Foreword

The topic of mobile sensing in psychology may seem to be a new field powered by recent technology, but the quest for more ecological data to measure mood, behavior, and cognition has been an old one. No doubt, Freud wondered about the relationship of what he observed in the consulting room to what was happening in the real world outside. And both clinicians and scientists since have wished for better insight into the real-world experience of people in psychological distress.

There is a highly apocryphal story about the scientist who devotes her life to creating the ideal mobile sensing tools, only to pass away before seeing these tools adopted in clinical use. The story is that such virtuous work is rewarded by St. Peter who, because of her exemplary dedication to improving the human condition, offers her an audience with God and an opportunity to ask the Almighty a single question. With some trepidation, she pops the question, "Father, will we ever have a mobile sensing device that is adopted by patients and providers?" Allegedly, God responds, "Yes, my child. But not in my lifetime."

At the outset of this important volume on mobile sensing, it's important to realize that the task for mobile sensing is neither easy nor quick. It's really two tasks, both covered extensively in this volume. First is the challenge of validation. Do the signals on a wearable or smartphone provide high-quality data, and can those data be tied to some ground truth? Acquiring high-quality signals in a world of variance, interference, and nonadherence feels like one of those "not in my lifetime" challenges. But several chapters in this book demonstrate that we can collect high-quality data on location, activity, emotion, and more. Smartphones, wearables, and social media provide an unprecedented scale of data, capturing the world outside of the consulting room or psychology lab. Yes, we need to create standards for quality and we need to integrate mobile sensing data with other measures, but already we can see the value of this new world of data for giving us insight into a person's *umwelt*.

The second task, the ground truth problem, is arguably more difficult. For measures of mood or cognition, what constitutes ground truth? Should we train algorithms to self-report scales, to diaries of activity and mood, or to clinical ratings? If we are limited to these measures, is the field of mobile sensing destined to be no better than the subjective tools we've been using for decades? Here the analytic tools may help. As Part II of this book makes clear, increasingly sophisticated analytic approaches may help us refine the signals so that they are more informative than traditional measures and ultimately may offer a new kind of ground truth. But mobile sensing data, in the near term, will be adjunctive and not replacements for more conventional measures, remembering that more objective measures are not inherently more valid measures.

These challenges of data collection and data analysis need to be put into the context of clinical need, as noted in Part III of this book. Beyond the importance of psychological research, we find ourselves in a mental health crisis with rising rates of suicide, drug-overdose deaths, and depression in youth. The world of mental health care is supported by dedicated professionals who generally work in a data-free zone, without objective data on what is happening outside of the clinic. They may ask about sleep, activity, social contact, and mood without any objective data on these highly quantitative variables. Imagine helping someone with diabetes without measuring blood sugar (now trackable with a continuous glucose monitor) or someone with hypertension without measuring blood pressure (now trackable with home monitoring systems).

To be clear, our mental health crisis is not caused by this data desert, but better measurement can be part of the solution. More than half of the population with a mental disorder are not in care. Remote monitoring can detect a problem and connect people to care. For those who receive care, diagnosis is largely based on subjective reports in a single visit. Remote monitoring can provide objective data on how someone is thinking, feeling, and behaving in the real world, leading to more precise diagnosis. And for those in treatment, there is a surprising absence of monitoring progress, what the field calls "measurement-based care." There is a saying in business that we can't manage what we can't measure. For mental health care to begin to resolve the mental health crisis, we will need to bake measurement into all aspects of care. Mobile sensing can help to solve this data desert passively, ecologically, and continuously, at scale.

I stress this clinical need and the promise of remote sensing because we seem to be in a world in which worries about perils can stifle the promises of innovation. Yes, we must be mindful of privacy and data provenance. We need to build "with," not just "for," users. Transparency, integrity, and equity are fundamental concerns and essential for success. But in order for these concerns to be welcomed with creative and compassionate solutions, they must not become threats to the overall enterprise of using innovation to solve a public health crisis. We must remember that we face a formidable mental health challenge, which can be solved only via innovations like mobile sensing.

Will this happen in our lifetimes? Bill Gates famously noted, "We always overestimate the change that will occur in the next 2 years and underestimate the change that will occur in the next 10." With recent advances in sensor technology, artificial intelligence, and image analysis, we may be closer than we think. This timely volume provides a comprehensive picture of just how close we are and what remains to be done.

Thomas Insel, MD
Executive Chair, Vanna Health
Former Director (2002–2015), National Institute of Mental Health
Author of *Healing: Our Path from Mental Illness to Mental Health*

CHAPTER 1

How to Conduct Mobile Sensing Research

Gabriella M. Harari, Serena Soh, and Lara Kroencke

CHAPTER OVERVIEW

Mobile sensing is a methodological approach that leverages digital devices and platforms to collect data about human behavior. This chapter provides a starting point for researchers interested in conducting mobile sensing research in psychological science by describing how to conduct sensing studies with smartphones. First, we consider a series of questions that will help determine whether mobile sensing is the right methodological approach for a given study, set of research questions, and target sample of research participants. Next, we review a series of considerations that will help shape the specific study implementation, such as the resources available, the platform used for data collection, and some of the basic features of the study design (e.g., study duration, sampling rate, strategies for participant engagement, ethical considerations). Finally, we discuss some recommended practices for data monitoring, data cleaning, and data analysis, while highlighting the need for standardized guidelines and best practices for conducting mobile sensing research.

Introduction

Mobile sensing is a methodological approach that leverages digital devices and platforms to collect data about human behavior. Mobile sensing is used in studies across a broad range of scientific disciplines (e.g., computer science and engineering, psychological science) to answer research questions in both technical and substantive domains. In the technical domain, mobile sensing research often focuses on software development or activity recognition in an effort to improve the capabilities of sensing technologies. In the substantive domain, mobile sensing research often focuses on assessing behaviors and/or environments to understand people's daily lives and psychological experiences.

The goal of this chapter is to provide a starting point for researchers interested in conducting mobile sensing research in psychological science. Our aim here is to provide a roadmap for those who are considering or preparing to launch a mobile sensing study by describing how to conduct sensing studies with smartphones in particular. We focus on smartphones because they are the prototype mobile sensing device and the one most commonly used in mobile sensing research to date. However, many of the considerations outlined here also apply to the design of studies that use other sensing technologies to collect sensing data from participants' wearables (e.g., smartwatches, fitness trackers) and smart home appliances (e.g., smart speakers).

First, we consider a series of questions that will help determine whether mobile sensing is the right methodological approach for a given study, set of research questions, and target sample of research participants. Next, we review a series of considerations that will help shape the specific study implementation, such as the resources available, the platform used for data collection, and some of the basic features of the study design (e.g., study duration, sampling rate, strategies for participant engagement). Finally, we discuss some recommended practices for data monitoring, data cleaning, and data analysis. Overall, this chapter lays the foundation for the more advanced chapters in Part II ("Mobile Sensors: Technological Know-How and Methodological How-To") and Part III ("Analysis of Mobile Sensing Data") by outlining the basic steps involved in conducting mobile sensing research. Figure 1.1 provides an overview of the key steps and considerations that shape mobile sensing studies.

Questions to Consider Before You Get Started

Before getting started with mobile sensing research, it is helpful to consider a series of conceptual questions to determine whether mobile sensing is the best or "right" approach for a given study. As with any method, there are several benefits and costs associated with adoption of mobile sensing in research studies. The benefits of adopting mobile sensing primarily stem from the potential to collect large-scale, fine-grained, real-world naturalistic observations of people's behaviors and environments, and to a lesser extent of people's verbalized thoughts and feelings. This window into the daily lives of research participants provides an unprecedented view that is unparalleled when compared to other methodologies. The costs of adopting mobile sensing stem from the logistical (e.g., resources available) and practical hurdles (e.g., analyzing intensive repeated measures data) that must be overcome to successfully design and conduct a mobile sensing study. Whether the benefits outweigh the costs for any given study will largely depend on the research questions one hopes to address and the characteristics of the target population one hopes to study.

What Are the Research Questions and Target Variables?

The research questions one hopes to address and the phenomenon of interest are two key factors that can help determine whether mobile sensing methods are appropriate. Generally, research questions that have a temporal component and are focused on understanding

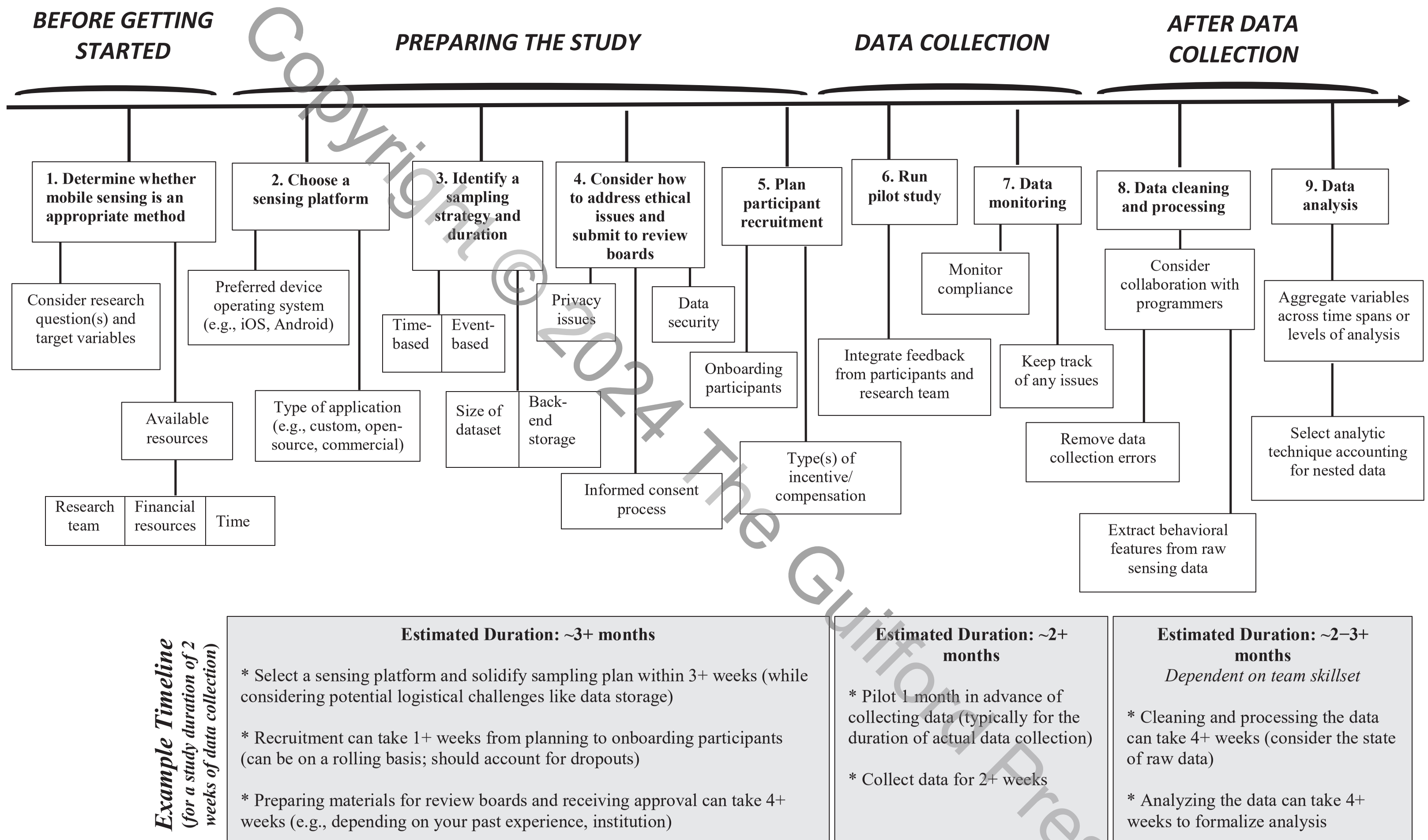

FIGURE 1.1. Overview of steps to consider in mobile sensing research.

some phenomenon over varying units of time (e.g., momentary, hourly, daily, weekly) are most suitable for mobile sensing study design. In addition, any questions about the degree to which people engage in behavior (e.g., frequency or duration of social interactions) are well suited to mobile sensing study design, whereas, at the time of this writing, research questions focused on more subjective aspects of behavior (e.g., quality of social interactions) are more challenging to address with mobile sensing studies. For example, several studies have focused on understanding the behavioral factors associated with college student well-being and academic performance during the academic term (e.g., Doryab et al., 2019; Wang et al., 2014, 2018; Wang, Harari, Hao, Zhou, & Campbell, 2015). In such studies, mobile sensing methods are well suited to addressing the research questions because they permit objective assessments of behaviors that are known to shape well-being and performance, such as the degree to which students engage in physical activity and social interactions, and exhibit certain sleeping patterns. Moreover, the studies benefit from the fact that continuous data are collected to measure the behaviors of interest. This permits the research team to aggregate the timestamped data in different ways and allows for multiple investigations of the research question using different approaches and analytic techniques to obtain a more complete understanding of the phenomena of interest. For example, some research studies have focused on a broad array of student behavior (e.g., physical activity, conversations, studying, partying) at different times of day and across entire academic terms to understand the factors associated with student well-being and academic performance (Wang et al., 2014, 2015). In contrast, other studies focused more narrowly on specific behaviors, such as social behavior (Harari, Müller, Aung, & Rentfrow, 2017; Harari, Müller, Stachl, et al., 2020) or mobility behavior (Müller, Peters, Matz, Wang, & Harari, 2020; Saeb, Lattie, Schueller, Kording, & Mohr, 2016). These examples highlight the opportunities introduced by using mobile sensing for answering research questions about human behavior over time. But it is worth noting that these studies focused on quantified estimates of the behaviors of interest and did not assess qualitative information about the behaviors observed (e.g., quality of social interactions or sleep).

Research questions with a temporal component also include research questions about dynamic intraindividual processes (Kuper, Modersitzki, Phan, & Rauthmann, 2021). For instance, researchers might examine how social behaviors are related to well-being states on the within-person level (i.e., whether individuals feel better after engaging in a conversation compared to how they normally feel) and individual differences therein. These within-person dynamics can best be investigated if the same individuals are observed repeatedly over time, which is typically the case in mobile sensing research. However, mobile sensing studies need not be solely focused on intraindividual processes.

Another area of opportunity presented by mobile sensing data is in understanding and objectively assessing interindividual differences, such as people's characteristic patterns of behaving over time (i.e., dispositional tendencies; Buss & Craik, 1980). If collected over long periods of time in which many types of situations are encountered, researchers can obtain estimates of people's behavioral tendencies by aggregating continuous sensing data at the within-person level over many days, weeks, or months for use in analyses at the between-person level. One point of caution with regard to deriving estimates of behavioral tendencies is that the research team should consider the implicit assumption that participants experienced a representative sampling of situations during the data collection period (e.g., weak and strong situations; Blum, Rauthmann, Göllner, Lischetzke, & Schmitt, 2018). For example, sensing studies conducted during the COVID-19 pandemic (e.g., Huckins et al., 2020) likely reflect a strong situational effect on social behavior that could affect behavioral estimates of face-to-face interaction and computer-mediated communication. These sensed behavioral tendencies can be used in place of self-reported behavioral tendencies to obtain objective estimates that quantify how a person actually tends to socialize, be physically active, and engage in various daily life activities over time. In past studies adopting this approach, the behavioral tendencies derived from sensing data have been examined in relation to self-reported personality traits (e.g., conversation, calling, texting, and app use tendencies; Harari, Müller, Stachl, et al., 2020; Stachl et al., 2017) and have even been used to predict self-reported personality traits alongside other sensing features (e.g., Mønsted, Mollgaard, & Mathiesen, 2018; Stachl et al., 2020). Chapter 20 provides a review of personality research in this domain.

In terms of target variables of interest, mobile sensing studies can provide information about people's inferred thoughts and feelings, as well as their observed behaviors and surrounding environments. However, they are best suited to providing objective assessments of behavioral and environmental information that can reflect the surrounding situation. The behavioral information that can be obtained from mobile sensing studies includes measures of human movement from accelerometers and Global Positioning System (GPS) data (e.g., physical activity, mobility patterns; see Chapters 4 and 5), social interactions from phone usage data (e.g., call and short messaging service [SMS] logs and app use logs; see Chapters 7 and 8), and various daily activities that are often measured in time use studies (e.g., some of which can be sensed like eating, sleeping, playing games, and listening to music; Harari, Müller, Mishra, et al., 2017; Sonnenberg, Riediger, Wrzus, & Wagner, 2012; White & Dolan, 2009). The environmental information that can be obtained from mobile sensing studies includes measures of ambience (e.g., light, noise, temperature), location (e.g., indoor vs. outdoors, places visited), and proximity to others (e.g., isolation vs. co-location; Harari, Müller, & Gosling, 2020). People's thoughts and feelings can also be inferred to some extent using sensing data, primarily by relying on verbal behavior collected from language data from social media (see Chapter 9) or audio data collected from microphones (see Chapter 10). But given that thoughts and feelings are inherently subjective phenomena, self-report methods may be a more effective and/ or convenient assessment approach for research focused on such constructs. Table 1.1 provides an overview of the different target variables of interest that can be derived from sensing data and the data sources needed to obtain them.

Who Are the Research Participants?

Another factor to consider when deciding whether to adopt mobile sensing as a data collection method for one's study is the target research population one plans to recruit. Much of the first wave of mobile sensing research was conducted with college-age young adults, with the aim of understanding the behaviors that shape their well-being. Targeting young adults as research participants in mobile sensing studies comes with several conveniences—they are generally readily available on university campuses where the research is being conducted, they are tech-savvy and already own smartphones, and they may be interested in participating in studies that collect data from their digital devices. For example, one study of student motivations to self-track showed that young adults

TABLE 1.1. Overview of Types of Data in Mobile Sensing Research

Data types	Description	Type of information assessed: Thoughts	Feelings	Behaviors	Environment
Mobile sensors					
Accelerometer	Orients the phone display horizontally or vertically; can record duration and degree of physical activity or movement			✓	✓
Bluetooth radio (BT)	Allows the phone to exchange data with other BT-enabled devices; can record the number of unique and repeated interaction partners and devices and co-located individuals			✓	✓
Global Positioning System (GPS) scans	Obtains the phone location from satellites; can record latitude and longitude coordinates			✓	✓
Light sensor	Monitors the brightness of the environment to adjust phone display; can record degree of ambient light or darkness			✓	✓
Microphone	Permits audio for calls; can record duration and frequency of conversations, degree of ambient silence or noise	✓	✓	✓	✓
Wi-Fi scans	Permits the phone to connect to a wireless network; can record location information based on the Wi-Fi network and crowds via the number of unique scans			✓	✓
Other types of data					
Cameras	Records images or video; can take pictures or videos periodically or semicontinuously		✓	✓	✓
Phone use logs	Records usage patterns such as notifications			✓	
App use logs	Records social interactions, entertainment, information-seeking behavior			✓	
Language data	Obtained from text data collected from the keyboard	✓	✓	✓	

Note. The first two columns of this table are adapted from tables presented in Harari et al. (2016) and Harari, Stachl, et al. (2021).

were interested in collecting data from their digital devices (e.g., smartphones, wearables) to improve their productivity and well-being, monitor their mood and daily activities, or improve their social lives (Harari, Müller, Mishra, et al., 2017).

Of course, not all research questions are about the lives of young people or about those young people who happen to be enrolled in universities. In such cases, more thought may need to be given as to how to go about recruiting and incentivizing the target group to participate in the study (see the section "How to Recruit and Incentivize Participants" later in this chapter).

Preparing a Mobile Sensing Study

Having determined that mobile sensing is the right methodology for your research questions and study, the next step is to consider a series of logistical issues that will help shape the design of the study. Mobile sensing studies are generally time and resource intensive, longitudinal in nature, and require careful thought to decisions that can affect the success of the study. Next, we outline how the resources one has available can shape subsequent decisions regarding the key features of the study design, such as the mobile sensing platform used for data collection and whether participants are engaged with the study. Ultimately, the logistical considerations and design decisions made at this step in the research planning will affect the quality of the resulting dataset.

What Resources Are Available?

The resources one has at hand to support the launch and completion of the study are a critical factor in study planning. Three main resources to consider are (1) the individual members and skillsets of the research team, (2) the financial resources available to support the study, and (3) the amount of time available to conduct the research.

The research team is a crucial factor in study planning for mobile sensing studies. The composition of the team and individual skills each member brings to the study will determine how responsibilities are distributed throughout the study period. In general, every sensing study involves several components that require oversight (sometimes simultaneously) and iteratively inform one another (e.g., pilot testing, data monitoring, participant interaction, data processing and analysis), making such studies nearly impossible to conduct by an individual alone. Sensing studies are a team effort, but whether that team is composed of individual students and research assistants or hired staff is a decision to be made early on in the study planning. Students and research assistants may be more motivated and invested in the study success given their likely involvement in the research planning process. However, if accountability is necessary, then hired staff may be a more reliable source of research support. Ultimately, this decision is contingent on the resources available.

In terms of skillsets, it is helpful to have team members who are familiar with the technical aspects of the sensing software being used (whether it be a custom, open-source, or commercial sensing application) and who are experienced in data science and programming to facilitate handling large-scale datasets. In addition, it is important to encourage open communication among the members of the research team throughout the study planning and data collection stages (e.g., via weekly meetings and/or other forms of synchronous and asynchronous interaction).

The financial resources available to help support the study are also important considerations when designing a sensing study. The amount of funding available can influence many of the decisions that must be made during study planning, such as study duration, number of participants to recruit, and type of sensing software used for data collection. For example, the study duration influences the amount of funding needed to pay staff (e.g., graduate students or research assistants hired to work on the project) and the amount of data that is collected, although the latter also depends on the number of participants recruited, the number of sensing data types collected, and the sampling frequency used during data collection. Generally, a study that runs for 1 week and only collects metadata from phone logs (e.g., calls, SMS, and app usage) is going to be less costly than a study that runs for 1 month and frequently collects raw sensor data (e.g., accelerometer, GPS). This is, in part, due to the storage requirements for such data, which drive up costs during data collection and subsequent analyses. The number of participants recruited will also affect the amount of funding needed if individuals are being financially compensated for their participation (see the section "How to Recruit and Incentivize Participants?" for alternative types of compensation). In addition, the decision to use a custom application (specifically developed for the study) or an open-source app (configured based on freely available software) may be a reasonable solution for research teams with the funds to hire people who can handle the more technical aspects of managing sensing software. Using custom or open-source software can permit more flexibility in that features can be customized to the needs of a given study, but this approach simultaneously introduces a great deal of technical complexity and requires more time for preparing and piloting the study to ensure the software is working as it should. Similarly, the decision to use a commercial app may come down to whether one can afford the expenses associated with running a sensing study with a given company. Several commercial sensing apps are available on the market, with each company naturally offering different rates for their services and having their own expenses to consider in providing their services. Some companies charge researchers based on specific study design characteristics, while others charge a flat service fee based on a subscription model (for a brief discussion of academic vs. corporate sensing research, see Chapter 33). In our own work, we have seen commercial companies quote anywhere from several hundred (e.g., ~$500 for a 2-day study collecting experience sampling and GPS data from 200 participants) to tens of thousands of U.S. dollars for sensing studies (e.g., ~$25,000 for a 4-week study collecting experience sampling reports and a full suite of many different types of sensing data from 1,000 participants). Beyond the study duration, the types of data collected and the sampling frequency can also affect the cost of running a study with a commercial company. So, given the variation in pricing we have observed in working with commercial companies, we generally encourage researchers interested in using a commercial app to speak with representatives of several companies to get estimated quotes for the cost of running a study that meets their desired specifications. To illustrate these points with more concrete examples, in Table 1.2 we briefly summarize our recent experiences and approach to conducting two different mobile sensing studies.

Another main resource required to effectively conduct a mobile sensing study is time (see Figure 1.1 for example estimates). Running a mobile sensing study (with any team and set of financial resources) will involve an intensive time commitment during the various stages of the study, from design to data collection to analysis. Thoughtful planning and discussion during the initial stages of the study will be required when the research team is deciding on the study design characteristics, testing and selecting platforms, and preparing materials for ethical review boards. Once the study is designed, the data collection stage is also demanding and can easily become a full-time job for individual members of the research team when accounting for the data monitoring and participant interactions required to ensure high data quality. So, it can be helpful for one or more team members to take the lead on different parts of the study. For example, one person might be responsible for running a pilot study with the research team to test the sensing software before the study launches, another person might be responsible for communicating with and onboarding participants during the study, while another person might be

TABLE 1.2. Study Design Considerations and Examples from Recent Mobile Sensing Studies

	Study names	
Considerations	**COVID-19 Smartphone Sensing Study (Talaifar et al., 2021)**	**Coping with Corona Project (Back et al., 2021)**
Study duration	3 weeks	4 weeks
Recruitment process	Through an online participant recruitment platform (*Prolific*) and university psychology course	Through a university psychology course
Number of participants	300+ students and adults	1,000+ students
Compensation	Course credit or monetary compensation ($10/week) and weekly feedback reports	Course credit and weekly feedback reports
Sensing software	Open-source app (*Beiwe*)	Commercial App (*Ksana Health*)
Sensing data collected	Accelerometer, battery state, Bluetooth, GPS, gyroscope, microphone, phone use logs, screen time, Wi-Fi	Accelerometer, ambient light, battery state, GPS, music, phone use logs
Self-reported data collected	Presurvey; two experience sampling surveys per day at set times; daily audio clip submissions; weekly app usage screenshots	Presurvey; eight experience sampling surveys per day at random times; postsurvey
Members of the research team responsible for data collection	Professors (2); doctoral students (3); undergraduate research assistants (3)	Professor (1); postdoctoral scholar (1); doctoral students (2); undergraduate research assistants (5)
Total cost	~$6,000 (mainly from participant compensation cost and recruitment platform fees)	~$30,000 (mainly from data collection platform fees)

responsible for monitoring the quality of the incoming data during and after the study. Of course, many of the tasks required to efficiently design and conduct sensing studies do ultimately require a collaborative effort. But we have found that many research teams are able to efficiently conduct studies with this kind of delegation of responsibility, so that there is a point of contact for troubleshooting issues that may arise with each aspect of the study.

How to Select a Mobile Sensing Platform?

The selection of a specific mobile sensing platform to use for data collection involves two key factors—the preferred device operating system (e.g., iOS [internet operating system], Android) and the type of application (e.g., custom, open-source, or commercial). The operating system and application selection should be determined based on considerations about the target research participants, the kinds of data needed for the study, and the resources available to the research team.

The selection of operating systems is consequential in that it shapes who can participate in the study and the kinds of sensing data that can be collected. As of 2021, Android and iOS jointly control approximately 99% of the global market share (Statista, 2021);

we therefore limit our discussion to these two mobile operating systems. It is also worth noting that the vast majority of sensing studies to date use applications that run on the iOS and/or Android phones. If participants are expected to use their own smartphones during the study, the research team must also consider the type of operating systems most used by their target sample. Past work has found that iOS users tend to have higher education levels, compared to Android users (Götz, Stieger, & Reips, 2017). But this demographic difference may not necessarily hold in all countries. In fact, Android phones are the most widely used phones around the world, having about a 72% share of the mobile operating system (OS) market (Statista, 2021).

The operating system also influences the kinds of data that can be collected by the sensing application. Generally speaking, iOS is more restrictive than Android in terms of the breadth and granularity of data sources that can be collected. This is in part due to the way that the two OS's allow third-party apps to access and collect data from the user's device. For example, third-party apps on iOS phones are not permitted to access the user's application usage logs at the time of this writing, but these sources of data can be accessed on Android phones (see Chapter 8 for more information about collecting app use data). So if a sensing study is designed to answer questions about the kinds of apps people use, the research team will need to identify a sensing platform that runs on Android phones and focus their recruitment efforts on participants who own Android phones. However, some data sources are commonly collected across both iOS and Android operating systems. These common sources of sensing data include accelerometer sensor data and activity classifications (e.g., stationary, walking, running), as well as GPS data.

The type of application used is also consequential because different sensing applications require different levels of support from the research team. A custom application is one that is designed specifically for and by a research team, and it is typically used in collaboration with computer scientists (e.g., EmotionSense, StudentLife; Rachuri et al., 2010; Wang et al., 2014). An open-source sensing application, such as AWARE[1] (Ferreira, Kostakos, & Dey, 2015) and Beiwe[2] (Torous, Kiang, Lorme, & Onnela, 2016), is one that is freely available for use by researchers. To effectively conduct a mobile sensing study with a custom or open-source application requires technical knowledge about how the sensing software operates. This is because if and when issues arise during data collection, someone on the research team needs to be able to troubleshoot and find a solution to address the issue. In contrast, a commercial sensing application is one that is operated and maintained for profit by a company (e.g., Ethica Data, Ksana Health). Conducting a mobile sensing study with a commercial application requires financial resources, but the benefits can outweigh the costs if the research team is not particularly interested in, skilled, or cares to be responsible for the technical details of how sensing systems operate.

How to Decide on a Sampling Strategy and Study Duration?

When selecting a sampling strategy, researchers must take into account many of the considerations introduced thus far, such as research questions, target variables, populations of interest, and available resources. Sampling strategies in mobile sensing often take the form of time-based sampling, such as continuous or periodic, or event-based data collection. Continuous and periodic sampling refers to schedules that collect data consistently at fixed times or within specifically set intervals, while event-based sampling refers to schedules that collect data contingent on the occurrence of certain events.

Time-based strategies include continuous and periodic sampling, which are often used in mobile sensing research. While continuous collection provides researchers with a wealth of data leading up to, during, and after the occurrence of the phenomenon being studied, periodic sampling enables researchers to decide how often and at what intervals the data are to be collected, depending on the objectives of the study or research question. Although the frequent and consistent nature of mobile sensing methods is considered to be one of its primary benefits, continuous sampling is not necessarily the best option for every study. For example, in the case of GPS data, sampling continuously (e.g., every minute) would lead to large datasets, challenges in data storage, and additional inconvenience to participants due to faster battery drainage. Moreover, participants' locations may not change very frequently during certain hours (e.g., during the work day if they are employed), which means that continuous sampling could result in obtaining redundant information. Rather, collecting GPS data every set interval of minutes within an hour (e.g., every 10 minutes) via periodic sampling may be a more appropriate and appealing option for both researchers and participants. For such reasons, studies like one examining the behavioral trends of college students through smartphones opt to use periodic GPS samples when computing outdoor mobility such as traveled distances (Wang et al., 2014).

Another strategy used in mobile sensing studies is event-based sampling where data collection is triggered by a predefined event. This strategy is most appropriate when examining specific phenomena that do not take place at regularly timed intervals, and it requires researchers to define the events that trigger data collection beforehand. Researchers often apply this strategy when studying smartphone use behaviors through metadata logs, which record events as they occur (e.g., a push notification is logged when it is received; calls and texts are logged as they are made or received). This sampling strategy is also commonly used when collecting movement or location-related data. By setting the events to be significant changes in GPS, the accelerometer, or the Wi-Fi network, data collected in those instances enable researchers to focus on and identify significant patterns in either activity or location changes. For example, this strategy has been observed in a study in which smartphone sensing data were used to predict clinical depression and researchers programmed event-based sampling for iOS users to study location trends by setting distance filters (Farhan et al., 2016). Event-based sampling helps ensure that data collection occurs at necessary times, but researchers must be prepared for the potential technological challenges that may arise. For example, the program may define events too generally and trigger data collection at unintended times, or technological glitches in the software may occur as other types of sampling are simpler in terms of data collection parameters. In the case of using location-related event-based sampling, unintended data collection may occur if data collection is triggered every time a participant is near the target location rather than when they are at the target location. Furthermore, GPS-based sampling can be difficult to program and implement, and additional testing of the location-contingent sampling will be required to identify potential bugs that might unexpectedly hinder data collection. This is why aspects such as ensuring the defined events are specific enough (e.g., precise distance filters for location-related events) and running pilot studies with a smaller pool of participants are especially important to these studies.

Once a sampling strategy is decided on, an appropriate study duration should then be considered given that together they determine the eventual size of the dataset. While mobile sensing studies typically last weeks to months, certain main considerations must be kept in mind when deciding on the length of a study. For instance, researchers must select a length (and sampling frequency) that enables them to answer their research question in terms of whether it examines momentary, hourly, daily, or weekly behavioral trends. If the study revolves around understanding how smartphone use behavior relates to well-being at the momentary level, the study duration can be shorter than a similar study focused on this relation over longer periods of time (e.g., understanding how well-being changes over the academic term). Lastly, from a logistical standpoint, it is important to consider that the combination of sampling rate and study duration determines eventual dataset size and statistical power. Sampling strategies that lead to a high frequency of data collection paired with long study durations, for example, could pose challenges for storing, processing, and analyzing the datasets, which may require the research team to have more advanced technical skills for large-scale analysis. Nonetheless, high sampling frequencies and long study durations have the benefit of increasing the power of the statistical analyses conducted. For example, in a recent 4-week study with ~700 participants, we collected around 4 terabytes of data. Even with an experienced and dedicated research team, we have spent a great deal of effort and time deciding on and implementing a workflow regarding how to aggregate, process, and analyze the data. To this point, the decision on the length of the study should also be made while acknowledging research team bandwidth and resource limitations. Longer durations often require more work on the part of the researchers either in monitoring data collection or analyzing the data afterward, as well as additional resources whether it be the monetary compensation for participants or costs associated with storing, managing, and analyzing large datasets.

How to Address Ethical Issues?

Conducting mobile sensing research introduces a host of new ethical quandaries for the social scientist. How can one respect individual privacy while collecting mobile sensing data from personal devices? How can the data be managed in a secure fashion? How can the study plans be best communicated to ensure appropriate oversight by relevant ethical review boards? As illustrated in the sections above, a great deal of data can be collected that provides detailed information about a person's behaviors (and to some degree, psychological experiences) in context. This is exciting for scientific discovery, while simultaneously concerning with regard to its potential negative effects for the individual participants. In this section we outline some of the main considerations in the ethical domain for getting started with mobile sensing research. However, we point interested readers to Chapter 2 for more detailed discussion of privacy issues and Chapter 3 for discussion of ethical issues as they relate to transparency and reproducibility in this research area.

Privacy issues are one of the most salient ethical concerns with regard to mobile sensing research. This is because sensing methods permit the collection of fine-grained *personal data*, which refers to "any kind of log or sensor data that directly describes an individual" (Wiese, Das, Hong, & Zimmerman, 2017, p. 452). The effects of participation on the individual privacy of the participant depend in large part on (1) the perceptions and concerns of the participants, and (2) the design of the study and the data management and analysis plan established by the research team. With regard to the participants, it is important to consider that people may be uncertain about their privacy preferences and the consequences of their behavior (Acquisti, Brandimarte, & Loewenstein, 2015). For example, participants may be unaware or unsure about the kinds of information they are providing about themselves when they permit collection of GPS data (De Montjoye, Hidalgo, Verleysen, & Blondel, 2013) or metadata from phone logs (Mayer, Mutchler, & Mitchell, 2016), both of which have been shown to be quite revealing about people's everyday behaviors. With regard to the design, some factors to consider are the types of data being collected, the sampling frequency being adopted, and the format of the data when it is collected. Generally, collecting and analyzing raw data is more sensitive than collecting and analyzing processed data. For example, collecting the *content* of communications is obviously more intrusive of participant privacy than collecting information about the *frequency* of communications. Similarly, raw GPS data (i.e., latitude and longitude coordinates) do not appear particularly sensitive in their raw format, but with additional preprocessing a person's home or work location could be inferred. A more privacy-preserving way of storing such location information would be to store the data as a categorical variable labeling the place a person was in (e.g., indexing a person was "home" or at "work"). In contrast, a threat to participant privacy would occur if such information were stored as the real address of the person's home or workplace. Given that participants may find sensing methods to be potentially invasive, special attention should be paid to facilitating transparency about the data being collected, participant control over personal data, and generally treating informed consent as a process (e.g., Harari, 2020; Kreuter, Haas, Keusch, Bähr, & Trappmann, 2020; Nebeker et al., 2016).

Data security is another aspect of the data management and analysis plan that is important to consider. Ensuring data security in a given study will be somewhat contingent on where the study is taking place (e.g., the institution, country), but some practices are relevant to almost all sensing studies. For example, with regard to the data management and analysis plan, some factors to consider are the people who will have access to the collected data and the strategy for processing and analyzing the data—for instance, ensuring that only key research personnel have access to personally identifying information about participants and that safeguards such as using secure servers for data storage and analysis can minimize potential concerns on behalf of participants and ethical review boards. When submitting mobile sensing research for ethical board review, several key things should be reported to ensure transparency about the design and research plans. In particular, we recommend describing the types of sensing data being collected, the format of the data, the location of where the data are stored, and the personnel who will have access to the files.

How to Recruit and Incentivize Participants?

Participant recruitment and compliance largely depend on the perceived benefits and costs of taking part in the study from the perspectives of the participants as well as their ability to fully participate. Because the cost of participating in a mobile sensing study tends to seem higher than that of other studies and because technologies (e.g., smartphones, wearables) or services (e.g., reliable internet access) are required, incentivizing individuals to make participation more appealing and providing participants with everything they need to actively participate are key to the success of a given study. In general, participant recruitment tends to be more challenging as people typically have concerns regarding privacy, personal data collection, data security, and data storage practices (see Chapter 2). Nevertheless, past mobile sensing studies have successfully recruited research participants from the student population, the general adult and elderly populations (e.g., Rachuri et al., 2010; Röcke, Katana, Fillekes, Martin, & Weibel, 2018; Saeb et al., 2015;

Stieger et al., 2021), and clinical populations (e.g., individuals undergoing chemotherapy, or those diagnosed with schizophrenia or bipolar disorder; Ben-Zeev et al., 2017; Low, 2020; Low et al., 2017; Matthews et al., 2016; Wang et al., 2017). In some cases, additional steps were taken to recruit participants (e.g., from hospitals and treatment centers) and onboard study participants to orient them to the goals and procedure of the study.

Furthermore, as mobile sensing studies require technologies and services that are not accessible to everyone, recruiting participants from rural areas, low-income communities, or developing nations may prove more challenging. According to Pew Research, smartphone adoption is growing in countries around the world, but countries with advanced economies have higher rates of ownership (e.g., in South Korea, Australia, and France, 75–95% of adults own a smartphone), compared to countries with emerging economies (e.g., in India, Indonesia, and South Africa, 24–60% of adults own a smartphone; Silver, 2019). However, with some creative planning in advance of the study launch, there are several ways to work around such constraints. For example, participants can be provided with the devices they need to participate (e.g., smartphones, wearables) and/or the services required for data collection for the study duration period (e.g., data plan for their phone). Providing such devices and services ensures that participants have the basic technical requirements needed to effectively participate in the study. It also can be a way to recruit participants from populations that do not readily have such technologies available to them, and it may help to target non-WEIRD (Western, educated, industrialized, rich, and democratic) samples (Henrich, Heine, & Norenzayan, 2010).

Once the target participants have been recruited into the study, keeping them incentivized and engaged with the study is another factor to consider. Motivations for participating and types of incentives preferred will vary by individual, but past studies have used monetary compensation, university credit, feedback reports, and lottery systems with varying levels of success (Farhan et al., 2016; Harari, Müller, Mishra, et al., 2017; Wang et al., 2014). Given their longitudinal nature and tendency to span weeks or months in duration, many sensing studies suffer from attrition due to participants dropping out over time, which can have negative impacts on the resulting dataset. Additional research is needed to better understand which incentives are most effective in maintaining high compliance rates. However, findings thus far suggest that adjusting self-tracking goals to align with participants' motivations and providing personalized feedback reports as an incentive (in addition to other forms of compensation like course credit, money, or prize lotteries) may help with compliance (Harari, Müller, Mishra, et al., 2017).

To keep attrition rates low, researchers should also consider how to balance study length with participant incentives. The success of the study and data collection efforts are impacted by rates of participation, so research teams have tested out different methods of incentive dispersion to sustain participant interest over time. For example, incentives can be spread out over the duration of the study—every few days, weeks, after every completed task, or all at once poststudy completion (Farhan et al., 2016; Wang et al., 2014). In a smartphone sensing study conducted within the Coping with Corona project in the fall of 2020 and spring of 2021 (Back et al., 2021; described in more detail in Table 1.2), the sample of university students recruited to participate received weekly feedback reports on their psychological states and behavior tendencies based on their sensing data and experience sampling reports. Students also received course credit after participating in each of the three steps in the study (i.e., completing a presurvey, self-tracking for 2 weeks, and reflecting on the study experience in a postsurvey). In a second COVID-19

Smartphone Sensing Study (Talaifar et al., 2021), we used a combination of monetary compensation and feedback reports as incentives for adults recruited from the community, and course credit and feedback reports as incentives for university students. Because adult participants were recruited through an online participant recruitment site, payment disbursements occurred when an individual either decided to no longer participate in the study or at the end of the study. The amount of compensation was dependent on the amount of time the individual spent participating. Feedback reports were also shared with participants weekly and included personalized information on their psychological states and behaviors.

These motivations and incentives should be substantial enough to outweigh the potential burden of participating whether that burden be the need to follow data uploading protocols, deal with app crashes or bugs, and, in some cases, use another device. As is the case with any mobile sensing study, typically participants must consistently follow procedures such as connecting to Wi-Fi and charging one's device regularly to upload their data. Additionally, there is a high likelihood that crashes and bugs in the mobile sensing platforms will arise and require individuals to troubleshoot with the guidance of the research team. These events are generally unavoidable, though they may pose negligible to varying amounts of burden among individuals in the population of interest and influence their decision to continue with the study in different manners. Also, researchers may decide to provide participants with a preprogrammed sensing device (Wang et al., 2014) rather than have them download a mobile sensing app on their personal device. This choice has some benefits, such as greater involvement from participant groups who do not have access to smartphones and services, as well as a standardization in device models or software, which ensures that all participants have devices with the same sensors necessary for some studies. At the same time, having some participants carry around a device second to their personal one may add yet another burden for them and lead to less accurate and missing data (e.g., phone logs; Harari et al., 2016). As providing a device also becomes more difficult with resource limitations and large samples of participants, most research teams opt for having participants use their own device when possible.

Furthermore, participant recruitment and incentives depend heavily on the context and nature of the study, which is why researchers often conduct pilot studies as a smaller-scale, shorter experiment to gauge what works and what does not. For example, based on pilot study recruitment and compliance statistics, researchers have general insight into whether (1) the recruitment strategy is effective, (2) people would be interested in and willing to participate, and (3) the current incentives are adequate. This also provides an opportunity to identify technology-related issues that need immediate attention before involving a large sample of participants or that the research team should be prepared to help troubleshoot.

Recommendations During and After Data Collection

Once the mobile sensing study has been designed, the next set of recommendations is more practical and focuses on the steps involved in conducting the study (e.g., monitoring data quality) and working with the data collected (e.g., data cleaning, processing, and analysis). Next, we outline our key recommendations, but for more detailed information we point interested readers to our past work on this topic (see Harari et al., 2016).

How to Check Participant Compliance and Data Quality?

Data monitoring involves checking compliance and data quality throughout the study. It is particularly important in sensing studies due to the technical demands and the unique challenges of the study design. First, sensing data are typically collected passively (i.e., without participant engagement), so any irregularities might go unnoticed by participants. Second, sensing data are collected continuously (i.e., with a high sampling frequency over uninterrupted periods of time), so problems must be detected quickly to intervene before the data quality is compromised. Third, while there is no need for active engagement with sensing apps in order for them to collect data, there are certain requirements for the app to function properly. For instance, all participants who take part in the sensing study must have their phone turned on and carry their phone with them as often as possible. Moreover, participants are often required to charge their phones and are connected to Wi-Fi regularly so their data can be uploaded. Lastly, some operating systems close apps that run in the background for too long, so participants have to regularly interact with the app to keep it running. In sum, it is important to regularly check the incoming data and to remind participants of the app's requirements.

Data monitoring involves downloading the sensor data and calculating and visualizing summary statistics, such as rates of uploads to the server or number of hours uploaded per day (Harari et al., 2016). Ideally, summary statistics should be calculated separately per sensor, as there may be problems with particular data sources. Some commercial platforms (e.g., Ethica Data, Ksana Health) provide data monitoring dashboards, which display data visualizations to researchers. We recommend checking the incoming data repeatedly throughout the study (e.g., at the end of each day) and contacting participants with missing data.

When monitoring the uploaded data, it is crucial to keep track of any problems that arise during the study. We recommend creating a data monitoring spreadsheet to document any issues that occurred during data collection. A rigorous documentation of problems will help to describe the study procedures later. Moreover, it is a crucial prerequisite for data cleaning.

How to Clean and Process the Data?

Sensing data are typically messy and should be cleaned before analyses. The data cleaning step is sometimes the most difficult step in the analysis, but it is also one of the most important steps. The choice of data cleaning procedures and their ordering can significantly impact the results of further analyses. Therefore, researchers should not use arbitrary data cleaning procedures (e.g., removing outliers when they could be real values) but should carefully think about data cleaning decisions before any analyses are run, and ideally, all decisions should be preregistered when possible.

Different types of data collection errors can compromise the quality of the data. With technically demanding data collections, error often results from technical problems. For instance, the sensing app might crash, or specific sensors might not be working properly (e.g., the GPS signal might be distorted; Müller et al., 2022). Moreover, there may be a lot of missing data if participants turn off their phones or accidentally close the app.

Different techniques are available to identify data collection errors. Unfortunately, only a few guidelines for data cleaning exist, and the decisions will always depend on the unique conditions of the study. Some authors have provided lists of problems they noticed when cleaning their own data and have provided recommendations for how to deal with these problems. For instance, in past work we have recommended removing: inaccurate or unrealistic data points (e.g., when two events occur simultaneously that do not seem possible, such as being in two different locations that are physically far apart within a very short time span); data points with missing timestamps or observations; duplicated data points; outliers (e.g., values above or below three standard deviations from the mean); and days or participants with too little data (e.g., less than 15 hours of data for a given day, or participants with only 1 day of data; Harari, Vaid, et al., 2020; Müller et al., 2022). These papers include relevant R code that provides more information about how one might go about executing these steps. The chapters in Part II and Part III of this handbook should also prove valuable for thinking through data cleaning steps for different types of data and for different analytic techniques.

After data cleaning, the raw sensing data have to be processed before any analysis can be run. The most common data processing process is to extract behavioral features. Feature extraction involves computing psychologically meaningful variables that can be used in further analyses, such as extracting locations visited from GPS data. For instance, in GPS data, psychologically meaningful locations (e.g., an individual's home) are typically represented by many different latitude and longitude coordinates. To extract mobility features for future analyses, researchers first determine key locations for every participant by clustering data points that are in close proximity to each other (for relevant R packages, see Müller et al., 2022). Next, researchers can interpret the locations (e.g., the home is often defined as the cluster where participants spend most of their time during the night) and calculate mobility features, such as the time spent in different locations based on the timestamps (Müller et al., 2022).

As another example, metadata logs (e.g., calls and app usage logs) typically consist of a list of timestamped events, such as when an app is opened or when an incoming call is received. Based on the number of entries and the associated timestamps, researchers can calculate frequencies (e.g., how often participants open an app or receives a phone call) and durations of events (e.g., Harari, Müller, Stachl, et al., 2020). Depending on the research question at hand, the features can be computed for different time intervals (e.g., across days, times of the day, or days of the week). For instance, researchers may calculate the frequency of calls for a given day and then average across days to obtain an estimate representative of a person's typical daily social tendencies (Harari, Müller, Stachl, et al., 2020).

Data from different sensors sometimes have to be combined to derive more complex features that rely on different sources of information (e.g., engaging in conversations in specific places). Sensing data can also be merged with self-report data, such as experience sampling reports. For instance, researchers may use smartphone sensing to obtain objective information about a person's behaviors or situational context, and experience sampling to ask participants about their subjective thoughts or feelings (Harari, Stachl, Müller, & Gosling, 2021). A detailed overview of all available features is beyond the scope of this chapter. However, it should be noted that the datasets are often very large (up to several gigabytes per participant) and that feature extraction requires advanced programming and analytical skills. Therefore, we recommend that psychological researchers interested in working with the unprocessed, raw sensing data refer to the mobile sensing literature for guidance on how to extract the variables of interest. As a starting point, we direct readers to the Reproducible Analysis Pipeline for Data Streams (RAPIDS) website.[3] This comprehensive resource provides an overview of different features and the code needed to compute them.

How to Analyze the Data?

After data cleaning and feature extraction, the data have to be prepared for analysis. Often, researchers have to aggregate their variables across different time spans (e.g., hourly, daily, weekly level) or levels of analysis (e.g., within-person vs. between-person) to answer the research question at hand. After data aggregation, researchers should check the distributions and psychometric properties (e.g., reliability) of all variables and select an appropriate analytic technique.

Because intensive longitudinal datasets consist of repeated observations from the same individuals, the analysis approach has to account for the nested structure of the data. Nested data are often analyzed using multilevel modeling (MLM; also called hierarchical linear modeling or random coefficient modeling; Hox, Moerbeek, & van de Schoot, 2018; Snijders & Bosker, 2012). Multilevel growth curve models (Bolger & Laurenceau, 2013) are one of several techniques to model intraindividual changes in variables across time. By using multilevel growth curve models, researchers can examine how behaviors change across different time spans (e.g., hours of the day, days of the week, or weeks of the academic semester) and examine different forms of change (e.g., linear, curvilinear, discontinuous). Importantly, MLM allows researchers to describe both normative behavior trajectories (e.g., how social behaviors change across the academic semester on average) as well as interindividual differences in these trajectories (to what extent the change trajectories differ between people) and how they are related to other individual difference variables (e.g., whether the differences in trajectories are predicted by personality traits).

In addition to research questions about the effects of time, intensive longitudinal studies are suited for research questions that focus on relationships between momentary states or momentary states and situational variables. Here, MLM allows researchers to disentangle effects on different levels of the analysis (Enders & Tofighi, 2007; Hamaker & Muthén, 2019). Specifically, when multiple measurements are collected from the same individuals, it is possible to analyze effects on both the within- and between-person levels. Within-person effects capture how time-point specific deviations from a person's average tendency in one variable are related to similar deviations in another variable. For instance, in a study that repeatedly assessed individuals' social behaviors (via sensing) and their mood (via the experience sampling method [ESM]), researchers might examine whether a given individual feels better after engaging in a social interaction compared to how they normally feel. Within-person relationships are particularly important when the focus is on intraindividual dynamics and individual differences therein (Kuper et al., 2021).

In addition to within-person relationships, researchers can examine between-person differences in behavioral tendencies. Between-person effects are obtained by aggregating the continuous sensing data on the person level (e.g., how much a person socializes on average) and using the behavioral aggregate instead of a self-report variable in further analyses. These aggregates serve as more objective estimates of how a person actually tends to behave in their everyday lives (as opposed to how they perceive themselves to behave).

Beyond MLM, there are more advanced techniques such as dynamic structural equation modeling, dynamic network analysis, person-centered/ideographic modeling, and machine learning. We point interested readers to Part III of this book for more information on these techniques for mobile sensing research. No matter the analytic technique selected to answer one's research questions, thorough and clear reporting of the data cleaning, processing, and analysis decisions is crucial for enhancing transparency and reproducibility in mobile sensing research (see Chapter 3 for more details).

Conclusions

Mobile sensing holds much promise for improving naturalistic observation in psychological science. The first wave of research studies at the intersection of psychology and computer science has showcased what is possible using these methods. However, a main factor that seems to be impeding the widespread use of these methods in the field more broadly is the lack of know-how regarding the steps involved in conducting a mobile sensing study. This chapter aims to address this knowledge gap by providing a starting point for those interested in or getting ready to launch a sensing study. In the future, more work needs to be done in the field to develop standardized guidelines and best practices for conducting mobile sensing research.

Notes

1. https://awareframework.com.
2. www.beiwe.org.
3. www.rapids.science/1.6.

REFERENCES

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509–514.

Back, M., Bühner, M., Geukes, K., Gosling, S. D., Harari, G. M., Kroencke, L., et al. (2021, December 15). *Coping with Corona (CoCo).* osf.io/2tzp9.

Ben-Zeev, D., Brian, R., Wang, R., Wang, W., Campbell, A. T., Aung, M. S. H., et al. (2017). CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatric Rehabilitation Journal, 40*(3), 266–275.

Blum, G. S., Rauthmann, J. F., Göllner, R., Lischetzke, T., & Schmitt, M. (2018). The nonlinear interaction of person and situation (NIPS) model: Theory and empirical evidence. *European Journal of Personality, 32*(3), 286–305.

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research.* Guilford Press.

Buss, D. M., & Craik, K. H. (1980). The frequency concept of disposition: Dominance and prototypically dominant acts. *Journal of Personality, 48*(3), 379–392.

De Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports, 3*, 1–5.

Doryab, A., Villalba, D. K., Chikersal, P., Dutcher, J. M., Tumminia, M., Liu, X., et al. (2019). Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and fitbit data. *JMIR mHealth and uHealth, 7*(7), 1–19.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121–138.

Farhan, A. A., Yue, C., Morillo, R., Ware, S., Lu, J., Bi, J., et al. (2016). Behavior vs. introspection: Refining prediction of clinical depression via smartphone sensing data. *2016 IEEE Wireless Health (WH)*, pp. 1–8.

Ferreira, D., Kostakos, V., & Dey, A. K. (2015). AWARE: Mobile context instrumentation framework. *Frontiers in ICT, 2*, 1–9.

Götz, F. M., Stieger, S., & Reips, U.-D. (2017). Users of the main smartphone operating systems (iOS, Android) differ only little in personality. *PloS One, 12*(5), 1–18.

Hamaker, E. L., & Muthén, B. (2019). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods, 25*(3), 365–379.

Harari, G. M. (2020). A process-oriented approach to respecting privacy in the context of mobile phone tracking. *Current Opinion in Psychology, 31*, 141–147.

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science, 11*(6), 838–854.

Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences, 18*, 83–90.

Harari, G. M., Müller, S. R., & Gosling, S. D. (2020). Naturalistic assessment of situations using mobile sensing methods. In J. F. Rauthmann, R. A. Sherman, & D. C. Funder (Eds.), *The Oxford handbook of psychological situations* (pp. 299–311). Oxford University Press.

Harari, G. M., Müller, S. R., Mishra, V., Wang, R., Campbell, A. T., Rentfrow, P. J., et al. (2017). An evaluation of students' interest in and compliance with self-tracking methods: Recommendations for incentives based on three smartphone sensing studies. *Social Psychological and Personality Science, 8*(5), 479–492.

Harari, G. M., Müller, S. R., Stachl, C., Wang, R., Wang, W., Bühner, M., et al. (2020). Sensing sociability: Individual differences in young adults' conversation, calling, texting, and app use behaviors in daily life. *Journal of Personality and Social Psychology, 119*(1), 204–228.

Harari, G. M., Stachl, C., Müller, S. R., & Gosling, S. D. (2021). Mobile sensing for studying personality dynamics in daily life. In J. F. Rauthmann (Ed.), *The handbook of personality dynamics and processes* (pp. 763–790). Academic Press.

Harari, G. M., Vaid, S. S., Müller, S. R., Stachl, C., Marrero, Z., Schoedel, R., et al. (2020). Personality sensing for theory development and assessment in the digital age. *European Journal of Personality, 34*(5), 649–669.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences, 33*(2–3), 111–135.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.

Huckins, J. F., DaSilva, A. W., Wang, W., Hedlund, E., Rogers, C., Nepal, S. K., et al. (2020). Mental health and behavior of college students during the early phases of the COVID-19 pandemic: Longitudinal smartphone and ecological momentary assessment study. *Journal of Medical Internet Research, 22*(6), 1–13.

Kreuter, F., Haas, G.-C., Keusch, F., Bähr, S., & Trappmann, M. (2020). Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent. *Social Science Computer Review, 38*(5), 533–549.

Kuper, N., Modersitzki, N., Phan, L. V., & Rauthmann, J. F. (2021). The dynamics, processes, mechanisms, and functioning of personality: An overview of the field. *British Journal of Psychology, 112*(1), 1–51.

Low, C. A. (2020). Harnessing consumer smartphone and wearable sensors for clinical cancer research. *NPJ Digital Medicine, 3*, 1–7.

Low, C. A., Dey, A. K., Ferreira, D., Kamarck, T., Sun, W., Bae, S., et al. (2017). Estimation of symptom severity during chemotherapy from passively sensed data: Exploratory study. *Journal of Medical Internet Research, 19*(12), 1–11.

Matthews, M., Abdullah, S., Murnane, E., Voida, S., Choudhury, T., Gay, G., et al. (2016). Development and evaluation of a smartphone-based measure of social rhythms for bipolar disorder. *Assessment, 23*(4), 472–483.

Mayer, J., Mutchler, P., & Mitchell, J. C. (2016). Evaluating the privacy properties of telephone metadata. *Proceedings of the National Academy of Sciences, 113*(20), 5536–5541.

Mønsted, B., Mollgaard, A., & Mathiesen, J. (2018). Phone-based metric as a predictor for basic personality traits. *Journal of Research in Personality, 74*, 16–22.

Müller, S. R., Bayer, J. B., Ross, M. Q., Mount, J., Stachl, C., Harari, G. M., et al. (2022). Analyzing GPS data for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *5*(2), 1–11.

Müller, S. R., Peters, H., Matz, S. C., Wang, W., & Harari, G. M. (2020). Investigating the relationships between mobility behaviours and indicators of subjective well-being using smartphone-based experience sampling and GPS tracking. *European Journal of Personality, 34*(5), 714–732.

Nebeker, C., Lagare, T., Takemoto, M., Lewars, B., Crist, K., Bloss, C. S., et al. (2016). Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. *Translational Behavioral Medicine, 6*(4), 577–586.

Rachuri, K. K., Musolesi, M., Mascolo, C., Rentfrow, P. J., Longworth, C., & Aucinas, A. (2010). EmotionSense: A mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (pp. 281–290).

Röcke, C., Katana, M., Fillekes, M., Martin, M., & Weibel, R. (2018). Mobility, physical activity and social interactions in the daily lives of healthy older adults: The MOASIS project. *Innovation in Aging, 2*(S1), 274.

Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., & Mohr, D. C. (2016). The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ, 4*, 1–15.

Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., et al. (2015). Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research, 17*(7), 1–12.

Silver, L. (2019, February 5). *Smartphone ownership is growing rapidly around the world, but not always equally*. www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally

Snijders, T. A. B., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.

Sonnenberg, B., Riediger, M., Wrzus, C., & Wagner, G. G. (2012). Measuring time use in surveys - How valid are time use questions in surveys?: Concordance of survey and experience sampling measures. *Social Science Research, 41*(5), 1037–1052.

Stachl, C., Au, Q., Schoedel, R., Gosling, S. D., Harari, G. M., Buschek, D., et al. (2020). Predicting personality from patterns of behavior collected with smartphones. *Proceedings of the National Academy of Sciences, 117*(30), 17680–17687.

Stachl, C., Hilbert, S., Au, Q., Buschek, D., De Luca, A., Bischl, B., et al. (2017). Personality traits predict smartphone usage. *European Journal of Personality, 31*(6), 701–722.

Statista. (2021, February 8). *Mobile operating systems' market share worldwide from January 2021 to January 2022*. www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009

Stieger, M., Flückiger, C., Rüegger, D., Kowatsch, T., Roberts, B. W., & Allemand, M. (2021). Changing personality traits with the help of a digital personality change intervention. *Proceedings of the National Academy of Sciences, 118*(8), 1–9.

Talaifar, S., Harari, G. M., Marrero, Z., Stachl, C., Vaid, S., & Gosling, S. D. (2021). *Using smartphone sensing to describe behavior during the COVID-19 pandemic*. Presentation at the virtual annual meeting of the Society for Personality and Social Psychology.

Torous, J., Kiang, M. V., Lorme, J., & Onnela, J.-P. (2016). New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Mental Health, 3*(2), 1–17.

Wang, R., Chen, F., Chen, Z., Li, T., Harari, G. M., Tignor, S., et al. (2014). StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 3–14.

Wang, R., Harari, G., Hao, P., Zhou, X., & Campbell, A. T. (2015). SmartGPA: How smartphones can assess and predict academic performance of college students. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 295–306.

Wang, R., Wang, W., Aung, M. S., Ben-Zeev, D., Brian, R., Campbell, A. T., et al. (2017). Predicting symptom trajectories of schizophrenia using mobile sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1*(3), 1–24.

Wang, R., Wang, W., DaSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., et al. (2018). Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2*(1), 1–26.

White, M. P., & Dolan, P. (2009). Accounting for the richness of daily activities. *Psychological Science, 20*(8), 1000–1008.

Wiese, J., Das, S., Hong, J. I., & Zimmerman, J. (2017). Evolving the ecosystem of personal behavioral data. *Human–Computer Interaction, 32*(5–6), 447–510.