13].

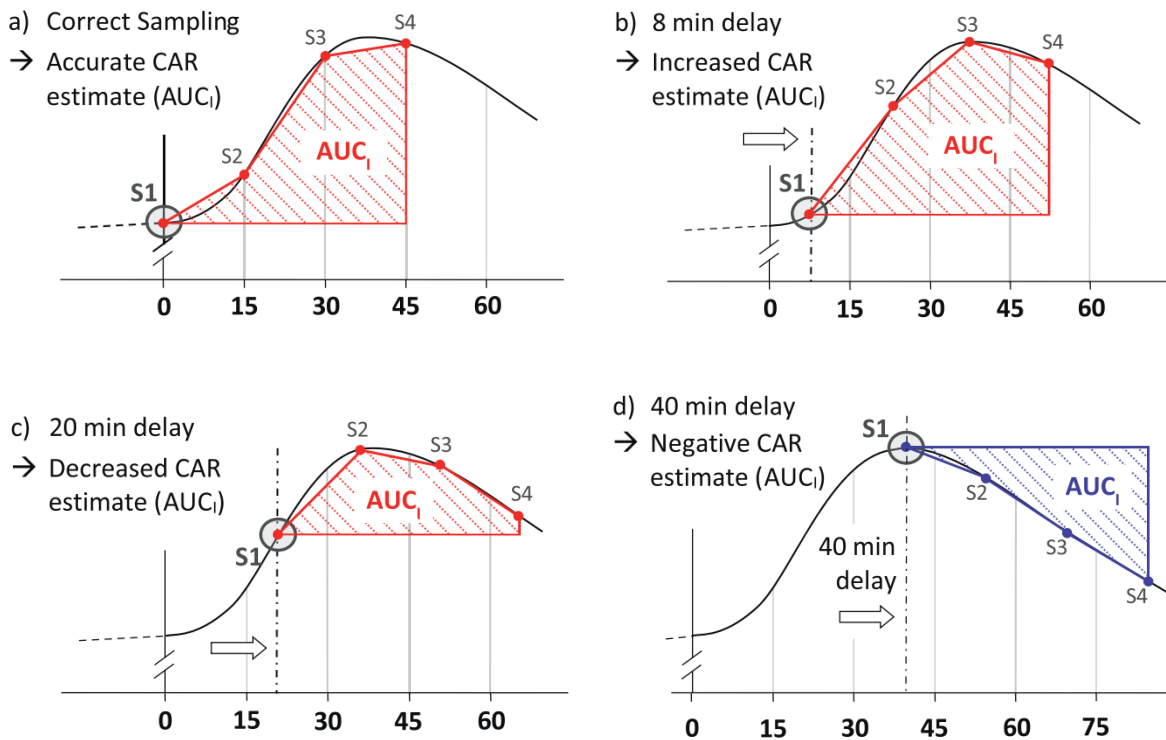


Figure 2.2: Illustration of the impact of delayed sampling on the computed  $AUC_1$ . The Figure depicts estimates given (a) correct sampling and for delays of (b) 8 min, (c) 20 min, and (d) 40 min between awakening and collecting S1 [Sta16].

Griefahn and Robens conducted 3 studies with 108 participants in total who were instructed to take samples on 6 to 8 days [Gri11]. They found that on 19.3 % of days there were delays in the first sample of 3 to 30 min and on 14.0 % of days there were delays of more than 30 min, which shows that delays occur frequently and can have a large impact on CAR measurements.

#### *Inaccurate post-awakening sampling times*

In addition to monitoring awakening time, it is also crucial to objectively track the post-awakening sampling times. Although objective control methods for sampling times are used more often than objective control methods for awakening times, most CAR studies still rely on self-reported sampling times without further verification [Sta16; Sta22].

Kudielka et al. conducted a study in which the participants had to collect 6 saliva samples throughout one day in intervals after awakening and at fixed times during the day [Kud03b]. An electronic monitoring device was used to measure the sampling times and compare them to the self-reported sampling times. 24 of the 47 participants were not informed about the additional

monitoring. 11 participants failed to obtain the sample at the correct time at least once with the informed group adhering to the sampling schedule more closely. The results also showed that noncompliance had a larger impact on the cortisol levels measured after awakening than those measured throughout the rest of the day.

Similar results were obtained by Broderick et al., who analyzed objective compliance and self-report compliance of 66 participants that had to take 5 samples over 7 days, respectively [Bro04]. The samples were scheduled immediately upon awakening, 45 min after awakening, and 16:00, 19:00, and 22:00. Non-compliance was defined as missing the defined sampling time by more than 15 min for the morning samples and 60 min for the afternoon samples. While self-reported compliance was 93 %, actual objective compliance was considerably lower at 71 % showing that participants tended to misjudge their adherence to the sampling schedule. Non-compliance resulted in a flatter cortisol slope showing that even liberal compliance windows of 15 min have an impact on the results.

The impact of non-compliance on the cortisol rise after awakening was investigated by Kudielka et al. [Kud07]. They showed that noncompliant participants had a significantly lower cortisol rise after awakening than compliant participants and 60 % of the participants took a sample outside of the specified 10-minute time window on at least one of the three sampling days. An even larger compliance issue was observed in a study conducted by Smith and Dougherty, who analyzed the sampling behavior of preschool-age children and their parents over two days [Smi14]. The sampling inaccuracy in this study was even higher with 44.3 % of the CAR samples being taken outside of the specified 10-minute time window, which resulted in a higher waking cortisol and lower CAR. Golden et al. conducted a larger-scale study with  $N = 935$  participants over 3 days with 6 samples per day [Gol14]. The first sample had to be taken directly after awakening, while a second one was scheduled 30 min later. The sampling delay of the first sample was larger than 15 min for 21.0 % of the samples and even 31.2 % of the second samples.

### *Confounding factors*

Due to the potential for errors, methodological standards for CAR assessments were published in 2016 and updated in 2022 [Sta16; Sta22]. Alongside issues resulting from inaccurate sample and awakening time reporting and methods to deal with these issues, several covariates that can influence the CAR directly were listed. According to the authors, these factors should be monitored and considered as potential confounds throughout the assessment. One of those parameters is the light level a person is exposed to in the morning with higher light levels being related to an elevated CAR and vice versa [Fig12; Tho04]. Researchers should therefore monitor the ambient light level by either obtaining self-report data from the participants (e.g., request information



about use of eye masks and lighting of bedroom) or measuring it (e.g., through small photosensor devices [Fig12]) [Sta16]. An increased CAR was also shown on weekdays compared to non-weekdays [Kun04; Sch04]. This might be related to the participants' anticipation of the day ahead and the perspective memory load [Bäu13; Gri11; Sta10; Mar15]. Alongside these covariates, Stalder et al. proposed several verification methods for awakening and sampling times that are introduced in the following Section [Sta16; Sta22].

## 2.2 Objective Verification Methods

### *Objective awakening time verification*

The gold standard to assess sleep parameters is PSG. Hereby, multiple body functions such as brain activity, eye movements, muscle activity, and heart rhythm are monitored and aggregated to detect sleep stages [VAN11]. Although PSG has already been used in different CAR studies [Oku10; Gri10; Gri11], it has some substantial disadvantages especially for ambulatory, large-scale studies, as it is costly and labor intensive. Furthermore, PSG is disruptive to participants' normal routine which can have an influence on the CAR.

A cheaper and less disruptive method is wrist actigraphy. This method utilizes wrist-worn devices with an in-built accelerometer with a high sensitivity to detect and monitor wrist movement [Lic06]. Scoring algorithms that use predefined activity count thresholds are employed to infer sleep-wake measures, which can be used to determine different sleep parameters such as latency to sleep, wake time after sleep onset, total sleep time, and awakening time [Sad02; Lic06].

In a systematic review, Haghayegh et al. suggested that recent-generation Fitbit devices that utilize heart rate as well as body movements do comparatively well in estimating the total sleep time [Hag19]. Similar results were achieved by Chinoy et al., who compared gold standard sleep measurement techniques and commercial wearable sleep tracking devices under naturalistic unrestricted home sleep conditions [Chi22]. The study included 21 participants whose sleep was monitored over one week. Four commercial sleep tracking devices and one actigraphy watch were tested and compared to a mobile sleep EEG headband device. The results indicated that the commercial sleep tracking devices performed equally well or better than the actigraphy device.

Another approach for determining the time of awakening is the use of a body movement monitor worn on the chest, which also records heart interbeat interval (IBI) data. Since waking from sleep is associated with an increase in heart rate [Hui94; Tri01; Tri03], available IBI and actigraphy data could help to determine the time of waking more accurately. Chest-worn motility monitors were already used in CAR research [Kup05; Sta11].

Apart from methodological strategies, awakening time accuracy can also be achieved through study design by forced awakening. Hereby, participants are externally woken by either other persons (e.g., study personnel or parents for children) or alarm clocks [Sta16]. Since the CAR is not affected by whether the participant woke up spontaneously or was externally woken [Wüs00; Sta09] there are no significant changes in the CAR profile to be expected. However, when using this strategy, researchers have to consider that the participants might wake up prior to the scheduled awakening time and therefore must provide an additional option to assess the awakening time [Sta16].

#### *Objective sampling time verification*

A second potential source of error in CAR surveys is inaccurate sampling times. Different strategies have been used to verify self-reported sampling times.

According to the Expert consensus guidelines published by Stalder et al., the current best practice is using screw top bottles that record the times when they have been opened (e.g., MEMS®Track Cap; AARDEX, Ltd., Zug, Switzerland<sup>1</sup>) or alternatively boxes with the same feature [Sta16]. The participants get one bottle/box per scheduled sample with the sampling devices being stored inside the bottle/box. They are also instructed to open the bottles/boxes just before they have to take the sample so that recorded timestamps are as close as possible to the actual sampling time. Although the method was already used in CAR assessments as well as in compliance/adherence studies [Kud07; Gol14], it is not protected against intentional misuse (e.g., if participants open the bottle and take the respective sample later) and becomes very expensive for large-scale studies as one bottle/box per sample is needed [Sta16]. In an update of the guidelines from 2022, the authors suggested the use of smartphones as an objective sampling time verification tool [Sta22]. Smartphones can be of great benefit for CAR assessments as most people have access to them, which means that no additional, cost-intensive hardware is required, and they provide multiple in-built functionalities such as camera and displaying notifications [Sta16; Sta22].

One method to confirm sampling times, that has already proven useful, was to set up a phone so that it gives an acoustic alarm in the morning until the participant engages with the phone. Then, a code is shown on the display, which has to be recorded on the label of the sampling device. The process is repeated each time a sample is due [Pow12; Pow15; Bed19]. This way, the participants are more likely to take the samples at the desired times and the sampling device can later be matched to the respective sampling time.

---

<sup>1</sup><https://aardexgroup.com/smart-pill-bottle/>

A similar method is to send text messages to the participants which they must reply to within a short time window [Rod18]. Here, the replies can be used as timestamp as well as a compliance marker. Hence, samples for which the reply is missing can be considered non-compliant and sorted out.

However, both methods described above are only advisable if the participants initiate the sampling procedure by themselves. If the reminders for the first samples are sent at pre-set, estimated awakening times and the participant woke up prior to the reminder, the procedure would be invalid [Sta22]. Zhu et al. utilized the camera functionality of smartphones by instructing the participants to take a photo upon awakening and each time they took a saliva sample. An automatically generated time stamp of the photo was used as sampling time assessment and adherence confirmation [Zhu19].

Haase et al. even provided a specifically designed tracking app for their study. The app guided the participants and displayed a 60-second timer for each sample. The participants were instructed to take a photo of the respective saliva sample and press a button once the timer expired. When the button was pressed, a timestamp was stored which could later be compared photo's time stamp to verify compliance [Haa24].

A similar low-cost solution that enables objective verification is CARWatch. CARWatch is an app that reminds the participants to take samples using the phones alarm functionality. To confirm that the sample was taken, the participant has to scan a barcode that is attached to the sampling device. A detailed description of the CARWatch framework can be found in Chapter 3. The app was evaluated in a study with a total of 288 participants of which 65 were instructed to use CARWatch. The results showed that using the app increased adherence to the sampling schedule and led to a more consistent sampling behavior [Ric23].

This thesis focuses on improving and extending the CARWatch framework to provide an open-source, modular, and flexible tool for improving CAR assessments by helping to objectively verify awakening and sampling times. Within this thesis, additional features that can further improve the assessments and prevent eventual confounding factors were included into the CARWatch framework. Additionally, a potential influence of the factors non-workday/workday (see Section 2.1) on the sampling behavior will further be investigated.



# Chapter 3

## Methods

The CARWatch framework provides a low-cost objective verification tool for sampling times that can easily be integrated into field studies. It can be used for any diurnal biomarker collection at home but was mainly designed to support the assessment of cortisol sampled during the CAR. Since the source code is publicly available<sup>1</sup>, researchers are also able to customize the framework for their own purposes.

This Chapter describes the functionalities and architecture of the CARWatch framework. The first part gives a brief overview of the core features of each component alongside a description of the component's usage. Afterwards, the user interface and features of the webapp are further explained and the last part describes the user interface of the Android app and functionalities.

### 3.1 Framework overview

The CARWatch framework consists of three applications: a smartphone app, a webapp, and a Python package. The individual applications are designed to be used together and fulfill different purposes (see Figure 3.1).

The webapp serves as a configuration and preparation tool for setting up and configuring studies. Information about the study is encoded into QR codes which are then parsed by the smartphone app to configure the app accordingly. The webapp was developed using the web framework svelte<sup>2</sup> and consists of a study configurator and an analysis tool. The study configurator contains multiple Hypertext Markup Language (HTML) forms in which the researchers can enter different study parameters such as the study name, the number of participants, the number of sampling days and

---

<sup>1</sup><https://github.com/mad-lab-fau/carwatch>

<sup>2</sup><https://svelte.dev/>

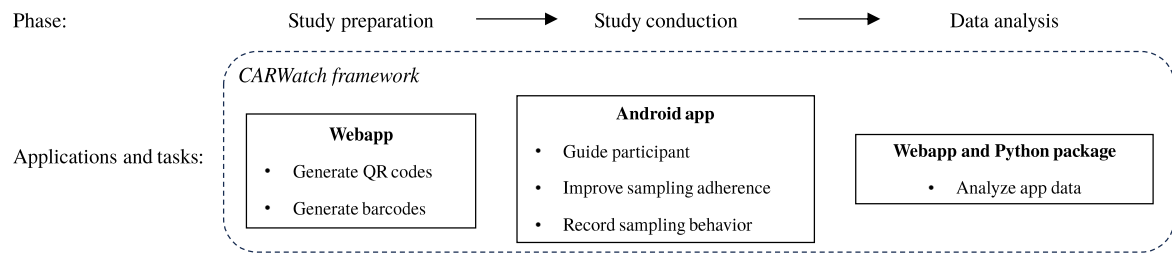


Figure 3.1: CARWatch framework applications during CAR assessment study.

the number of samples per day, the intended sampling times, etc. After all required information was added, the website generates individual QR codes for each study participant and barcodes for each saliva sample that can be printed out and attached to the sampling devices before they are handed out to the participants.

The CARWatch app is designed as an objective verification tool for sampling times and to increase the participants' adherence to the sampling schedule. It can be installed on Android-based smartphones through the Google Play Store. Android was chosen as platform since it is the most widely spread mobile operating system [Sta19] and because of its open-source characteristics [Ric23]. An iOS version of CARWatch is - while this thesis was written - under development. For CAR assessments, each participant needs to use one smartphone on which the CARWatch app is installed. For the initial setup of the app, the user needs to scan the QR code generated by the webapp so that the study configuration is loaded. The app includes an alarm feature that can be used to set a wake-up alarm for the study days. If no alarm is set, the participants need to self-report their awakening to initiate sampling for that day. After awakening was reported or the morning alarm was stopped, sample alarms that serve as sample collection reminders are scheduled. Taking a sample must be verified by using the app to scan the barcode attached to the sampling device. This procedure is repeated on each study day. Information about when the participant woke up, took the samples, and which barcodes were scanned are stored inside automatically generated log files. After the study was finished for the individual participant, the log files can be sent to the researchers to be further analyzed.

Both the webapp or the Python package can be used to summarize multiple log files into a single tabular file that contains the sampling information for all participants. The Python package provides extended analysis options like merging the sampling time with information about the cortisol measurements. Furthermore, it can also be used to generate the QR codes and barcodes. As no major adjustments were made to the Python package in the course of this thesis, it will not be considered further here.

As part of this work, both the Android app and the webapp were enhanced with various new features. The features added to the Android app include sleep data tracking, an introductory slideshow, absolute alarm times, a light intensity logger, and a reworked alarm user interface. The webapp was extended with several study configuration parameters to enable the configuration of the new Android app features. Additionally, the generation of the study material was made more customizable.

## 3.2 Webapp

The webapp consists of two components: the study configurator and the postprocessing tool. The latter can be used to analyze the logs generated by the Android app. For this, all log files need to be compressed into a ZIP file and uploaded to the postprocessing tool. In order for the postprocessing to work properly, the log file names should not be changed as the filenames are parsed to extract the participant identifiers (IDs), which is required for subsequent matching with other information. The webapp then summarizes the log file data into a single CSV file which contains one row per participant available in the dataset. Each row includes information about awakening time and type of awakening (woken up by alarm or self-reported awakening) and the samples taken by the participant (sampling times and scanned barcode) for every study day. In this way, researchers can get an overview without having to extract the information manually from the log files.

The second component, the study configurator, consists of three HTML forms. In the first form (Figure 3.2) general study parameters such as study name, number of sampling days, number of participants, and participant IDs can be entered. Furthermore, the number of samples per day can be specified and information on whether an additional sample should be collected just before going to bed (also called *evening sample*) can be provided.

The second form (Figure 3.3) can be used to customize the layout of the barcode sheets that are generated. This is especially useful if the barcodes should be printed onto pre-cut print label sheets with specific dimensions. The user can choose between the two different formats A4 (international standard) and ANSI letter (standard paper format in North America), set the number of barcodes that are printed onto one page, and customize the distances between the barcodes. For each paper format, a preset that is already adjusted to a specific label template is available. If the study does not require barcodes the user can also choose to print only the IDs of each sample or just skip this step.

In the final step (Figure 3.4), additional study parameters for the app configuration like an email address to which the final log files will be sent, can be entered. Researchers can also decide if the

**Step 1**

### Study Details

Study name  
Test

Number of sampling days: 4  
Number of biomarker samples per day: 4  
Prefix for biomarker IDs (e.g. 'S' for saliva): S

Read participant IDs from file

Number of participants: 50  
Participant ID prefix (displayed on labels): P\_

Add additional evening sample (SE)  
 Start sample counter from zero (S0 instead of S1)

Study Type  
CAR Study

← Back      Next →

Figure 3.2: Study configuration step 1.

**Step 2**

### Barcode Details

Generate labels for study

Add study name to labels  
 Add barcodes to labels

Print label layout:  
 Use ANSI letter format (11 in x 8.5 in) instead of A4 (297 mm x 210 mm)

Layout preset  
Avery Zweckform J4791-25 (48 labels per ...)

Number of columns: 4  
Number of rows: 12  
Distance between columns: 2,5 mm  
Distance between rows: 0 mm

Left margin: 9,8 mm  
Right margin: 9,8 mm  
Top margin: 21,2 mm  
Bottom margin: 21,2 mm

← Back      Next →

Figure 3.3: Study configuration step 2.

participant ID should be included in the generated QR code, whether scanning the same barcode twice is prohibited, and if additional sleep data should be collected from Google Fit throughout the study. In this form, the researchers must also define the sampling schedule. To do this, they can specify how many samples should be taken at a fixed time of the day. The remaining samples are scheduled based on the participant's awakening time. Depending on whether the sampling time is



Figure 3.4: Study configuration step 3.

fixed, the researchers must either select a time or a time distance in minutes, which determines the interval between the respective sample and the previous sample, except for the first sample, which is scheduled based on the participant’s awakening. In the last part of the form, the print layout can be customized similar to the barcode print layout customization by choosing the page format and the number of rows and columns that are printed onto a single page. By clicking “Complete”, the study material is generated and can be opened via the respective button.

An individual barcode is generated for each sample in the study, encoding an eight-digit combination (Figure 3.5 shows the generated barcode sheet and Figure 3.6 a single barcode). The first three digits identify the participant to whom the barcode belongs. The next two digits specify the study day on which the sample should be taken, followed by two digits indicating the specific

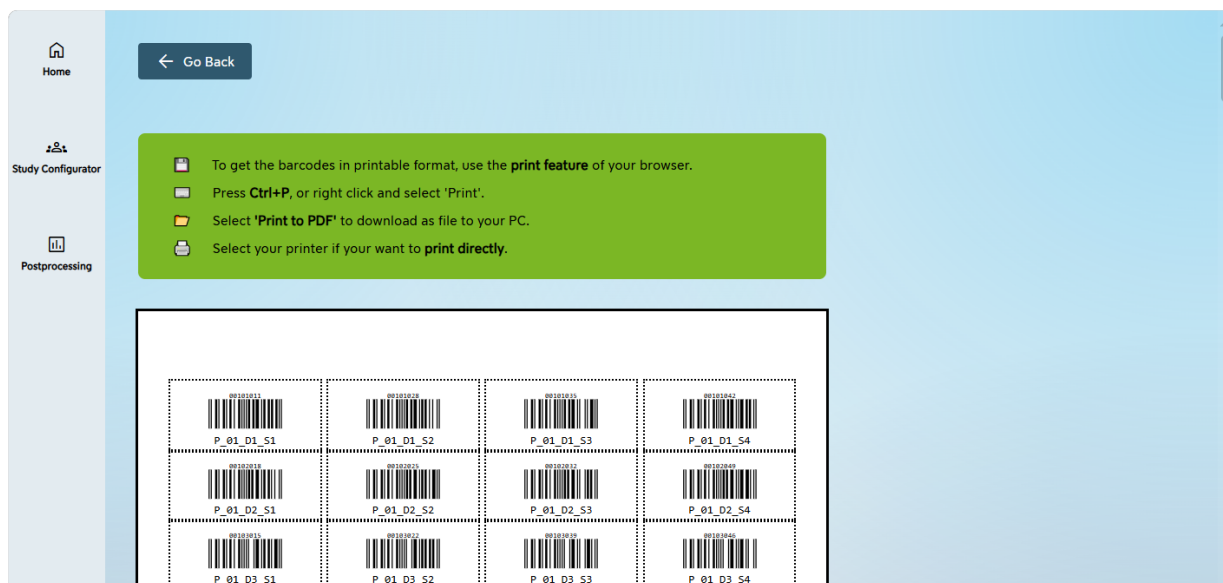


Figure 3.5: Barcodes generated by the webapp.



Figure 3.6: Generated barcode. The digits above the code are the barcode value. The label below the barcode provides information about the participant and the sample to which the barcode belongs.

sample of the day. The final digit serves as a control number according to the EAN-8 barcode standard<sup>3</sup>. Following this encoding scheme, a barcode with a value of “00903028” 3.6 would identify the second sample of participant “009”, collected on the third sample day.

The QR codes contain the study parameters that were entered in the form of the webapp. For each participant, one QR code is generated. Depending on whether the participant ID is included, each QR code is equal or belongs to a specific participant whose ID is displayed below the QR code in the generated print preview. The QR code encodes a string that consists of multiple *key:value* combinations that are separated by semicolons. The decoded string of the QR code depicted in Figure 3.7 contains the following information:

”CARWATCH;N:Test;D:5;NP:50;SS:S0;T:0,30,15;A:1500;E:0;M:foo@bar.de;  
FD:1;GF:1;V:1.1.0;PID:P\_03”

<sup>3</sup><https://www.iso.org/standard/46143.html>



P\_03

Figure 3.7: Generated QR code containing the study configuration. The participant ID that is encoded in the QR code is displayed below the QR code.

- “CARWATCH” → serves as control variable to ensure the validity of the QR code
- “N:Test” → study name is “Test”
- “D:5” → samples should be taken over 5 days
- “NP:50” → 50 study participants
- “SS:S1” → the first sample ID is “S1”
- “T:0,30,15” → time distances between samples (samples S1, S2, and S3 should be taken 0 / 30 / 45 min after awakening)
- “A:1500” → absolute times: one sample must be taken at 15:00
- “E:0” → no evening sample is required
- “M:foo@bar.de” → email address is “foo@bar.de”
- “FD:1” → scanning the same barcode twice is not allowed
- “GF:1” → enable the retrieval of Google Fit sleep data
- “V:1.1.0” → webapp version the QR code was generated with
- “PID:P\_09” → participant ID is “P\_09”

The QR codes and barcodes can be printed by using the print function that is supported by most modern web browsers.

## 3.3 App

The Android app offers functions to improve the sampling behavior of the participants, to monitor the sampling, and to record relevant information. It was developed in Java using the Android Java API framework<sup>4</sup>. The first Section of this Chapter describes the app's user interface and the functionalities of the different screens. In the second Section, the core features of the app are listed and explained in more detail.

### 3.3.1 User Interface

The main app screen consists of a header, a content area, and a bottom navigation bar (see Figure 3.8). The header contains the current screen name and a button that opens the menu (Figure 3.9). The menu entry “Share logs” can be used to send the generated log files to the researchers after the participant has completed the study. The “Delete logs” option deletes all log and data files generated by the app and was added so that already exported files can be deleted if the smartphone is used for several participants in succession. The entry must be clicked multiple times, so that the log files are not deleted by mistake since it cannot be undone afterwards. With the “Kill Alarms” option, the user can deactivate all active alarms. The “Reregister” option resets the stored study configuration and opens a QR code scanner so that the user can load a new configuration. Clicking “Show Tutorial”, restarts the tutorial slides of the slide show (see Section 3.3.2). The last menu entry “Info” shows information about the current app version.

To navigate between the different screens, the user must click the respective entry of the bottom navigation bar. The wake-up screen showed in Figure 3.8 lets the user report their awakening in case no wake-up alarm was set, or the participant woke up prior to the alarm. Clicking the “Yes” button initiates the morning procedure provided it has not been initiated earlier. If the first sample must be taken immediately after awakening (sampling time distance in study configuration is 0), the sampling procedure is started. Additionally, the remaining sample alarms for the day are set in the background.

The alarm screen (Figure 3.10) provides an overview of the current alarms. The wake-up alarm for the next morning can be set by clicking the top time. After the morning procedure was started, the sample alarms are displayed below the wake-up alarm. Each alarm can be activated and deactivated by clicking the switch button in front of the alarm time. For the sample alarms, the label of the sample and the status of the sample are shown next to the alarm. A green checkmark symbol indicates that the sample was already taken and verified by scanning the respective barcode.

---

<sup>4</sup><https://developer.android.com/>

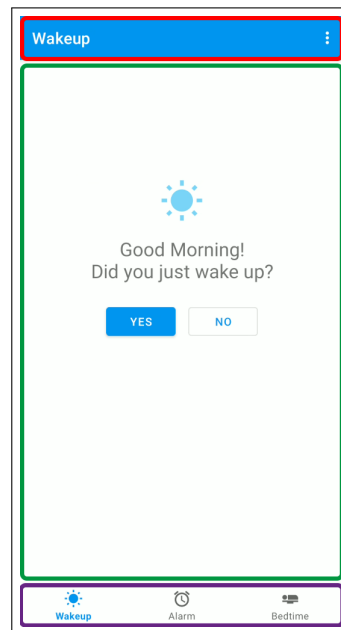


Figure 3.8: Main screen of the app consisting of the header (highlighted red), the content area that holds the content screens (highlighted green), and the bottom navigation bar (highlighted purple). The content screen displayed is the wakeup screen.

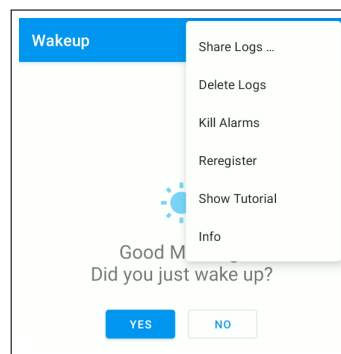


Figure 3.9: Unfolded app menu.

The orange clock symbol shows that should have been taken already but the respective barcode was not scanned yet. Additionally, a barcode symbol is displayed if a sample still needs to be taken, regardless of when the sampling time is. Clicking the barcode symbol opens the barcode scanner for the respective sample.

The bedtime screen (Figure 3.11) can be used to activate the bedtime mode by clicking the “Yes” button which starts the light intensity recording (see Section 3.3.2). If the study includes an evening sample, the respective sampling procedure is started. Additionally, the screen contains a “Lights out” button that enables the dark mode.

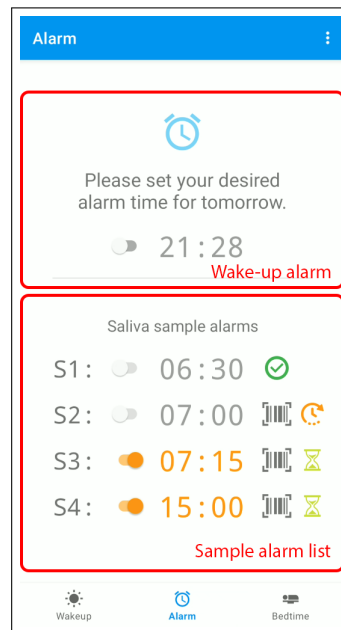


Figure 3.10: Alarm screen with wake-up alarm and sample alarms. Active alarms are highlighted in orange. The symbols behind the alarm times indicate whether the sample was taken (checkmark), should have been already taken (orange clock), or is scheduled for later (sand clock). The barcode symbol can be used to open the barcode scanner.

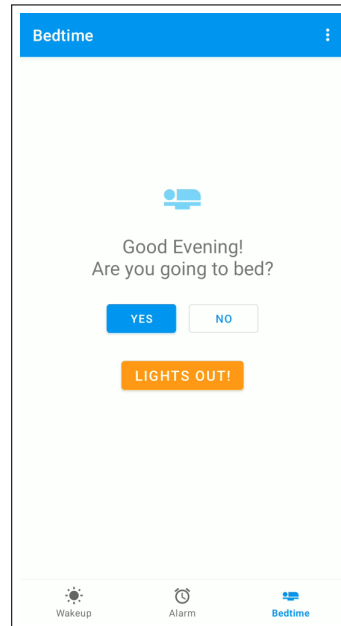


Figure 3.11: Bedtime screen that can be used to initiate the evening procedure by clicking the “Yes” button.

### 3.3.2 Features

The CARWatch app features can be grouped into different categories. The wake-up alarms, sample alarms, and the barcode scanner actively support the participant and serve as objective verification for awakening and sampling times. The sleep data reader and the light intensity passively collect information about the participant's sleep and environment factors, and the slide show was designed to help the user set up the app and provide an overview of how to use it. This Section describes the features listed above, their functions and their architecture in more detail. The app also contains other features that are not listed here as they were not considerably changed in the course of this work.

#### *Alarm*

Alarms serve different purposes. The wake-up alarm follows the forced awakening strategy (see Section 2.2) by waking the participant externally and thereby objectively determining the awakening time, which is logged in the background. The sample alarms serve as reminders and help the participants to adhere to the sampling schedule. Based on the study configuration the wake-up alarm can also be a sample alarm given that the first sample must be taken immediately after awakening. Unlike the awakening alarm, the sample alarm cannot be rescheduled. When an alarm goes off, the user is alerted by an acoustical alarm signal and a visual notification which can be clicked to stop the alarm (see Figure 3.12). Upon stopping a sample alarm, the barcode scanner, shown in Figure 3.13, is automatically opened, and the participant has one minute to scan the barcode attached to the sampling device. The sampling procedure is finished if the scanned barcode is valid. If the participant does not manage to take the sample within the time window, an additional reminder alarm is triggered. After the second alarm was stopped, the barcode scanner can only be opened by going to the alarm screen and clicking the barcode symbol (see Section 3.3.1).

#### *Welcome slide show*

The welcome slide show is automatically shown when the app is opened for the first time after installation and serves as initial guide for the app. It contains multiple slides that are shown in sequence and must be run through before the app can be used (see Figure 3.14). The slides are embedded into a single screen which consists of the slide area, and a bottom bar (see Figure 3.15). The dots in the center of the bottom bar indicate which slide is currently being displayed. The user can advance the slides by either pressing the "Next" button or swiping left. Once the slideshow is finished, it can always be restarted via the top right menu (see Section 3.3.1). In the first two slides

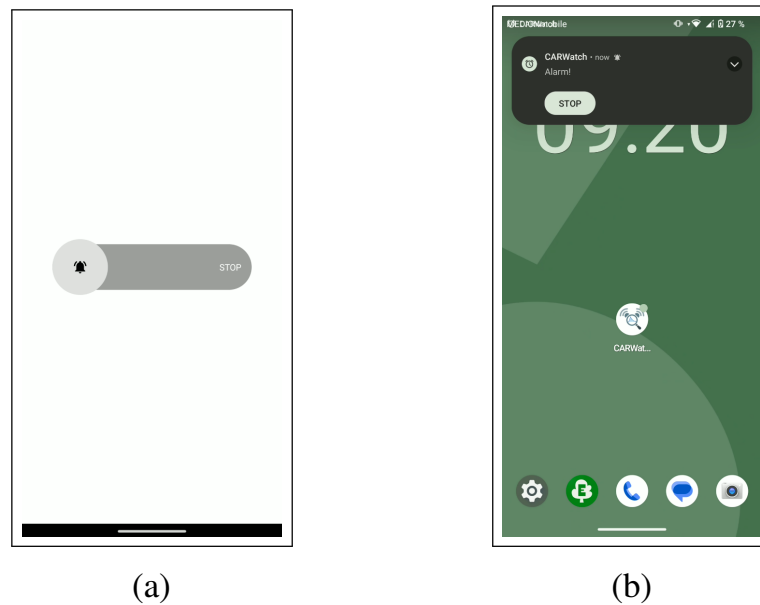


Figure 3.12: Alarm notification if display is locked (a) and if display is unlocked (b).

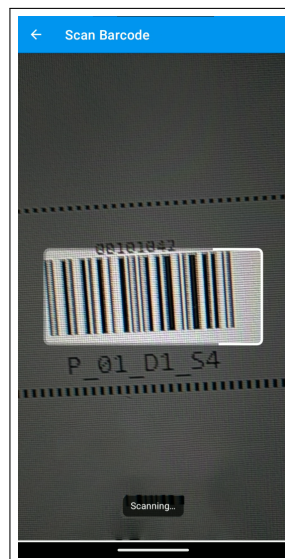


Figure 3.13: Scanner for sample device barcodes. The highlighted area must be pointed at the barcode.

of the slideshow, the user is asked to give the required app permissions. If the permissions are granted, the camera is opened and must be pointed at the study configuration QR code generated by the webapp. After the study configuration was loaded, the user can advance to the next slide on which they must enter their participant ID. This slide is skipped if the QR code already contained the ID. If the researchers activated the Google Fit sleep data collection, the next slide contains



a dialog in which the user is asked to log into their Google account and allow the app to access their sleep data. The last few slides (referred to as *tutorial slides*) provide an overview of the app structure and core functionalities. Each tutorial slide consists of a screenshot and text that explains the respective screen and its elements (see Figure 3.15b). The tutorial slides can be skipped by using the button in the bottom left corner and the user is also able to go to the previous slide by swiping right. Finishing the slide show automatically redirects the user to the main screen.

The slide show is implemented as a single Android *Activity*. Figure 3.16 shows the relations of the classes that belong to the slide show. The *SlideShowActivity* class creates the slide objects in the private method *initializeSlides()*, holds the individual slides in the *slides:List<WelcomeSlide>* attribute, and operates the slide navigation with *nextSlide()* and *previousSlide()*. A *FragmentContainerView* was used as layout element to embed the respective slide UI in the *Activity*. Based on the parameters that are passed when starting the *SlideShowActivity*, it initially creates all slides, only the setup slides, or only the tutorial slides. Most slides are objects of different classes which differ in terms of layout and functionality embedded in the slide. To ensure that a slide class implements all functions that are required by the *SlideShowActivity*, each slide class must implement the *WelcomeSlide* interface. The abstract class *BaseWelcomeSlide* provides a default implementation of all functions of the *WelcomeSlide* interface and was added to reduce redundancies. The *BaseWelcomeSlide* class also extends the *Fragment* class since the *FragmentContainerView* of the *SlideShowActivity* can only display *Fragment* objects. In this way, each object of a class that extends the *BaseWelcomeSlide* class is also a *Fragment* object which is why the *BaseWelcomeSlide#getFragment():Fragment* method just returns the object itself. In total, seven non-abstract slide classes were implemented, each of which representing one or more slides in the slide show:

- *WelcomeText* - first slide, shows greeting text
- *PermissionRequest* - in this slide, the user is asked for the app permissions
- *QrFragment* - contains the QR code scanner
- *ParticipantIdQuery* - provides an input field for the participant ID
- *GoogleFitAuthentication* - lets the user log into their Google Fit account
- *EndTutorialSlide* - last slide, shows text
- *TutorialSlide* - class that can be used to create tutorial slides

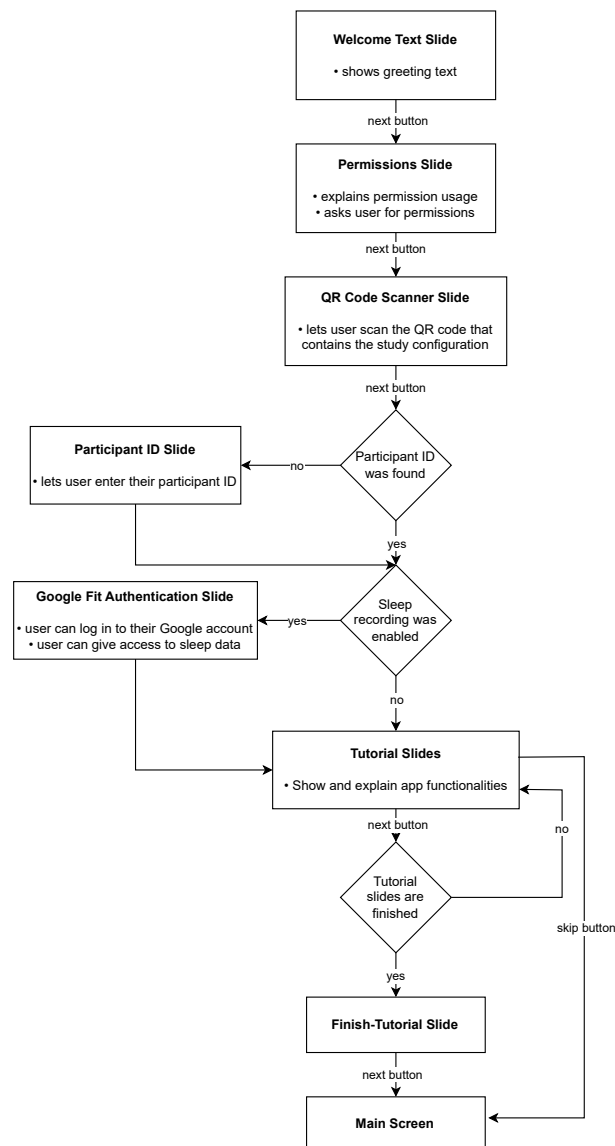


Figure 3.14: Slide sequence of slide show after app installation. If the slide show is started via the “Reregister” menu entry, it starts with the QR code scanner slide and finished before the tutorial slides would be shown. If the tutorial option in the menu is clicked, the slide show starts and finishes with the tutorial slides.

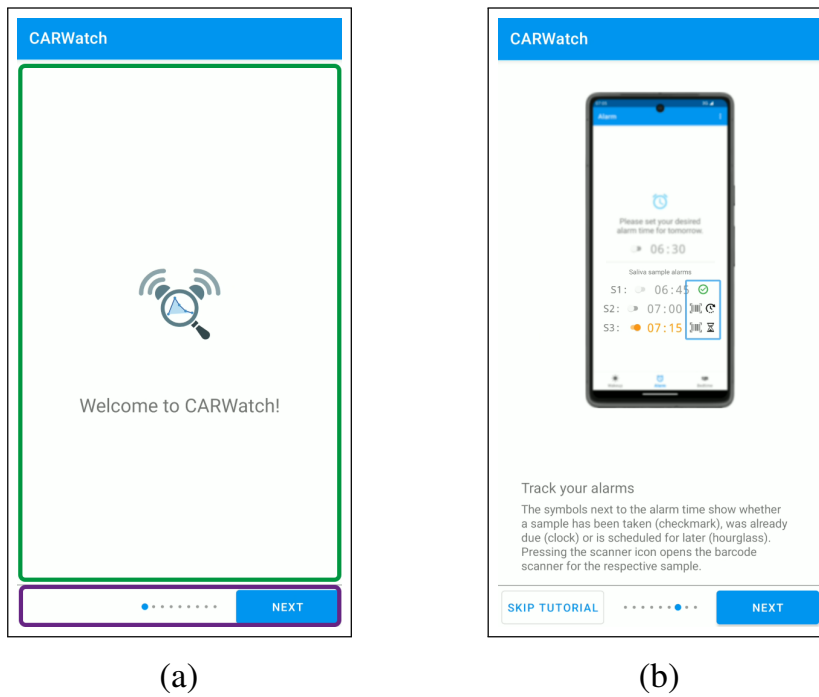


Figure 3.15: (a) First slide that is shown when the app is opened for the first time. The green area highlights the slide content, the purple area highlights the slide navigation bar. (b) Tutorial slide that explains the alarm screen. Each tutorial slide has the same structure and contains a screenshot of the feature that is explained as well as an explanatory text. The tutorial slides can be skipped with the button in the lower left corner.

All slide objects that belong to the app tutorial are of the *TutorialSlide* class as they all have the same layout consisting of an image, a headline, and a description. These three parameters must be specified when creating a *TutorialSlide* object.

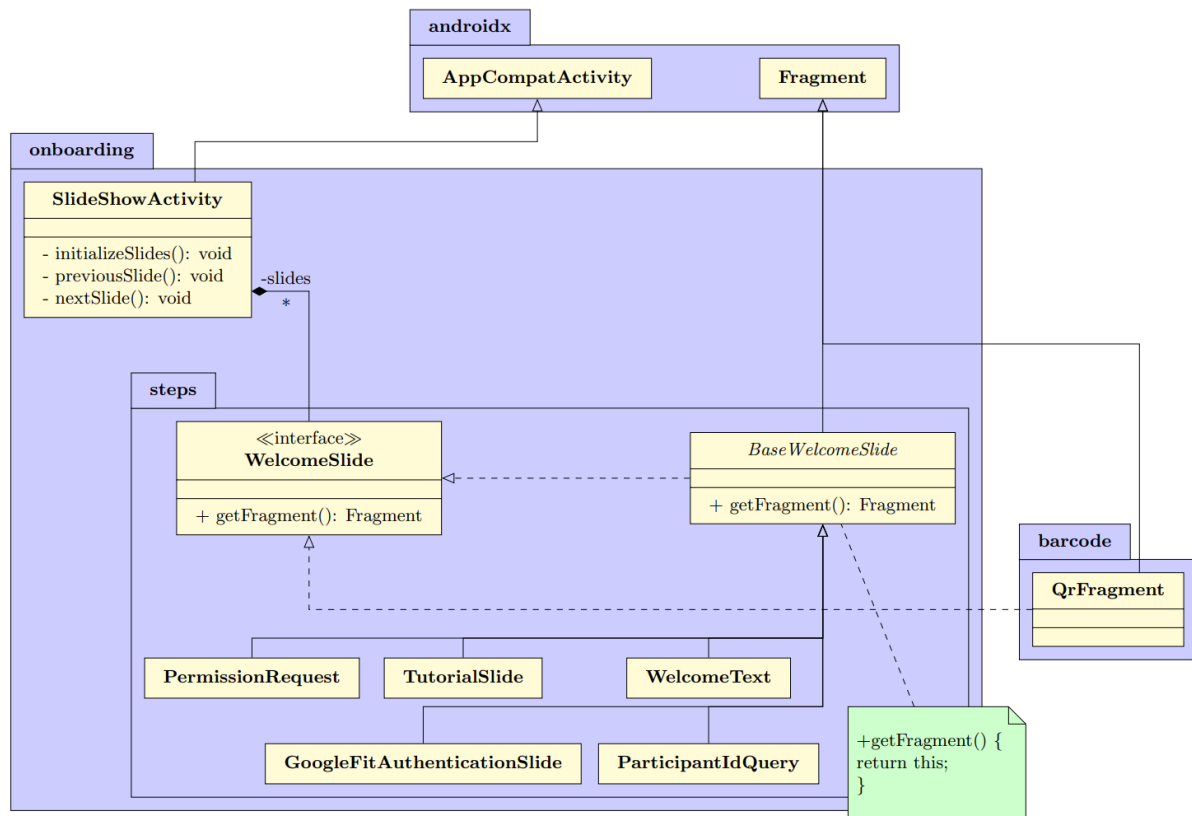


Figure 3.16: Slide show classes and their relations. The diagram does not contain all class attributes and methods, only those mentioned in the slide show Section.

### *Sleep data reader*

The sleep data reader was added as an additional method for determining the awakening time. It enabled the use of external wearables such as smartwatches or smart rings. The sleep data reader is enabled if the respective option was selected in the study configuration and the user added their Google account during the app setup. Additionally, the Google Fit App<sup>5</sup> must be installed and the user's sleep data must have been transferred to the Google Fit App. To determine the awakening time, the app fetches the user's sleep data from the Google Fit API<sup>6</sup>. When starting the app, a request is sent to the API to retrieve all sleep data from the past 24 hours. Upon successful completion, the API returns one or more sessions that include all sleep phases within the specified interval.

<sup>5</sup>[https://www.google.com/intl/de\\_de/fit/](https://www.google.com/intl/de_de/fit/)

<sup>6</sup><https://developers.google.com/fit>

Each sleep phase has a start and end time and can be of one of the following types, which describes the sleep stage:

- Awake (during sleep) - Waking phases during the sleep session
- Light sleep
- Deep sleep
- REM sleep

The end time of the last sleep phase (last phase of type “Light sleep”, “Deep sleep”, or “REM sleep”) is logged as recorded wake up time in the main log-file each day. Additionally, a consecutive CSV file that contains all sleep phases recorded during the study is maintained. This enables the researchers to further analyze the participant’s sleep behavior. The CSV file has one row per sleep phase, containing its start and end time as UNIX timestamp and written out and the sleep stage the participant is in during the phase. Exemplary sleep data recording with a Samsung Galaxy Watch 4 is shown in Table 3.1. The Table shows the last five sleep phases detected before awakening. The inferred awakening time of this example would be the end time of the second to last sleep phase (05:55) as the last sleep phase is of stage “Awake (during sleep)”. The sleep data file is exported along with the log-files.

Start date & time	End date & time	Sleep phase
29.05.2024 05:44	29.05.2024 05:45	Light sleep
29.05.2024 05:45	29.05.2024 05:46	REM sleep
29.05.2024 05:46	29.05.2024 05:47	Awake (during sleep)
29.05.2024 05:47	29.05.2024 05:55	Light sleep
29.05.2024 05:55	29.05.2024 06:00	Awake (during sleep)

Table 3.1: Exemplary sleep phases before awakening.

#### *Light intensity logger*

The light intensity logger was added to collect more information about the participants’ sleep environment. As described in Section 2.1, the light level does influence the CAR and should therefore be controlled or at least monitored. This monitoring can be achieved using the light intensity logger feature of the CARWatch app, which utilizes the Android sensor framework<sup>7</sup> to access the

<sup>7</sup>[https://developer.android.com/develop/sensors-and-location/sensors/sensors\\_overview](https://developer.android.com/develop/sensors-and-location/sensors/sensors_overview)

device's built-in sensors. The framework allows the app to register so-called *SensorEventListeners* that are automatically informed whenever the respective sensor value changes.

For monitoring the ambient light level, the classes *LightIntensityLogger* and *ProximitySensorListener* were added. The *LightIntensityLogger* class implements the *SensorEventListener* interface, enabling it to receive updates from the device's light sensor. The light sensor measures the ambient light level in lux (lx), and the *LightIntensityLogger* logs this data for further analysis. The data is logged in an additional CSV file stored in the same directory as the regular log files, with the following columns:

- “unix\_time”: Contains the unix timestamp of when the light intensity was measured
- “date\_time”: Contains the measuring date and time in a human-readable format
- “light\_intensity\_in\_lx”: Measured light intensity in lx
- “is\_object\_near”: Contains “1” if the distance between the phone's proximity sensor and the nearest object is below a certain threshold and “0” if not.

The values in the last column are determined using the *ProximitySensorListener* class, which is responsible for monitoring the device's proximity sensor. The proximity sensor measures the distance to an object, usually in centimeters. This data can be used to determine whether the device is being held close to the user's face or whether it is lying flat on a table, which would invalidate the measured light intensity. The *ProximitySensorListener* class uses a threshold to infer whether an object is close to the device. The threshold is determined based on the maximum range of the proximity sensor, but is a maximum of 5 cm. Upon adding a new light intensity entry to the data file, the *LightIntensityLogger* fetches the current “is\_object\_near” value from the *ProximitySensorListener*. The light intensity logging can be activated or deactivated through the webapp for the whole study. To start the actual recording in the app, the participant must click the “Yes” button on the bedtime screen (see Section 3.3.1). The light levels are logged every 5 minutes throughout the night. Logging is stopped after the last sample that was scheduled based on the awakening time was taken or upon awakening if no such sample was scheduled.

### 3.4 Evaluation

The study was evaluated using data from the TriASS, a research project conducted by the Trier University under the direction of Prof. Dr. Gregor Domes. The CARWatch system was utilized to verify sample and awakening times.

### 3.4.1 Study design

Samples were collected at the participants' homes. The study involved 30 participants aged 18 to 34, with an average of  $25.6 \pm 3.5$  comprising 10 male participants, 19 female participants, and 1 diverse participant. The data collection period extended between 08.02.2024 and 18.03.2024. Each participant provided data over 4 days, consisting of 2 non-workdays and 2 workdays, which were not necessarily in sequence.

Researchers provided each participant with a Xiaomi Redmi 9AT smartphone running Android 10 (API level 29). The CARWatch application versions used were 2.3.1\_20240209\_1054-969c953 and 2.3.0\_20240201\_1712-57e3b2b. In addition to the sample collection, participants were required to complete a questionnaire about their general life situation, mental and physical well-being, and behavior in various situations.

Participants were free to wake up on their own or set an alarm clock. Saliva samples were collected four times in the morning: immediately upon awakening and then at 30 min, 45 min, and 60 min minutes post-awakening (Samples S1, S2, S3, and S4). In the following, these samples are referred to as *CAR samples*. Additionally, two more samples were collected at 15:00 (S5) and 20:00 (S6), the *fixed-time samples*. Sleep data recordings were not included as no additional sleep tracking devices were available.

Each day, participants filled out a protocol (see Figure 3.17) recording their sampling times and answering four questions about their sleep and their expectations for the upcoming day. Additionally, the participants were instructed to report their awakening in the app or use the in-built alarm so that the sample alarms were scheduled and the barcodes on the saliva devices could be scanned on sampling.

### 3.4.2 Data analysis

Prior to the analysis, the data was cleaned by filtering the cortisol samples for statistical outliers and values outside the measurable range of  $0.33\text{-}82.80 \text{ nmol L}^{-1}$  as suggested by Stalder et al. [Sta16; Sta22]. *CAR samples* and *fixed-time samples* were cleaned separately. For the *CAR samples*, only sampling days that contained all samples S1 to S4 were included in the analysis.

Three reporting strategies were specified to compare different sampling time assessment approaches: *Self-report*, *App*, and *Naïve*. The reporting strategy *Self-report* consisted of the sampling times retrieved from the protocol. As no additional awakening times were assessed, the first sampling time was used as the awakening time. For the *App* strategy, the timestamp of when the awakening was reported in the *App*, or the awakening alarm timestamp was extracted as the

# Protokollkarte



Probandencode: \_\_\_\_\_

Datum: \_\_\_\_\_

Ruhetag:

Arbeitstag:

Salivette	Uhrzeit	Bemerkung/ Besonderes
Aufwachen		
+30 min		
+45 min		
+60 min		

Sind Sie heute Nacht aufgewacht?

Ja  Nein

Wie gut haben Sie geschlafen?

Sehr schlecht ①②③④⑤ Sehr gut

Wie anstrengend wird der Tag heute?

Gar nicht ①②③④⑤ Sehr

Wie sind Ihre Erwartungen an den Tag?

Sehr schlecht ①②③④⑤ Sehr gut

Salivette	Uhrzeit	Bemerkung/ Besonderes
15.00 Uhr		
20.00 Uhr		

Figure 3.17: Empty protocol card for a single day of sample collection.

awakening time. Sampling times were determined from the log entries of when the respective saliva device barcode was scanned. For this, a Python script was used to scan the log files and match the extracted timestamps to the respective samples. Afterwards, the samples were visually scanned to check if sample IDs and times were matched correctly. For the *Naïve* strategy, the awakening times from the *Self-report* strategy were used. If no awakening time was available in the protocols, the awakening time from the *App* strategy was used. The sampling times were determined based on the intended sampling intervals (awakening time + 0 / 30 / 45 / 60 min, and 15:00 and 20:00 for S1, S2, S3, S4, S5, and S6 respectively). This strategy simulated the naïve assumption of participants taking the samples perfectly according to the schedule without any monitoring.



A repeated measures ANOVA (rmANOVA) was performed to analyze the impact of sampling inaccuracies on the CAR metrics  $AUC_G$  and  $AUC_I$ . Hereby, the reporting strategy was used as a within-subject variable and  $AUC_G$  and  $AUC_I$  as dependent variables. Additionally, pairwise t-tests were performed as post-hoc tests.

A mixed ANOVA was conducted to analyze the CAR on workdays and non-workdays, using reporting strategy as a within-subject variable, the day type (workday or non-workday) as a between-subject variable, and the CAR metrics as dependent variables. The mixed ANOVA was performed on all samples as well as on a subset of samples which included samples of all employed participants. The significance level was set to  $\alpha = 0.05$  and effect sizes were reported as  $\eta_p^2$  and Hedges'  $g$  for ANOVA and pairwise t-tests, respectively. All statistical analyses were performed using the Python packages BioPsyKit [Ric21] and pingouin [Val18].

Regarding the sampling times, only samples of the reporting strategies *App* and *Self-report* were analyzed. To show the difference between the intended sampling time and the actual sampling time, the *sampling delay* (or *delays*) was computed. For the *CAR samples*, the *sampling delay* is the difference between the reported sampling time and the sum of the reported awakening time (of the respective reporting strategy) and the intended sampling interval. For the *fixed-time samples*, the *delay* is the difference between the intended time and the actual sampling time. The following groups were compared regarding their *sampling delays*:

- Reporting strategy *App* and reporting strategy *Self-report*
- Samples taken on workdays and samples taken on non-workdays
- Samples taken by employed participants and samples taken by students



# Chapter 4

## Results

A total of 30 participants took part in the study, with 6 being employed, one being in an apprenticeship, one being in an internship, and 22 being students. The ages of the participants ranged from 18 to 34 years, with a mean age of  $25.6 \pm 3.5$  years. The gender distribution was 19 female, 10 male, and 1 diverse participant.

In total, 720 samples were collected, comprising 480 *CAR samples* and 240 *fixed-time samples* over 120 days. *CAR samples* and *fixed-time samples* were cleaned separately.

### 4.1 Data cleaning

For the *CAR samples*, 4 sampling days were excluded due to samples falling outside the measurable range of  $0.33\text{-}82.80 \frac{\text{nmol}}{\text{L}}$  and another 5 sampling days were excluded for containing samples that deviated more than 3 standard deviations from the mean [Sta16; Sta22]. The same criteria were applied to the *fixed-time samples*, resulting in 8 days excluded for measurements out of the measurable range and 2 days excluded as samples deviated more than 3 standard deviations from the mean.

Subsequently, samples without matching sampling times in either the protocol or the log files were removed. *fixed-time samples* scanned outside a range of 180 min before or after the scheduled time were also excluded. Lastly, incomplete CAR measurements, where any of the samples S1 to S4 were missing, were excluded.

After cleaning the data, the final dataset consisted of:

- *CAR samples*:
  - 102 sampling days with 4 samples each for the reporting strategy *Self-report*

- 108 sampling days with 4 samples each for the reporting strategy *Naïve*
- 91 sampling days with 4 samples each for the reporting strategy *App*
- *Fixed-time samples:*
  - 95 sampling days with a total of 181 samples for the reporting strategy *App*
  - 102 sampling days with a total of 199 samples for the reporting strategy *Self-report*
  - 102 sampling days with a total of 200 samples for the reporting strategy *Naïve*

## 4.2 Analysis of CAR features

This Section presents the results of the analysis of the CAR features. Specifically, the features  $AUC_G$  and  $AUC_I$  were analyzed. The first part examines the influence of the reporting type on these features, while the second part explores the influence of the day type on the CAR.

### 4.2.1 CAR throughout the different reporting strategies

The rmANOVA revealed that the reporting strategy had a statistically significant effect on  $AUC_G$ , with  $F(2, 168) = 23.6, p < 0.05$ . The results of the paired t-test indicated a significant difference with a very low effect size between the reporting strategies *App* and *Naïve* (corrected  $p < 0.05$  and Hedges'  $g = 0.05$ ), as well as between the reporting strategies *App* and *Self-report* (corrected  $p < 0.05$  and Hedges'  $g = 0.04$ ). This is also illustrated in Figure 4.1a, which compares the  $AUC_G$  values of the different reporting types. The Figure shows that while the means of the reporting strategies differ only slightly, there are often variations between individual  $AUC_G$  values.

For  $AUC_I$ , the  $F$  value was much lower ( $F(2, 168) = 4.3, p < 0.05$ ) and the results of the paired t-test suggested no significant differences between the individual reporting strategies, as all corrected p-values were greater than 0.05. Unlike in the  $AUC_G$  comparison, there are almost no differences between the individual values, as shown in Figure 4.1b.

### 4.2.2 CAR on workdays compared to CAR on non-workdays

A mixed ANOVA was conducted to compare the CAR on non-weekdays and weekdays. The analysis showed that the day type had no significant effect on either  $AUC_G$  and  $AUC_I$  with  $p > 0.05$  for both respectively.

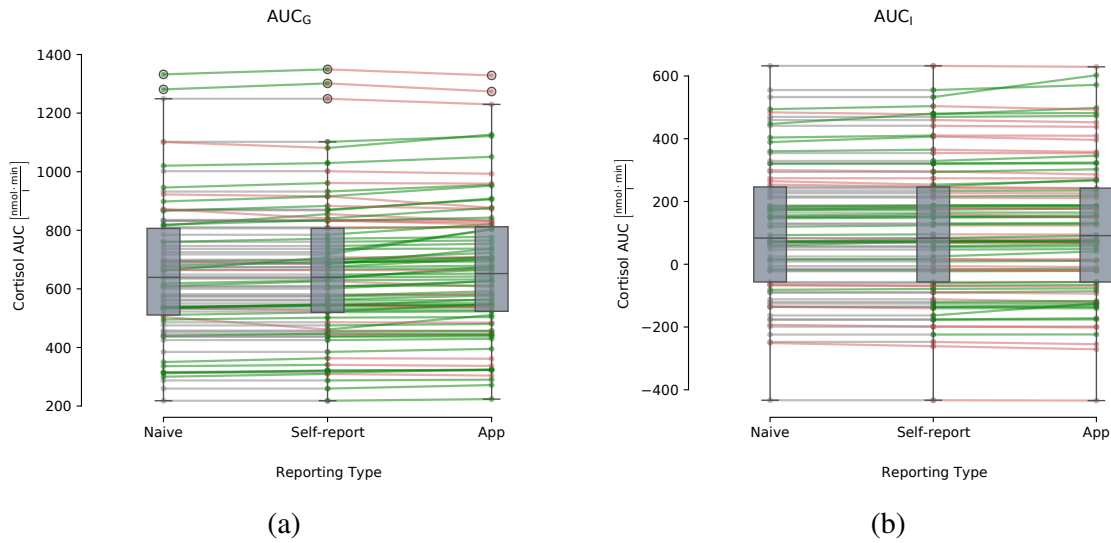


Figure 4.1:  $AUC_G$  (a) and  $AUC_I$  (b) computed from sampling times using different reporting strategies. The green lines between reporting strategies indicate an increase of  $AUC_I$  values, red lines indicate a decrease, gray lines indicate no change.

When focusing on employed participants, the mixed ANOVA still indicated no significant effect on the  $AUC_G$  ( $F(1, 16) = 0.835, p = 0.38$ ). However, for  $AUC_I$ , the results were  $F(1, 16) = 5.85, p = 0.03$ , indicating a significant effect of whether it was a workday or non-workday on  $AUC_I$ , with a partial eta squared  $\eta_p^2 = 0.27$ .

The mean values and standard deviation over all samples of the reporting strategies *App* and *Self-report* for the respective day type are presented in Table 4.1. As depicted in the Table, the mean values for  $AUC_I$  are particularly different between workdays and non-workdays. Additionally, all values show a large standard deviation, indicating that there are substantial differences within the groups of workdays and non-workdays.

## 4.3 Analysis of sampling times

### 4.3.1 Sampling times throughout the different reporting strategies

#### *CAR samples*

The delays between reported sampling times and naïve sampling times for the reporting strategies *App* and *Self-report* are depicted in Figure 4.2. As there was no separate awakening time assessed, the sampling delay of S1 for reporting strategy *Self-report* was assumed to be 0 min for all sampling days. The delays for the subsequent samples remained minimal, as most self-reported sampling

Metric	Workday	Non-workday
$AUC_G$	687.30 (219.80)	565.81 (261.11)
$AUC_I$	256.76 (230.89)	21.50 (134.11)
Sample size (sampling days)	20	18

Table 4.1:  $AUC_G$  and  $AUC_I$  in  $\frac{nmol}{L}$  of employed participants on workdays and non-workdays reported as median (Interquartile range (IQR)). Only sampling days of the reporting strategies *Self-report* and *App* were analyzed.

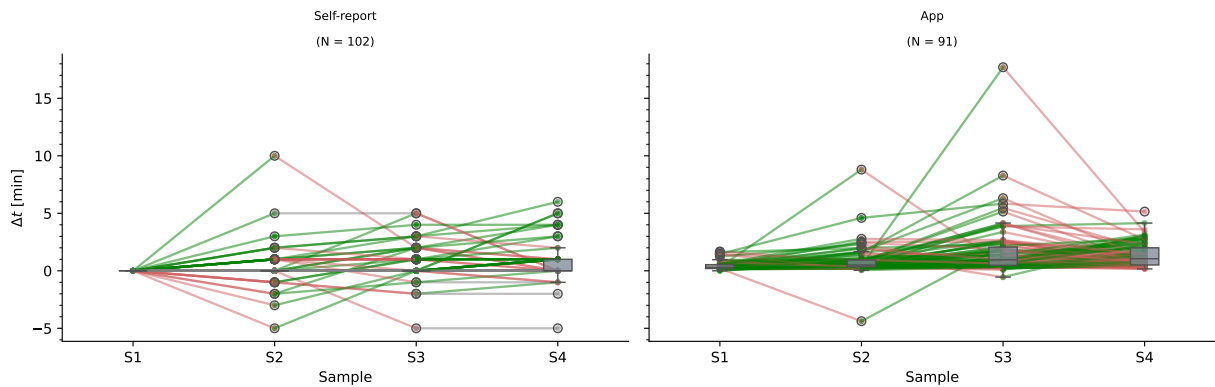


Figure 4.2: Sampling delays of individual participants grouped by the reporting strategies. The Figure includes all *CAR samples*.

times were close to the schedule sampling times. The sample with the largest median delay across all self-reported sampling times was S4 with a delay of 0.66 min. Figure 4.2 shows that while there were some delays for samples S2 and S3, almost none exceeded 5 min.

For the reporting strategy *App*, sampling delays were slightly higher (see Table 4.2 median and IQR). Unlike the delays of reporting strategy *Self-report*, which were nearly negligible, the app-reported delays were mostly positive and included more instances where delays exceeded 5 min. The largest delay of app-reported sampling times was the one of S3 with a median delay of more than 1.5 min. The IQR of delays of the *App* reporting strategies were also higher, with the largest one being the one from S3 (1.58 min).

Reporting strategy	S1	S2	S3	S4
Self-report	0.0 (0.00)	0.0 (0.00)	0.00 (0.00)	0.00 (1.00)
App	0.3 (0.33)	0.5 (0.58)	0.95 (1.58)	1.05 (1.49)

Table 4.2: Sampling delays of *CAR samples* grouped by the reporting strategies; values are reported as median (IQR) in minutes

Reporting strategy	S5	S6
Self-report	0.00 (1.00)	0.00 (1.00)
App	1.05 (0.92)	0.95 (1.03)

Table 4.3: Sampling delays of *fixed-time samples* grouped by the reporting strategies; values are reported as median (IQR) in minutes.

#### *Fixed-time samples*

Outliers were more pronounced in the *fixed-time samples* compared to the CAR samples for both reporting strategies, with some samples exhibiting delays of over 1 hour, as shown in Figure 4.3. However, since most samples were taken close to the sampling schedule, the mean sampling delay was very low. As depicted in the histogram in Figure 4.4, more than 60 % had a maximum delay of 1 min for the reporting strategy *Self-report*. For the reporting strategy *App*, approximately 20 % of samples had a maximum delay of 1 min, and over 50 % had a delay of 1 to 2 min. The resulting median and IQR values are shown in Table 4.3.

### 4.3.2 Sampling delays on workdays compared to sampling delays on non-workdays

The overall difference between sampling delays on workdays and non-workdays for *CAR samples* was very low (Table 4.4). As depicted in Figure 4.5, sampling delays only differed in the outliers, with more negative outliers on non-workdays compared to workdays. On non-workdays, the median delay increased from S1 to S4, while on workdays, the median delay reached its maximum at S3 and then slightly decreased again.

The median and IQR values for sampling time delays of the *CAR samples* are very similar when comparing within the same reporting strategy across different day types (Table 4.5). For the reporting strategy *App* consistent median delays and IQRs for all sampling times were captured. Similarly, the median sampling delays for the reporting strategy *Self-report* were 0 for all samples with negligible IQR values, indicating no considerable variation in sampling delays between workdays and non-workdays within each reporting strategy.

Similar results were obtained for the absolute samples, which also only showed small differences between the day types and within the reporting strategies (see Table 4.6).

A greater difference was observed after further dividing the samples of the individual day types into the two subgroups *employed* (containing samples of participants who are in employment or training) and *student* (containing samples of participants who are studying). Table 4.7 indicates

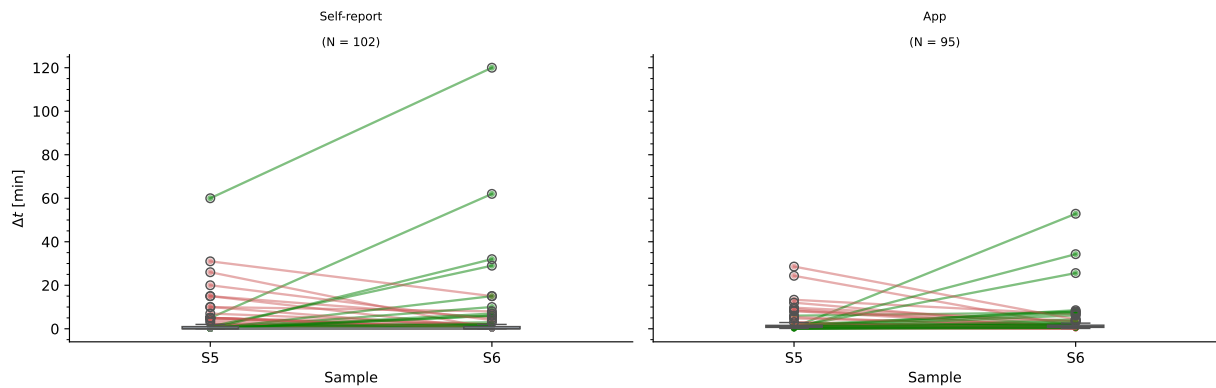


Figure 4.3: Sampling delays of individual participants grouped by the reporting strategies. The Figure includes all *fixed-time samples*.

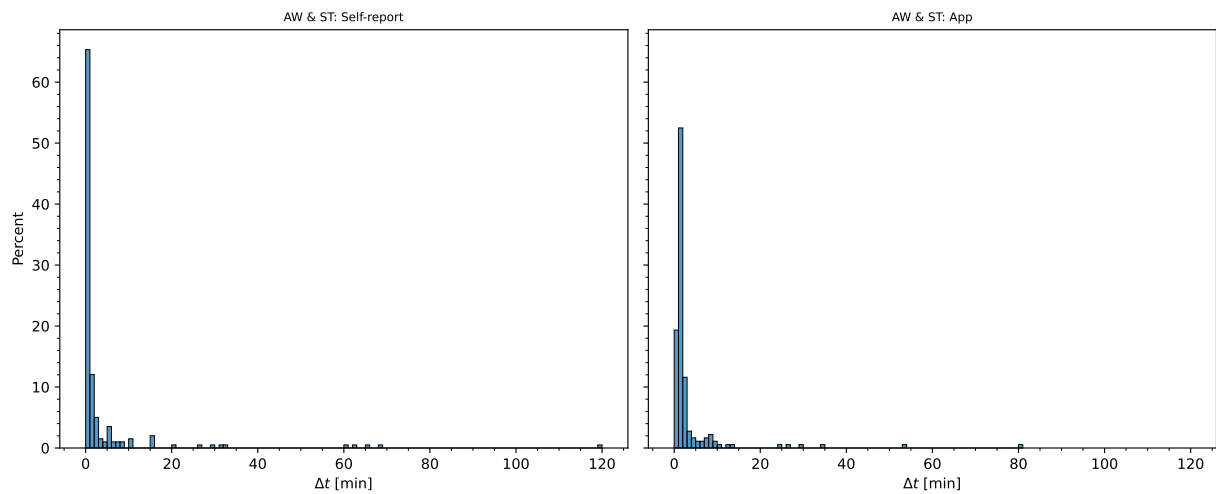


Figure 4.4: Distribution of sampling delays of reporting strategies *Self-report* and *App* for *fixed-time samples*.

Day type	S1	S2	S3	S4
Workday	0.00 (0.29)	0.29 (0.79)	0.37 (1.36)	0.43 (1.16)
Non-workday	0.00 (0.25)	0.20 (0.62)	0.43 (1.07)	0.52 (1.30)

Table 4.4: Sampling delays of *CAR samples* grouped by the day type; values are reported as median (IQR) in minutes

that for the reporting strategy *Self-report*, the delays across all samples (S1 to S4) on both workdays and non-workdays were minimal for both participant groups as the median delay was consistently at 0 min, with low IQR values.



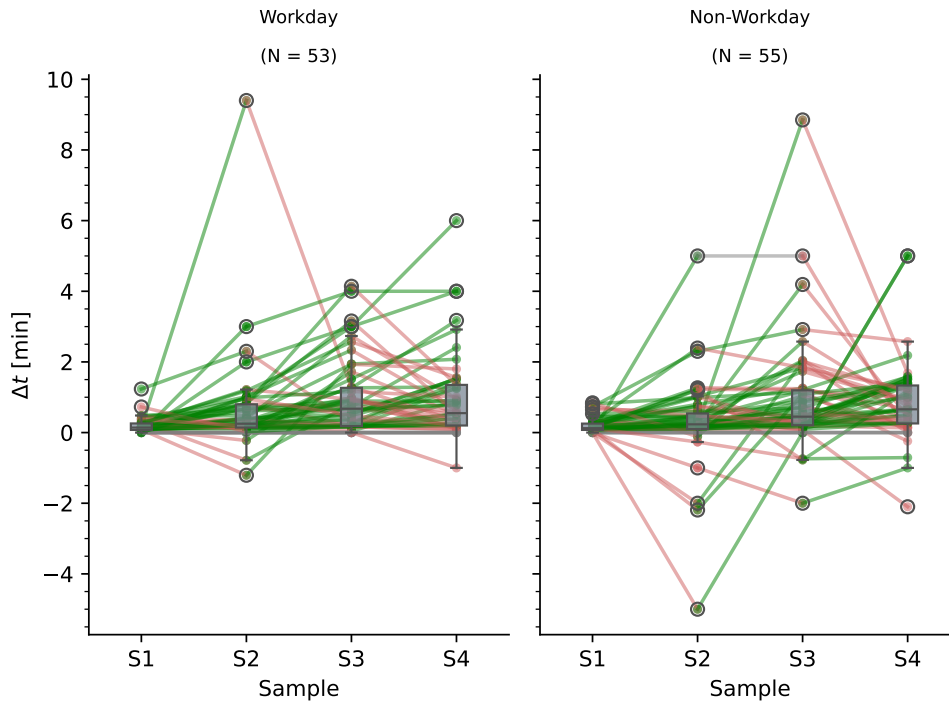


Figure 4.5: Sampling delays of individual participants grouped by the day type. The Figure includes all *CAR samples* of the reporting strategies *App* and *Self-report*.

Day type	Reporting strategy	S1	S2	S3	S4
Workday	App	0.31 (0.39)	0.48 (0.51)	0.97 (1.53)	1.03 (1.41)
Workday	Self-report	0.00 (0.00)	0.00 (0.00)	0.00 (1.00)	0.00 (1.00)
Non-workday	App	0.28 (0.31)	0.50 (0.58)	0.93 (1.73)	1.05 (1.57)
Non-workday	Self-report	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.75)

Table 4.5: Sampling delays of *CAR samples* grouped by day type and reporting strategy; values are reported as median (IQR) in minutes

Day type	Reporting strategy	S5	S6
Workday	App	1.09 (1.13)	0.83 (0.85)
Non-workday	App	1.02 (0.8)	1.18 (1.18)
Workday	Self-report	0.0 (2.0)	0.0 (1.0)
Non-workday	Self-report	0.0 (1.0)	0.0 (1.5)

Table 4.6: Sampling delays of *fixed samples* grouped by day type and reporting strategy; values are reported as median (IQR) in minutes

Day type	Profession	Reporting strategy	S1	S2	S3	S4
Workday	Employed	App	0.48 (0.45)	0.37 (0.7)	0.57 (0.87)	1.1 (0.98)
Workday	Employed	Self-report	0.0 (0.0)	0.0 (0.5)	0.0 (1.5)	0.0 (1.5)
Workday	Student	App	0.27 (0.27)	0.46 (0.39)	0.88 (1.47)	0.73 (1.4)
Workday	Student	Self-report	0.0 (0.0)	0.0 (0.0)	0.0 (0.75)	0.0 (0.75)
Non-workday	Employed	App	0.23 (0.15)	0.5 (0.47)	1.05 (1.08)	2.37 (1.87)
Non-workday	Employed	Self-report	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.5)
Non-workday	Student	App	0.33 (0.32)	0.49 (0.55)	0.93 (1.85)	0.89 (1.03)
Non-workday	Student	Self-report	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.5)

Table 4.7: Sampling delays of *CAR samples* grouped by the day type, profession, and reporting strategy; values are reported as median (IQR) in minutes

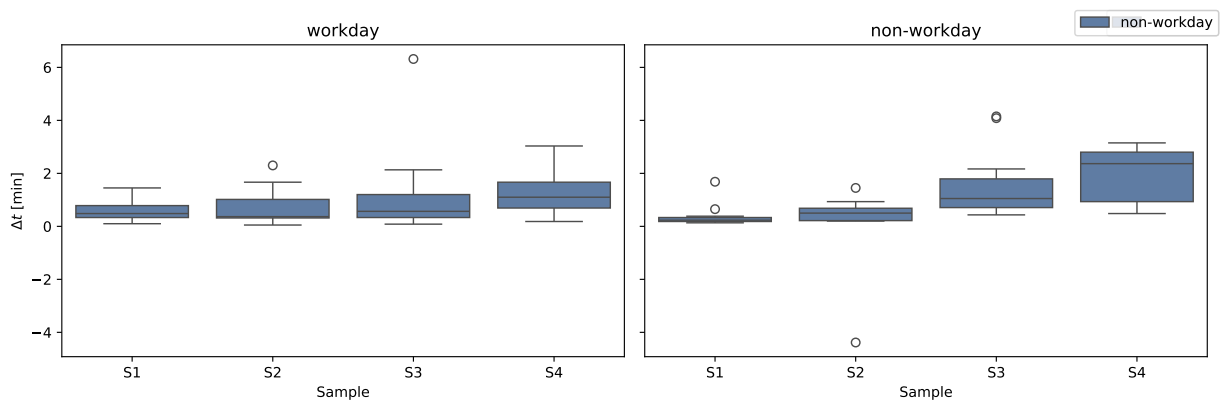


Figure 4.6: Sampling delays of employed participants grouped by the day type. Only samples of the reporting strategy *App* are included.

Regarding the reporting strategy *App*, time delays for employed participants on non-workdays were higher and also showed a stronger increase compared to the time delays on workdays. The longest median delay and largest IQR were observed for S4, with a delay of 2.37 min and an IQR of 1.87 min. In contrast, on workdays, the median delay for the same group started at 0.48 min for S1 and increased to 1.1 min for S4, showing a more moderate rise. The delay times for S3 and S4 on workdays were only half of those on non-workdays as shown in Figure 4.6.

For students, the delays were also higher on non-workdays compared to workdays, but the difference was less pronounced than for employed participants. On non-workdays, the median delay for students rose from 0.33 min for S1 to 0.89 min for S4, while on workdays, it ranged from 0.27 min for S1 to 0.73 min for S4.

In summary, the analysis highlights a difference in sampling delays between employed participants and students regarding the reporting strategy *App* on non-workdays. The delays for the reporting strategy *Self-report*, however, were minimal regardless of the day type or participant profession.

The results for samples S5 and S6 were similar to those observed for the earlier samples. For the reporting strategy *Self-report*, there was almost no delay in sampling times, except for S6 for employed participants on non-workdays. This instance marked the largest delay for the reporting strategy *Self-report*, with a median delay of 1.0 min and an IQR of 5.75 min. For the reporting strategy *App*, delays for fixed-time samples were slightly higher compared to the *CAR samples*. The longest median delay and largest IQR were observed for S6 on non-workdays for employed participants, with a median delay of 1.85 min and an IQR of 2.98 min. The median delays and IQRs for the reporting strategy *App* for samples S5 and S6 were somewhat higher than those observed for the *CAR samples*. Although the *App* reporting strategy showed slightly higher delays for fixed time samples (S5 and S6) compared to *CAR samples*, the delays were still minimal overall.



# Chapter 5

## Discussion

The following enhancements were made in the scope of this thesis. The *sleep data reader* can be used to integrate external sleep recording devices. This is especially useful if participants do not set an alarm or wake up prior to the set alarm, as it provides an objective method to record the awakening time so that the self-reported awakening time can be verified. Additionally, it can be used to analyze the participants' sleep quality, which can have an unwanted influence on the CAR [Las08]. However, the sleep data reader relies on the accuracy of the external device, which should be considered in studies utilizing this feature.

The *light intensity logger* was added to monitor the ambient light level, which is considered a potential variable that can lead to confounding [Sta16]. Unlike the sleep data reader, the light intensity logger uses the smartphone's internal sensors. Therefore, researchers should consider the accuracy of these sensors, as they can vary between different smartphones. Furthermore, the participants should be instructed on how to activate the light intensity logger before going to bed and that the light sensor should not be covered by any objects.

The *welcome slideshow* simplifies the first steps that are done after the installation of the Android app and serves as an introduction to the app. Although it does not directly improve the objectiveness of sampling and awakening time or the monitoring of covariates, it could prevent user errors.

The rework of the alarm overview increased its clarity by including all alarms and showing additional information on the alarms with the help of the symbols next to the alarms. Although not examined in the scope of this thesis, this can improve the participants' sampling behavior as they now always have access to the sampling times in a single app screen.

Before discussing the data collected in the TriASS, some issues regarding the study design should be mentioned. Firstly, there was no objective assessment of awakening time, as participants

were not instructed to use the awakening alarm provided by the app or any other objective verification method for awakening time. They were also not instructed to record their awakening time separately from the sampling times, which could have led to errors. When comparing the delays of this study to those measured in the proof-of-concept study for the *CARWatch* app [Ric23], a noticeable difference in S1 delays was observed. In Richer’s study, the self-reported S1 delay and S1 delay obtained from the app were approximately 1 min, while in this study, the S1 delays were 0 and 0.3 min. Therefore, awakening times should be measured objectively or at least reported separately from sampling times. A few issues with the app, such as sample alarms not going off, were also identified. This resulted in missing sampling times in the app logs. Consequently, out of the 111 days remaining after excluding cortisol outliers, 20 more days had to be excluded because not all *CAR samples* had matching timestamps in the app logs.

Overall, the results indicated that sampling adherence was generally high, and sampling delays were low across both reporting strategies and for both *CAR* and *fixed-time samples*. However, participants tended to slightly overestimate their adherence to the sampling schedule, consistent with findings by Broderick et al. [Bro04]. This overestimation was indicated by computed sampling delays for self-reported sampling being smaller than those for the reporting strategy *App*. These small differences align with the findings of Richer et al. [Ric23]. In relation to compliance windows used in existing research (e.g., 15 min [Bro04] or 10 min [Kud07]) the median sampling delay in this study was rather low, with almost no sampling day having *CAR* samples taken outside these windows. This consistency is also evident when examining the statistical impact of the reporting strategies on the computed *CAR* curve. The findings of the rmANOVA revealed that the reporting strategy had a significant, though low, impact on the computed  $AUC_G$  and  $AUC_I$ . Notably, pairwise test results indicated no significant difference between the reporting strategies *Naïve* and *Self-report*, but a significant difference between the strategies *App* and *Self-report*, underscoring the assumption that participants tend to overestimate their adherence and that sampling times should be assessed objectively as a supplement to self-reported times.

Contrary to the literature [Kun04; Sch04], the type of day (workday or non-workday) had no statistically significant effect on the *CAR*, possibly influenced by the participants’ employment status, as most were studying at the time of assessment. However, when performing the same analysis on employed participants only, a significant workday-non-workday difference in the  $AUC_I$  was found, as shown by the results of the mixed ANOVA. This difference was visible in the median values for the respective day types, which were much lower on non-workdays than on workdays. Nonetheless, the high IQR values indicate a wide spread of values, which could have been affected by the small sample size, as complete measurements were available for only 10 workdays and 9

non-workdays. Similar findings were observed for the sampling times. While the differences in sampling delays between non-workdays and workdays for all participants were low, the differences for employed participants were more pronounced and increased throughout the morning. This could be because workday mornings are more stressful, leaving less time available, which causes samples to be postponed. As already mentioned for the CAR curve, these results could have been influenced by the small sample size, as only a few complete sampling days were available for employed participants. However, even these larger sampling delays and differences between app-reported sampling intervals and self-reported sampling intervals were probably too low to impact the resulting CAR metrics.

In summary, the study faced a few design and app-related problems that could have affected the results, such as the lack of objective awakening time assessment, app-related issues, and a small sample size. Despite these issues, the overall adherence to the sampling schedule was high, with minimal delays. However, participants tended to overestimate their adherence, highlighting the need for objective assessment methods. The type of day had no significant impact on the CAR for most participants, but differences were observed for employed participants, possibly due to the stress of workday mornings. Overall, the CARWatch framework serves as a supportive tool for CAR assessments. Though improvements in robustness are still needed, the verification of sampling times works well, and samples with a large difference between the self-reported and app-obtained sampling times can be sorted out. However, when conducting a study, it is essential to either include awakening by alarm or an additional sensor that measures awakening time, as awakening time cannot be objectively verified through CARWatch by solely using the self-report functionality.





# Chapter 6

## Conclusion & Outlook

As part of this work, the CARWatch framework was enhanced with several functionalities. The sleep data reader was added as an additional objective verification method for awakening times and can further be used to measure the participant's sleep behavior. The light intensity logger was added to measure the light level a participant is exposed to in the morning, as it is a potential confounding that can have an unwanted effect on the CAR. Reworked Android app screens and an Android app introduction were implemented to improve the participant's sampling behavior and the user experience. Furthermore, smaller adjustments were made to increase the customizability and flexibility of the framework.

The CARWatch framework served as an objective sampling time verification tool in the TriASS. The study data were examined to uncover participants' sampling behavior and its impact on the derived CAR curve. Moreover, the analysis included differences in the CAR and sampling patterns between workdays and non-workdays. The general sampling behavior was highly consistent with minimal differences between self-reported sampling times and sampling times recorded by the Android app. Participants also showed a high sampling adherence with minimal delays between scheduled sampling times and recorded sampling times. Differences between non-workdays and workdays for  $AUC_G$ ,  $AUC_I$ , and sampling times were minimal. Further analyzing separate subgroups revealed an increased CAR and longer sampling delays on non-workdays for employed participants. However, only 6 employees participated in the study, which decreases the significance of the findings. Therefore, future studies should further investigate the differences in the CAR between groups at different educational phases, especially focusing on the different day types and stress levels of subgroups. Overall, the results showed that CARWatch is capable of supporting

CAR studies as an objective verification tool that can easily be integrated into and adjusted to the study. Additionally, CARWatch can be downloaded for free, making it economically valid even for large-scale studies.

Nevertheless, some improvements would be beneficial for future studies. These include improving the robustness to reduce the proportion of missing sampling times and adding functionality to embed supplementary questionnaires so that participant and sampling day-related data can be assessed directly through the CARWatch Android app. The former could be directly integrated into the welcome slideshow. The latter could include questions regarding sleep quality and expectations for the day ahead, similar to the protocol questions used in the TriASS, as these are covariates of the CAR [Sta16]. Additionally, the awakening time should be assessed independently if the awakening was manually reported by the participant and is not measured by an external device.

In summary, the CARWatch framework, which has already proven to be a reliable objective verification tool for sampling times, was extended by new data collection methods. However, these methods still need to be evaluated and additional improvements might increase the framework's benefit for CAR studies. Furthermore, the findings of this thesis should be further investigated in future studies.

# List of Figures

2.1	Illustration of the impact of delayed awakening time reporting on the estimated CAR profile [Smy13]. . . . .	6
2.2	Illustration of the impact of delayed sampling on the computed $AUC_I$ . The Figure depicts estimates given (a) correct sampling and for delays of (b) 8 min, (c) 20 min, and (d) 40 min between awakening and collecting S1 [Sta16]. . . . .	7
3.1	CARWatch framework applications during CAR assessment study. . . . .	14
3.2	Study configuration step 1. . . . .	16
3.3	Study configuration step 2. . . . .	16
3.4	Study configuration step 3. . . . .	17
3.5	Barcodes generated by the webapp. . . . .	18
3.6	Generated barcode. The digits above the code are the barcode value. The label below the barcode provides information about the participant and the sample to which the barcode belongs. . . . .	18
3.7	Generated QR code containing the study configuration. The participant ID that is encoded in the QR code is displayed below the QR code. . . . .	19
3.8	Main screen of the app consisting of the header (highlighted red), the content area that holds the content screens (highlighted green), and the bottom navigation bar (highlighted purple). The content screen displayed is the wakeup screen. . . . .	21
3.9	Unfolded app menu. . . . .	21
3.10	Alarm screen with wake-up alarm and sample alarms. Active alarms are highlighted in orange. The symbols behind the alarm times indicate whether the sample was taken (checkmark), should have been already taken (orange clock), or is scheduled for later (sand clock). The barcode symbol can be used to open the barcode scanner. . . . .	22

3.11	Bedtime screen that can be used to initiate the evening procedure by clicking the “Yes” button. . . . .	22
3.12	Alarm notification if display is locked (a) and if display is unlocked (b). . . . .	24
3.13	Scanner for sample device barcodes. The highlighted area must be pointed at the barcode. . . . .	24
3.14	Slide sequence of slide show after app installation. If the slide show is started via the “Reregister” menu entry, it starts with the QR code scanner slide and finished before the tutorial slides would be shown. If the tutorial option in the menu is clicked, the slide show starts and finishes with the tutorial slides. . . . .	26
3.15	(a) First slide that is shown when the app is opened for the first time. The green area highlights the slide content, the purple area highlights the slide navigation bar. (b) Tutorial slide that explains the alarm screen. Each tutorial slide has the same structure and contains a screenshot of the feature that is explained as well as an explanatory text. The tutorial slides can be skipped with the button in the lower left corner. . . . .	27
3.16	Slide show classes and their relations. The diagram does not contain all class attributes and methods, only those mentioned in the slide show Section. . . . .	28
3.17	Empty protocol card for a single day of sample collection. . . . .	32
4.1	$AUC_G$ (a) and $AUC_I$ (b) computed from sampling times using different reporting strategies. The green lines between reporting strategies indicate an increase of $AUC_I$ values, red lines indicate a decrease, gray lines indicate no change. . . . .	37
4.2	Sampling delays of individual participants grouped by the reporting strategies. The Figure includes all <i>CAR samples</i> . . . . .	38
4.3	Sampling delays of individual participants grouped by the reporting strategies. The Figure includes all <i>fixed-time samples</i> . . . . .	40
4.4	Distribution of sampling delays of reporting strategies <i>Self-report</i> and <i>App</i> for <i>fixed-time samples</i> . . . . .	40
4.5	Sampling delays of individual participants grouped by the day type. The Figure includes all <i>CAR samples</i> of the reporting strategies <i>App</i> and <i>Self-report</i> . . . . .	41
4.6	Sampling delays of employed participants grouped by the day type. Only samples of the reporting strategy <i>App</i> are included. . . . .	42

# List of Tables

3.1	Exemplary sleep phases before awakening. . . . .	29
4.1	AUC <sub>G</sub> and AUC <sub>I</sub> in $\frac{\text{nmol}}{\text{L}}$ of employed participants on workdays and non-workdays reported as median (IQR). Only sampling days of the reporting strategies <i>Self-report</i> and <i>App</i> were analyzed. . . . .	38
4.2	Sampling delays of <i>CAR samples</i> grouped by the reporting strategies; values are reported as median (IQR) in minutes . . . . .	38
4.3	Sampling delays of <i>fixed-time samples</i> grouped by the reporting strategies; values are reported as median (IQR) in minutes. . . . .	39
4.4	Sampling delays of <i>CAR samples</i> grouped by the day type; values are reported as median (IQR) in minutes . . . . .	40
4.5	Sampling delays of <i>CAR samples</i> grouped by day type and reporting strategy; values are reported as median (IQR) in minutes . . . . .	41
4.6	Sampling delays of <i>fixed samples</i> grouped by day type and reporting strategy; values are reported as median (IQR) in minutes . . . . .	41
4.7	Sampling delays of <i>CAR samples</i> grouped by the day type, profession, and reporting strategy; values are reported as median (IQR) in minutes . . . . .	42



# Bibliography

- [Ada09] Emma K. Adam and Meena Kumari. “Assessing salivary cortisol in large-scale, epidemiological research”. In: *Psychoneuroendocrinology* 34.10 (2009). tex.ids= Adam2009a publisher: Elsevier Ltd, pp. 1423–1436. issn: 03064530. doi: 10.1016/j.psyneuen.2009.06.011. url: <http://dx.doi.org/10.1016/j.psyneuen.2009.06.011>.
- [Bäu13] Damaris Bäuml, Clemens Kirschbaum, Matthias Kliegel, Nina Alexander, and Tobias Stalder. “The cortisol awakening response in toddlers and young children”. In: *Psychoneuroendocrinology* 38.11 (Nov. 2013), pp. 2485–2492. issn: 03064530. doi: 10.1016/j.psyneuen.2013.05.008. url: <https://linkinghub.elsevier.com/retrieve/pii/S0306453013001856> (visited on 04/16/2024).
- [Bed19] Theresa Beddig, Iris Reinhard, and Christine Kuehner. “Stress, mood, and cortisol during daily life in women with Premenstrual Dysphoric Disorder (PMDD)”. In: *Psychoneuroendocrinology* 109 (Nov. 2019). tex.ids= Beddig2019a, p. 104372. issn: 03064530. doi: 10.1016/j.psyneuen.2019.104372. url: <https://linkinghub.elsevier.com/retrieve/pii/S0306453019302100> (visited on 10/18/2022).
- [Bro04] Joan E. Broderick, Daniel Arnold, Brigitte M. Kudielka, and Clemens Kirschbaum. “Salivary cortisol sampling compliance: Comparison of patients and healthy volunteers”. In: *Psychoneuroendocrinology* 29.5 (2004). tex.ids= Broderick2004a, pp. 636–650. issn: 03064530. doi: 10.1016/S0306-4530(03)00093-3.
- [Chi09] Yoichi Chida and Andrew Steptoe. “Cortisol awakening response and psychosocial factors: A systematic review and meta-analysis”. In: *Biological Psychology* 80.3 (2009), pp. 265–278. issn: 03010511. doi: 10.1016/j.biopsycho.2008.10.004.
- [Chi22] Evan D Chinoy, Joseph A Cuellar, Jason T Jameson, and Rachel R Markwald. “Performance of four commercial wearable sleep-tracking devices tested under unrestricted conditions at home in healthy young adults”. In: *Nat. Sci. Sleep* 14 (Mar. 2022), pp. 493–516.

- [Cho22] Man Ho Choi. “Clinical and technical aspects in free cortisol measurement”. In: *Endocrinol. Metab. (Seoul)* 37.4 (Aug. 2022). Publisher: Korean Endocrine Society, pp. 599–607.
- [Clo04] Angela Clow, Lisa Thorn, Phil Evans, and Frank Hucklebridge. “The Awakening Cortisol Response: Methodological Issues and Significance”. In: *Stress* 7.1 (2004), pp. 29–37. ISSN: 10253890. DOI: 10.1080/10253890410001667205.
- [Clo10] Angela Clow, Frank Hucklebridge, Tobias Stalder, Phil Evans, and Lisa Thorn. “The cortisol awakening response: More than a measure of HPA axis function”. In: *Neuroscience and Biobehavioral Reviews* 35.1 (2010). Publisher: Elsevier Ltd, pp. 97–103. ISSN: 01497634. DOI: 10.1016/j.neubiorev.2009.12.011. URL: <http://dx.doi.org/10.1016/j.neubiorev.2009.12.011>.
- [Coh07] Sheldon Cohen, Denise Janicki-Deverts, and Gregory E. Miller. “Psychological Stress and Disease”. In: *JAMA* 298.14 (Oct. 2007). \_eprint: [https://jamanetwork.com/journals/jama/articlepdf/209083/jco70057\\_1685\\_1687.pdf](https://jamanetwork.com/journals/jama/articlepdf/209083/jco70057_1685_1687.pdf), pp. 1685–1687. ISSN: 0098-7484. DOI: 10.1001/jama.298.14.1685. URL: <https://doi.org/10.1001/jama.298.14.1685>.
- [Eld14] Greg J. Elder, Mark A. Wetherell, Nicola L. Barclay, and Jason G. Ellis. “The cortisol awakening response - Applications and implications for sleep medicine”. In: *Sleep Medicine Reviews* 18.3 (2014). Publisher: Elsevier Ltd, pp. 215–224. ISSN: 15322955. DOI: 10.1016/j.smr.2013.05.001. URL: <http://dx.doi.org/10.1016/j.smr.2013.05.001>.
- [Fig12] Mariana G. Figueiro and Mark S. Rea. “Short-Wavelength Light Enhances Cortisol Awakening Response in Sleep-Restricted Adolescents”. In: *International Journal of Endocrinology* 2012 (2012). Publisher: Hindawi Limited, pp. 1–7. ISSN: 1687-8345. DOI: 10.1155/2012/301935. URL: <http://dx.doi.org/10.1155/2012/301935>.
- [Fri09] Eva Fries, Lucia Dettenborn, and Clemens Kirschbaum. “The cortisol awakening response (CAR): Facts and future directions”. In: *International Journal of Psychophysiology* 72.1 (2009). Publisher: Elsevier B.V., pp. 67–73. ISSN: 01678760. DOI: 10.1016/j.ijpsycho.2008.03.014. URL: <http://dx.doi.org/10.1016/j.ijpsycho.2008.03.014>.
- [Gol14] Sherita Hill Golden, Brisa N. Sánchez, Amy S. DeSantis, Meihua Wu, Cecilia Castro, Teresa E. Seeman, Sameh Tadros, Sandi Shrager, and Ana V. Diez Roux. “Salivary cortisol protocol adherence and reliability by socio-demographic features: The Multi-Ethnic Study of Atherosclerosis”. In: *Psychoneuroendocrinology* 43 (May



- 2014), pp. 30–40. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2014.01.025. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453014000511> (visited on 06/12/2022).
- [Gri10] Barbara Griefahn and Sibylle Robens. “The normalization of the cortisol awakening response and of the cortisol shift profile across consecutive night shifts – An experimental study”. In: *Psychoneuroendocrinology* 35.10 (2010). Publisher: Elsevier Ltd, pp. 1501–1509. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2010.05.002. URL: <http://dx.doi.org/10.1016/j.psyneuen.2010.05.002>.
- [Gri11] Barbara Griefahn and Sibylle Robens. “Cortisol awakening response - Are sampling delays of 15 minutes acceptable?” In: *International Journal of Psychophysiology* 82.2 (2011). tex.ids= Griefahn2011a publisher: Elsevier B.V., pp. 202–205. ISSN: 01678760. DOI: 10.1016/j.ijpsycho.2011.08.005. URL: <http://dx.doi.org/10.1016/j.ijpsycho.2011.08.005>.
- [Haa24] Lisa Haase, Antonia Vehlen, Julia Strojny, and Gregor Domes. “Effects of menstrual cycle phase and ovulation on the salivary cortisol awakening response”. In: *Psychoneuroendocrinology* 160 (2024), p. 106669. ISSN: 0306-4530. DOI: <https://doi.org/10.1016/j.psyneuen.2023.106669>. URL: <https://www.sciencedirect.com/science/article/pii/S0306453023006479>.
- [Hag19] Shahab Haghayegh, Sepideh Khoshnevis, Michael H Smolensky, Kenneth R Diller, and Richard J Castriotta. “Accuracy of Wristband Fitbit Models in Assessing Sleep: Systematic Review and Meta-Analysis”. In: *J Med Internet Res* 21.11 (Nov. 28, 2019), e16273. ISSN: 1438-8871. DOI: 10.2196/16273. URL: <http://www.ncbi.nlm.nih.gov/pubmed/31778122>.
- [Hui94] H. V. Huikuri, M. J. Niemelä, S. Ojala, A. Rantala, M. J. Ikäheimo, and K. E. Airaksinen. “Circadian rhythms of frequency domain measures of heart rate variability in healthy subjects and patients with coronary artery disease. Effects of arousal and upright posture.” In: *Circulation* 90.1 (1994). \_eprint: <https://www.ahajournals.org/doi/pdf/10.1161/01.CIR.90.1.121>, pp. 121–126. DOI: 10.1161/01.CIR.90.1.121. URL: <https://www.ahajournals.org/doi/abs/10.1161/01.CIR.90.1.121>.
- [Kud03a] B.M. Kudielka and C. Kirschbaum. “Awakening cortisol responses are influenced by health status and awakening time but not by menstrual cycle phase”. In: *Psychoneuroendocrinology* 28.1 (Jan. 2003), pp. 35–47. ISSN: 03064530. DOI: 10.1016/

- S0306-4530(02)00008-2. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453002000082> (visited on 06/23/2022).
- [Kud03b] Brigitte M. Kudielka, Joan E. Broderick, and Clemens Kirschbaum. “Compliance With Saliva Sampling Protocols: Electronic Monitoring Reveals Invalid Cortisol Daytime Profiles in Noncompliant Subjects”. In: *Psychosomatic Medicine* 65.2 (Mar. 2003). tex.ids: Kudielka2003, pp. 313–319. ISSN: 0033-3174. DOI: 10.1097/01.PSY.0000058374.50240.BF. URL: <http://journals.lww.com/00006842-200303000-00019> (visited on 07/19/2020).
- [Kud07] Brigitte M. Kudielka, Louise C. Hawkley, Emma K. Adam, and John T. Cacioppo. “Compliance with ambulatory saliva sampling in the Chicago Health, Aging, and Social Relations Study and associations with social support”. In: *Annals of Behavioral Medicine* 34.2 (June 2007), pp. 209–216. ISSN: 0883-6612, 1532-4796. DOI: 10.1007/BF02872675. URL: <https://academic.oup.com/abm/article/34/2/209-216/4569389> (visited on 06/12/2022).
- [Kud10] Brigitte M. Kudielka and Stefan Wüst. “Human models in acute and chronic stress: Assessing determinants of individual hypothalamus–pituitary–adrenal axis activity and reactivity”. In: *Stress* 13.1 (Jan. 2010). tex.ids= Kudielka2010, pp. 1–14. ISSN: 1025-3890, 1607-8888. DOI: 10.3109/10253890902874913. URL: <http://www.tandfonline.com/doi/full/10.3109/10253890902874913> (visited on 06/23/2022).
- [Kun04] Sabine R Kunz-Ebrecht, Clemens Kirschbaum, Michael Marmot, and Andrew Steptoe. “Differences in cortisol awakening response on work days and weekends in women and men from the Whitehall II cohort”. In: *Psychoneuroendocrinology* 29.4 (May 2004), pp. 516–528. ISSN: 03064530. DOI: 10.1016/S0306-4530(03)00072-6. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453003000726> (visited on 04/16/2024).
- [Kup05] Nina Kupper, Eco J.C. de Geus, Mireille van Den Berg, Clemens Kirschbaum, Dorret I. Boomsma, and Gonneke Willemsen. “Familial influences on basal salivary cortisol in an adult population”. In: *Psychoneuroendocrinology* 30.9 (2005), pp. 857–868. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2005.04.003.
- [Las08] N. Lasikiewicz, H. Hendrickx, D. Talbot, and L. Dye. “Exploration of basal diurnal salivary cortisol profiles in middle-aged adults: Associations with sleep quality and metabolic parameters”. In: *Psychoneuroendocrinology* 33.2 (2008), pp. 143–151.

ISSN: 0306-4530. DOI: <https://doi.org/10.1016/j.psyneuen.2007.10.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0306453007002454>.

- [Lic06] Kenneth L. Lichstein, Kristen C. Stone, James Donaldson, Sidney D. Nau, James P. Soeffing, David Murray, Kristin W. Lester, and R. Neal Aguillard. “Actigraphy Validation with Insomnia”. In: *Sleep* 29.2 (Feb. 2006). \_eprint: <https://academic.oup.com/sleep/article-pdf/29/2/232/8502644/sleep-29-2-232.pdf>, pp. 232–239. ISSN: 0161-8105. DOI: 10.1093/sleep/29.2.232. URL: <https://doi.org/10.1093/sleep/29.2.232>.
- [Mar15] Brian Lovell Mark A. Wetherell and Michael A. Smith. “The effects of an anticipated challenge on diurnal cortisol secretion”. In: *Stress* 18.1 (2015). Publisher: Taylor & Francis \_eprint: <https://doi.org/10.3109/10253890.2014.993967>, pp. 42–48. DOI: 10.3109/10253890.2014.993967. URL: <https://doi.org/10.3109/10253890.2014.993967>.
- [Oku10] Michele L. Okun, Robert T. Krafty, Daniel J. Buysse, Timothy H. Monk, Charles F. Reynolds, Amy Begley, and Martica Hall. “What constitutes too long of a delay? Determining the cortisol awakening response (CAR) using self-report and PSG-assessed wake time”. In: *Psychoneuroendocrinology* 35.3 (Apr. 1, 2010). Publisher: Pergamon, pp. 460–468. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2009.08.017. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453009002625> (visited on 07/19/2020).
- [Pow12] Daniel J. Powell and Wolff Schlotz. “Daily Life Stress and the Cortisol Awakening Response: Testing the Anticipation Hypothesis”. In: *PLoS ONE* 7.12 (Dec. 20, 2012). Ed. by Tobias Deschner, e52067. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0052067. URL: <https://dx.plos.org/10.1371/journal.pone.0052067> (visited on 11/09/2020).
- [Pow15] Daniel J.H. Powell, Rona Moss-Morris, Christina Lioffi, and Wolff Schlotz. “Circadian cortisol and fatigue severity in relapsing-remitting multiple sclerosis”. In: *Psychoneuroendocrinology* 56 (June 2015). tex.ids= Powell2015a, pp. 120–131. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2015.03.010. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453015000979> (visited on 10/18/2022).
- [Pru03] Jens C. Pruessner, Clemens Kirschbaum, Gunther Meinlschmid, and Dirk H. Hellhammer. “Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change”. In: *Psychoneu-*

- roendocrinology* 28.7 (2003), pp. 916–931. ISSN: 03064530. DOI: 10.1016/S0306-4530(02)00108-7.
- [Pru97] Jens C Pruessner, Oliver T Wolf, Dirk H Hellhammer, Angelika Buske-Kirschbaum, Kristin Von Auer, Silke Jobst, Frank Kaspers, and Clemens Kirschbaum. “Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity”. In: *Life sciences* 61.26 (1997). tex.ids= PruessnerJC1, pruessner1997free publisher: Elsevier, pp. 2539–2549.
- [Ric21] Robert Richer, Arne Küderle, Martin Ullrich, Nicolas Rohleder, and Bjoern Eskofier. “BioPsyKit: A Python package for the analysis of biopsychological data”. In: *Journal of Open Source Software* 6.66 (Oct. 12, 2021). tex.ids= Richer, Richer2021, Richer2021b publisher: The Open Journal, p. 3702. ISSN: 2475-9066. DOI: 10.21105/joss.03702. URL: <https://joss.theoj.org/papers/10.21105/joss.03702> (visited on 10/16/2021).
- [Ric23] Robert Richer, Luca Abel, Arne Küderle, Bjoern M. Eskofier, and Nicolas Rohleder. “CARWatch – A smartphone application for improving the accuracy of cortisol awakening response sampling”. In: *Psychoneuroendocrinology* 151 (May 2023), p. 106073. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2023.106073. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453023000513> (visited on 03/06/2023).
- [Rod18] Kerri E. Rodriguez, Crystal I. Bryce, Douglas A. Granger, and Marguerite E. O’Haire. “The effect of a service dog on salivary cortisol awakening response in a military population with posttraumatic stress disorder (PTSD)”. In: *Psychoneuroendocrinology* 98 (Dec. 2018). tex.ids= Rodriguez2018a, pp. 202–210. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2018.04.026. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453018300441> (visited on 10/18/2022).
- [Sad02] Avi Sadeh and Christine Acebo. “The role of actigraphy in sleep medicine”. In: *Sleep Medicine Reviews* 6.2 (2002), pp. 113–124. ISSN: 1087-0792. DOI: <https://doi.org/10.1053/smr.2001.0182>. URL: <https://www.sciencedirect.com/science/article/pii/S1087079201901820>.
- [Sch04] Wolff Schlotz, Juliane Hellhammer, Peter Schulz, and Arthur A. Stone. “Perceived Work Overload and Chronic Worrying Predict Weekend-Weekday Differences in the Cortisol Awakening Response”. In: *Psychosomatic Medicine* 66.2 (2004), pp. 207–214. ISSN: 00333174. DOI: 10.1097/01.psy.0000116715.78238.56.

- [Smi14] Victoria C. Smith and Lea R. Dougherty. “Noisy spit: Parental noncompliance with child salivary cortisol sampling”. In: *Developmental Psychobiology* 56.4 (2014). \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/dev.21133>, pp. 647–656. DOI: <https://doi.org/10.1002/dev.21133>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/dev.21133>.
- [Smy13] Nina Smyth, Angela Clow, Lisa Thorn, Frank Hucklebridge, and Phil Evans. “Delays of 5–15min between awakening and the start of saliva sampling matter in assessment of the cortisol awakening response”. In: *Psychoneuroendocrinology* 38.9 (Sept. 2013). Publisher: Elsevier Ltd, pp. 1476–1483. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2012.12.013. URL: <http://dx.doi.org/10.1016/j.psyneuen.2012.12.013>.
- [Sta09] Tobias Stalder, Frank Hucklebridge, Phil Evans, and Angela Clow. “Use of a single case study design to examine state variation in the cortisol awakening response: Relationship with time of awakening”. In: *Psychoneuroendocrinology* 34.4 (May 2009), pp. 607–614. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2008.10.023. URL: 10.1016/j.psyneuen.2008.10.023 (visited on 02/16/2021).
- [Sta10] Tobias Stalder, Phil Evans, Frank Hucklebridge, and Angela Clow. “Associations between psychosocial state variables and the cortisol awakening response in a single case study”. In: *Psychoneuroendocrinology* 35.2 (2010), pp. 209–214. ISSN: 0306-4530. DOI: <https://doi.org/10.1016/j.psyneuen.2009.06.006>. URL: <https://www.sciencedirect.com/science/article/pii/S0306453009001966>.
- [Sta11] Tobias Stalder, Phil Evans, Frank Hucklebridge, and Angela Clow. “Associations between the cortisol awakening response and heart rate variability”. In: *Psychoneuroendocrinology* 36.4 (2011). tex.ids= Stalder2011a publisher: Elsevier Ltd, pp. 454–462. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2010.07.020. URL: <http://dx.doi.org/10.1016/j.psyneuen.2010.07.020>.
- [Sta16] Tobias Stalder, Clemens Kirschbaum, Brigitte M. Kudielka, Emma K. Adam, Jens C. Pruessner, Stefan Wüst, Samantha Dockray, Nina Smyth, Phil Evans, Dirk H. Hellhammer, Robert Miller, Mark A. Wetherell, Sonia J. Lupien, and Angela Clow. “Assessment of the cortisol awakening response: Expert consensus guidelines”. In: *Psychoneuroendocrinology* 63 (2016). tex.ids= Stalder2016a publisher: Elsevier Ltd, pp. 414–432. ISSN: 18733360. DOI: 10.1016/j.psyneuen.2015.10.010.

- [Sta19] Statcounter Global Stats. *Mobile Operating System Market Share Germany 2019* | Statcounter Global Stats. 2019. URL: <https://gs.statcounter.com/os-market-share/mobile/germany/2019> (visited on 06/28/2022).
- [Sta22] Tobias Stalder, Sonia J. Lupien, Brigitte M. Kudielka, Emma K. Adam, Jens C. Pruessner, Stefan Wüst, Samantha Dockray, Nina Smyth, Phil Evans, Clemens Kirschbaum, Robert Miller, Mark A. Wetherell, Johannes B. Finke, Tim Klucken, and Angela Clow. “Evaluation and update of the expert consensus guidelines for the assessment of the cortisol awakening response (CAR)”. In: *Psychoneuroendocrinology* 146 (Dec. 2022), p. 105946. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2022.105946. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453022002876> (visited on 10/18/2022).
- [Tho04] L. Thorn, F. Hucklebridge, A. Esgate, P. Evans, and A. Clow. “The effect of dawn simulation on the cortisol response to awakening in healthy participants”. In: *Psychoneuroendocrinology* 29.7 (2004), pp. 925–930. ISSN: 0306-4530. DOI: <https://doi.org/10.1016/j.psyneuen.2003.08.005>. URL: <https://www.sciencedirect.com/science/article/pii/S0306453003001604>.
- [Tri01] John Trinder, Marinella Padula, David Berlowitz, Jan Kleiman, Sibilah Breen, Peter Rochford, Christopher Worsnop, Bruce Thompson, and Robert Pierce. “Cardiac and respiratory activity at arousal from sleep under controlled ventilation conditions”. In: *Journal of Applied Physiology* 90.4 (Apr. 2001). Publisher: American Physiological Society, pp. 1455–1463. ISSN: 1522-1601. DOI: 10.1152/jappl.2001.90.4.1455. URL: <http://dx.doi.org/10.1152/jappl.2001.90.4.1455>.
- [Tri03] John Trinder, Nicholas Allen, Jan Kleiman, Vicki Kravlevski, Darci Kleverlaan, Katie Anson, and Young Kim. “On the nature of cardiovascular activation at an arousal from sleep”. In: *Sleep* 26.5 (Aug. 2003). Publisher: Oxford University Press (OUP), pp. 543–551.
- [Ulr09] Yvonne M. Ulrich-Lai and James P. Herman. “Neural regulation of endocrine and autonomic stress responses”. In: *Nature Reviews Neuroscience* 10.6 (June 1, 2009), pp. 397–409. ISSN: 1471-0048. DOI: 10.1038/nrn2647. URL: <https://doi.org/10.1038/nrn2647>.
- [Val18] Raphael Vallat. “Pingouin: statistics in Python”. In: *Journal of Open Source Software* 3.31 (2018). publisher: The Open Journal, p. 1026. DOI: 10.21105/joss.01026.

- [VAN11] ALEXANDER T. M. VAN DE WATER, ALISON HOLMES, and DEIRDRE A. HURLEY. “Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography – a systematic review”. In: *Journal of Sleep Research* 20.1 (2011). \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2869.2009.00814.x>, pp. 183–200. DOI: <https://doi.org/10.1111/j.1365-2869.2009.00814.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2869.2009.00814.x>.
- [Wüs00] Stefan Wüst, Jutta M. Wolf, Dirk H. Hellhammer, Ilona S. Federenko, Nicole C. Schommer, and Clemens Kirschbaum. “The cortisol awakening response - normal values and confounds”. In: *Noise & health* 2.7 (2000), pp. 79–88. ISSN: 1463-1741. URL: <http://www.ncbi.nlm.nih.gov/pubmed/12689474>.
- [Zhu19] Yannan Zhu, Xu Chen, Hui Zhao, Menglu Chen, Yanqiu Tian, Chao Liu, Zhuo Rachel Han, Xiuyun Lin, Jiang Qiu, Gui Xue, Hua Shu, and Shaozheng Qin. “Socioeconomic status disparities affect children’s anxiety and stress-sensitive cortisol awakening response through parental anxiety”. In: *Psychoneuroendocrinology* 103 (May 2019), pp. 96–103. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2019.01.008. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453018309004> (visited on 10/18/2022).





# Appendix A

## Acronyms

**CAR** cortisol awakening response

**PSG** polysomnography

**HPA** hypothalamic-pituitary-adrenocortical

**ANS** autonomic nervous system

**CVD** cardiovascular disease

**AUC<sub>G</sub>** integrated area under the curve

**AUC<sub>I</sub>** increase of cortisol from the level recorded on waking

**EEG** electroencephalography

**EOG** electrooculography

**EMG** electromyography

**ECG** electrocardiography

**TriASS** Trier Everyday stress study

**IBI** interbeat interval

**HTML** Hypertext Markup Language

**ID** identifier

**rmANOVA** repeated measures ANOVA

**IQR** Interquartile range