



Published in final edited form as:

Perspect Psychol Sci. 2016 November ; 11(6): 838–854. doi:10.1177/1745691616650285.

Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges

Gabriella M. Harari¹, Nicholas D. Lane^{2,3}, Rui Wang⁴, Benjamin S. Crosier⁵, Andrew T. Campbell⁴, and Samuel D. Gosling^{1,6}

¹Department of Psychology, The University of Texas at Austin ²Nokia Bell Labs, Cambridge, England ³Computer Science Department, University College London ⁴Department of Computer Science, Dartmouth College ⁵Center for Technology and Behavioral Health, Department of Biomedical Data Science, Dartmouth College ⁶Department of Psychology, University of Melbourne

Abstract

Smartphones now offer the promise of collecting behavioral data unobtrusively, in situ, as it unfolds in the course of daily life. Data can be collected from the onboard sensors and other phone logs embedded in today's off-the-shelf smartphone devices. These data permit fine-grained, continuous collection of people's social interactions (e.g., speaking rates in conversation, size of social groups, calls, and text messages), daily activities (e.g., physical activity and sleep), and mobility patterns (e.g., frequency and duration of time spent at various locations). In this article, we have drawn on the lessons from the first wave of smartphone-sensing research to highlight areas of opportunity for psychological research, present practical considerations for designing smartphone studies, and discuss the ongoing methodological and ethical challenges associated with research in this domain. It is our hope that these practical guidelines will facilitate the use of smartphones as a behavioral observation tool in psychological science.

Keywords

mobile sensing; smartphones; behavior; big data; research design

Nearly 60% of adults in the United States own smart-phones, with adoption rates increasing in countries around the world (Pew Research Center, 2014, 2015). These phones are sensor-rich, computationally powerful, and nearly constant companions to their owners, providing

Reprints and permissions: sagepub.com/journalsPermissions.nav

Corresponding Author: Gabriella M. Harari, Department of Psychology, The University of Texas at Austin, 1 University Station A8000, Austin, TX 78751. gabriella.harari@utexas.edu.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information may be found at <http://pps.sagepub.com/content/by/supplemental-data>

unparalleled access to people as they go about their daily lives (Lane et al., 2010). Moreover, smartphones can be used to query people about their subjective psychological states (via notifications to respond to survey questions). These features have paved the way for the use of smartphones as data-collection tools in psychological research (Gosling & Mason, 2015; Miller, 2012).

Researchers already have begun to experiment with smartphones as behavioral data-collection tools, and in the process, they have gained valuable experience in addressing the numerous practical challenges of undertaking successful studies. In this article, we have drawn lessons from the first generation of smartphone-sensing studies to offer researchers practical advice for implementing these methods. We first review the sensors available in today's off-the-shelf smartphones and point to some promising areas of opportunity for psychological research using smartphone-sensing methods. To facilitate research in this area, we present practical considerations for designing smartphone studies and discuss the ongoing methodological and ethical challenges associated with this kind of research. The Supplemental Materials include a more detailed technical discussion of the logistical setup needed for smartphone-sensing studies (see Supplements A and B).

Traditional Methods of Collecting Behavioral Data

In existing procedures for collecting data on behavior, researchers typically ask participants to estimate the frequency or duration of past or typical behaviors. For example, a person asked to report on sociability behaviors for a given time period might be asked, "How many people did you talk to?" (frequency), or "How many minutes did you spend in conversation?" (duration). However, these self-reporting procedures are associated with well-known biases, such as participants' lack of attention to critical behaviors, memory limitations, and socially desirable responding (Gosling, John, Craik, & Robins, 1998; Paulhus & Vazire, 2007). Other methods for estimating behaviors have focused on presenting participants with hypothetical scenarios or recording behaviors in contrived laboratory studies. Several commentators have lamented the widespread reliance on self-reports and artificial laboratory studies in psychological science, rather than on objective behaviors as they play out in the context of people's natural lives (e.g., Baumeister, Vohs, & Funder, 2007; Furr, 2009; Paulhus & Vazire, 2007; Reis & Gosling, 2010). However, for many decades, the existing methods for collecting behavioral data in the field have been difficult and time consuming to use and intrusive for the participants being observed (Craik, 2000). Consequently, as a discipline, psychology has a great deal of data on what people believe they do, derived from their self-reports, but little data on what people actually do, derived from direct observations of their daily behaviors (Baumeister et al., 2007).

Smartphone sensing methods are poised to address this gap in research by allowing researchers to collect records of naturalistic behavior relatively objectively and unobtrusively (Boase, 2013; Rachuri et al., 2010; Wrzus & Mehl, 2015). In doing so, these methods allow researchers to address some of the methodological shortcomings of retrospective self-reports and studies of behavior in artificial laboratory contexts (Baumeister et al., 2007; Funder, 2006; Furr, 2009; Paulhus & Vazire, 2007). Moreover, the rising adoption rates of smartphones across the world are set to help psychological researchers reach beyond

participants from WEIRD populations (i.e., Western, educated people from industrialized, rich, democratic countries; Henrich, Heine, & Norenzayan, 2010) to obtain more representative samples that produce generalizable findings about people's day-to-day behavioral tendencies. Together, these features mean that smart-phones have the potential to revolutionize how behavioral data are collected in psychological science (Gosling & Mason, 2015; Miller, 2012).

The Promise of Smartphone Sensing

Off-the-shelf smartphones already come equipped with the sensors needed to obtain a great deal of information about their owners' behavioral lifestyles. They routinely record sociability (who we interact with via calls, texts, and social media apps) and mobility behaviors (where we are via accelerometer, global positioning system [GPS], and WiFi) as part of their daily functioning. Smart-phone sensing methods make use of these behavioral records by implementing on-the-phone software apps that passively collect data from the native mobile sensors and system logs that come already embedded in the device. Table 1 presents an overview of the most common types of smartphone data, their function in the device, and a summary of the behaviors they have been used to infer in past research. Some of the most common sensors found in smartphones include the accelerometer, Bluetooth, GPS, light sensor, microphone, proximity sensor, and WiFi. Other types of smartphone data collected include call logs, short message service (SMS) logs, app-use logs, and battery-status logs.

Such smartphone data can be used to capture many behaviors, which we have organized here in terms of a framework derived from previous research on acoustic observations: social interactions, daily activities, and mobility patterns (Mehl, Gosling, & Pennebaker, 2006; Mehl & Pennebaker, 2003). Social-interaction behaviors include initiated and received communications (via call and text logs; Boase & Ling, 2013; Chittaranjan, Blom, & Gatica-Perez, 2011; Kobayashi & Boase, 2012), ambient conversation (via microphone; Lane et al., 2011; Rabbi, Ali, Choudhury, & Berke, 2011), speaking rates and turn taking in conversation (via microphone; Choudhury & Basu, 2004), and the size of in-person social groups (via Bluetooth scans; Chen et al., 2014). Daily activities include people's physical activity (via accelerometer; Miluzzo et al., 2008), sleeping patterns (via combination of light sensor and phone usage logs; Chen et al., 2013), and partying and studying habits (via combinations of GPS, microphone, and accelerometer; Wang, Harari, Hao, Zhou, & Campbell, 2015). Mobility patterns include people's duration of time spent in various places (like their home, gym, or local café), the frequency of visiting various places, the distance travelled in a given time period, and routines in mobility patterns (via GPS and WiFi scans; Farrahi & Gatica-Perez, 2008; Harari, Gosling, Wang, & Campbell, 2015; Wang et al., 2014).

It should be noted that smartphones are just one of the many mobile-sensing devices that can collect behavioral information with great ecological validity; other devices include wearable devices (e.g., smartwatches) and household items (e.g., smart thermometers). However, in light of their ubiquity and the fact that they come already equipped with numerous embedded sensors (Lane et al., 2010), smartphones are particularly well placed to address

many of the methodological challenges facing researchers in the field as they strive to make psychology a truly behavioral science (Miller, 2012).

Smartphone-sensing research is flourishing in the field of computer science but only recently has begun to enter the methodological toolkit for psychological researchers (Gosling & Mason, 2015; Miller, 2012; Wrzus & Mehl, 2015). Thus, there are many areas of opportunity for psychological researchers to use sensing methods to examine both new and existing research topics. Interdisciplinary research groups composed of psychologists and computer scientists already have incorporated sensing methods into studies of such varied topics as emotional variation in daily life (Rachuri et al., 2010; Sandstrom, Lathia, Mascolo, & Rentfrow, 2016), sleeping patterns and postures (Wrzus et al., 2012), and interpersonal behaviors in group settings (Mast, Gatica-Perez, Frauendorfer, Nguyen, & Choudhury, 2015). To help researchers think about how they might integrate sensing methods into their own research, we next present an illustrative range of three research domains (see Table 2 for a summary of these domains and suggested analytic techniques that could be used to explore them).

Describing behavioral lifestyle patterns over time

As a nominally behavioral science, the field of psychology has amassed surprisingly little knowledge about people's patterns of everyday behavior over time—their *behavioral lifestyles*. Sensing research, focused on longitudinal patterns of stability and change in behavioral lifestyles, can provide information about the behaviors associated with individual differences (e.g., demographic, personality, or well-being factors) and life stages (e.g., adolescence, adulthood, old age). Even a catalog of the basic behaviors in which people engage would provide a much-needed empirical foundation on which more sophisticated questions can be built. For instance, smart-phone data can be used to classify different types of behavioral lifestyles based on certain markers that characterize a person's or a group's behaviors over time. Such studies could be used to develop classification models to distinguish between-persons behavioral classes, such as the set of patterns that characterize a “working lifestyle” or a “student lifestyle,” based on social interaction (e.g., frequency of making and receiving SMS messages), daily activities (e.g., times of the day when sedentary vs. active), and mobility behaviors (e.g., regularity in mobility patterns to work or campus). It is likely that behavioral lifestyles also would emerge that describe within-person variations in behavior, such as activity patterns (e.g., people who show a highly sedentary lifestyle only on weekdays vs. those who show such a lifestyle throughout the week) or socializing patterns within a week (e.g., people who have a highly social lifestyle only on weekends vs. throughout the week). The identification of behavioral classes in this manner could aid in the development of interventions for those who deviate from normative or healthy behavioral lifestyles.

Predicting life outcomes and implementing mobile interventions

Clinicians and other health practitioners working with populations exhibiting problematic behaviors have long-recognized the opportunity for smartphone-based methods for improving health outcomes—an area of research known as *mobile health* (mHealth). A common goal for many mHealth studies is identification of behaviors associated with

positive and negative health outcomes so that behavior-change interventions may be designed and implemented on a large scale (Lathia et al., 2013).

Behavioral lifestyle data about social interactions, daily activities, and mobility patterns can be used to determine the key predictors of important life outcomes including physical health (e.g., heart disease or obesity), mental health (e.g., depression or anxiety), subjective well-being (e.g., mood or stress), and performance (e.g., academic or occupational). For instance, sensing methods are being used to describe the patterns of behavior and subjective experience associated with depressive symptoms (Saeb et al., 2015) and the behavioral and mood patterns associated with schizophrenia symptoms (Ben-Zeev et al., 2014). Other examples of mental health outcomes that sensing methods are particularly well placed to identify are the behavioral markers that precede manic or depressive episodes, alcohol or drug relapse among recovering addicts, and suicidal ideation or attempts. Researchers also have used sensing methods to build machine-learning models that use behavioral lifestyle data (e.g., sociability or studying trends) during an academic term to predict students' academic performance (as measured by grade point average) at the end of the term (Wang et al., 2015).

Descriptive research on the normative and nonnormative behavioral patterns of people's daily lives could point to the significant patterns indicating when an intervention might be delivered. For instance, researchers trying to predict a clinical episode in schizophrenic populations may find that the onset of an episode manifests via a change in a participant's daily social interactions, activities, or location patterns (Ben-Zeev et al., 2014). In other instances, such as researchers trying to predict relapse among recovering addicts, the presence of a problem behavior (e.g., location data showing that the participant prone to alcoholism is spending time in a bar) may be enough to trigger an intervention. In the same vein, for researchers trying to predict periods of severe depression among depressed populations, the absence of certain behaviors (e.g., not socializing with others or not leaving one's home) in a participant's behavioral records may indicate when an intervention is needed (e.g., Saeb et al., 2015).

These types of mHealth techniques hold much promise for increasing access to psychotherapy among diverse populations (Morris & Aguilera, 2012). Technically, these intervention strategies are possible now but are feasible only in small-scale controlled settings because constant monitoring of the incoming data is required for the interventions to be implemented effectively in real time. Moreover, to be truly effective, more descriptive research is needed from both normative and nonnormative populations. We expect the next few years to yield much data on behavioral patterns across a broad spectrum of psychological topics, paving the way for these large-scale but targeted and individualized interventions.

Examining social network systems

Social scientists long have been interested in social network systems because they provide a way to link micro- and macro-level processes within a larger social structure. For example, social network systems have been used to examine friendship groups (Eagle, Pentland, & Lazer, 2009), online social media interactions (Brown, Broderick, & Lee, 2007), and disease

transmission (Gardy et al., 2011). Traditionally, social networks have been studied with self-report methods; however, self-reported networks can provide data only on how people think their networks are structured, not how they are actually structured. Mental representations of networks can be biased by motives or memory and may not accurately depict actual behavior. Such biases were demonstrated when Bernard, Killworth, and Sailer (1980, 1982) asked members of social groups to report on their interactions and then compared these data with observational interaction data, finding that self-reports of communication patterns did not map onto actual behavioral communication patterns.

Smartphone-sensing methods can address this problem by providing a new way to collect and analyze social-interaction data. Instead of relying on recall, researchers can obtain actual communication records from many different phone-based sources. With participant consent, researchers can monitor call and SMS message logs for frequency, duration, and unique persons contacted in incoming and outgoing interactions (e.g., de Montjoye, Quoidbach, Robic, & Pentland, 2013) and can obtain information about online social networks by tapping into data from other applications installed on the phone (e.g., Contacts, Facebook, Twitter, or Gmail; Chittaranjan et al., 2011; LiKamWa, Liu, Lane, & Zhong, 2013). Moreover, researchers can also obtain data that serves as a proxy for estimating face-to-face interactions from mobile sensors such as Bluetooth and microphone data to infer when participants are with other people or engaged in conversation (e.g., Chen et al., 2014; Lane et al., 2011; Rabbi et al., 2011; Wang et al., 2014).

Unlike the social network information provided by recall, smartphone-based social structures do not depend on the limits of human memory. The direct assessment of social interactions with a sensing device combines the authenticity of a network built from recall with the accuracy of electronic assessment. These types of social interaction data (e.g., phone and SMS logs, Bluetooth, microphone, and social media app usage) can be combined with other data (e.g., self-report surveys) and synthesized with any number of other variables in highly descriptive models of behavior.

Practical Considerations for Making Key Design Decisions

The technology and software that permit smartphone-sensing research are changing so rapidly that it would be of little use to review or recommend specific products; however, there are a series of basic questions about the design that need to be addressed by anyone running a smartphone-sensing study. The answers to these questions will guide which smartphone devices and sensing software are adopted for the study, even as the capabilities of the specific products evolve. Therefore, to facilitate the use of smartphone-sensing methods, we next present a set of practical considerations for key design decisions that will need to be made in most smart phone-sensing studies. These tips are derived from our experiences implementing both small- and large-scale smartphone-sensing studies. Table 3 presents an overview of these decisions. Ultimately, of course, each of these design decisions will be guided by the research questions under study. We start by laying out the general structure of most smartphone-sensing studies.

How are smartphone-sensing systems set up?

Smartphone-sensing studies require the setup of a system that runs the app software and facilitates the collection, storage, and extraction of data during and after the study (see Table 4 for a summary of the features of a smart-phone-sensing study design). These smartphone-sensing systems consist of three main components: the front-end, the back-end, and the data-processing components.

The front-end component consists of the smartphone application software that is installed on the participant's phone (left panel of Fig. 1) and consists of the user interface with which the participant interacts with the app (e.g., to respond to survey notifications). The app software collects data by sampling from a series of sensors, apps, and phone logs. The back-end component consists of three major behind-the-scenes features that run on a server to facilitate data collection: the portal server, the participant manager, and the data storage feature (middle panel of Fig. 1). The portal server is the main node of the back-end component; it receives the data from the front-end component and checks it against the participant manager (which provides user authentication). The portal server stores the sensor data collected from apps in the data storage, which is typically a database that can handle very large data sets (e.g., MySQL, MongoDB). The database is a necessary feature of the back-end component because it allows researchers to query the data and, when necessary, apply transformations to the data to compute behavioral inferences from the sensor data (see "Developing behavioral measures from smartphone data" section). Any additional data collected during the study (e.g., pre- or postsurvey measures) can also be stored in the data storage.

The data-processing component consists of monitoring the data collection, preparing the sensor data for analysis, and formal analysis of the data (see right panel of Fig. 1). The monitoring of incoming data is a critical component of sensing-study design because the data are often being collected passively (i.e., automatically on the phone, without participant input) and somewhat continuously (e.g., every few hours or every few minutes), making it important that any problems with data collection are identified as they occur in real time. Generally, data-monitoring practices involve the extraction of the sensor data from the database in which they are stored and the application of several computer scripts (often written in Python or R) to the data to obtain summary statistics and visualizations of the data collected. Example data-monitoring tasks include, visualizing participation and attrition rates during the study, and estimating the amount of data collected as the study progresses. Such summaries are crucial indications of the application's performance and participants' engagement, permitting researchers to tweak the study design or contact participants about any observed gaps in data collection even as the project is underway. Readers interested in a more technical review of the sensing system should consult the Supplemental Materials (Supplement B).

Which device and sensing application should I use?

In smartphone-sensing studies, investigators typically deploy apps for devices running Android or iOS operating systems (OS). The decision to use either of these OSs will be influenced by several factors, including the sampling-rate constraints of the OS and whether

participants will be downloading the sensing app to their own device or be given a device to use for the study duration. As of March 2016, the Android OS permits third-party apps to sample from more sensors and system phone logs than apps running on iOS (for a detailed discussion of the pros and cons of these two OSs at the time of this writing, see Supplement A in the Supplemental Materials). Android also claims around 80% of the global market for smartphone devices, while iOS claims around 15% of the market (International Data Corp., 2015), suggesting that researchers conducting psychological studies that use Android devices will have access to more diverse and representative samples from populations around the world. However, the smartphone market in the United States is split close to evenly between Android and iOS users (Smith, 2013), and there are some demographic differences in use of the OSs (i.e., iOS users being of higher socioeconomic status), suggesting that U.S. sensing studies may have difficulty recruiting samples of only Android or iOS users if one OS is selected for the study design.

What types of smartphone data do I need for my study?—Some mobile-sensing apps vary in the breadth of sensors used, ranging from those that only collect data from one sensor (e.g., StressSense, in which microphone data were used to infer stress levels from features of a participant’s voice; Lu et al., 2012) to those that collect data from many sensors (e.g., StudentLife, in which accelerometer, Bluetooth, GPS, microphone, and WiFi were used to chart behaviors associated with well-being and performance; Wang et al., 2014). Some apps also have integrated sensing and ecological-momentary assessments (EMAs) as part of their data-collection process (e.g., Emotion Sense; Rachuri et al., 2010); such apps are useful for researchers that want to query participants about their subjective experiences while also collecting objective behavioral data from the smartphone. The right-hand column of Table 4 provides a list of some existing sensing applications for interested researchers.

Should participants use their own mobile device, or should I provide them with one?—One large benefit of participants using their own mobile devices is that participant recruitment can sample from diverse and larger populations if the application is made available publicly on app stores. For example, the Emotion Sense application (Rachuri et al., 2010) is available on Google Play, the store for Android applications, and has registered thousands of active users worldwide. An additional benefit to using participants’ own devices is that the behavioral data will have high fidelity and ecological validity because the data are collected from the participants’ primary device, which they already keep with them throughout the day (see Supplement A in the Supplemental Materials for a comparison of primary and secondary phone users in a study by Wang et al., 2014).

The main drawbacks to having participants use their own devices stem from the lack of standardization in the smartphone data collected, which is introduced when a mixture of devices and OSs are used. For example, the classifiers used to process the data (see “Developing behavioral measures from smartphone data” section) could introduce noise to the behavioral measures if they were developed for data collected from devices running one OS (e.g., Android), but are then used for data collected from devices running another OS (e.g., iOS). Such standardization issues even arise when a standard OS is used across devices made by different manufacturers or devices containing different makes of sensors (e.g.,

participants using different Android devices; Stisen et al., 2015). For this reason, in some smartphone-sensing studies, participants have been provided with mobile devices to use for the duration of the study so that all participants use the same device and OS (e.g., Wang et al., 2014); however, this approach may require participants to carry an extra phone with them and does not scale to studies with very large numbers of participants.

How long should my study run, and how often should I sample?

Smartphone-sensing studies tend to be longitudinal designs and may span from several hours to several months. Both the length of study and the sampling rate will play a key role in determining which app to use and the eventual size of the data set. For example, a study of daily fluctuations in activity would require (a) a longitudinal design of at least 1 week so that each day of the week would be represented and (b) a smartphone device and sensing app that permit semicontinuous collection of accelerometer sensor data.

Sampling rates vary from those that automatically monitor smartphone data continuously (e.g., every few minutes) to those that collect smartphone data only periodically, such as when a participant opens the app to respond to a survey notification. The sampling rate at which the smartphone data are collected has a big impact on the size of the resulting data set. Due to the continuous nature of data collection via smartphone, longer study durations with higher sampling rates will result in significantly larger data sets, with some reaching hundreds of gigabytes of sensor data. Thus, researchers must take care to ensure they have sufficient server space on the back-end component to handle such quantities of data, which are rare in conventional studies. High sampling rates can also result in the battery life of the device being drained rapidly, which can present problems for retaining participants who are likely to drop out of studies if the everyday use of their phone is impaired by the sensing application.

How do I obtain behavioral variables from the smartphone data?

Smartphone-sensing studies produce big data sets that require researchers to use some familiar and advanced techniques for processing data. It is possible that future sensing systems will eliminate the need for researchers to implement these processing techniques (e.g., by providing researchers with the desired behavioral variables already computed). However, to facilitate research in this domain, we present some current techniques for processing sensor data to ready it for more formal analyses. The techniques we review are used to create behavioral variables from sensor data, such as techniques for extracting behavioral inferences, inferring more complex behaviors, and preparing the data for analysis.

Extracting behavioral inferences—The application software used in the study typically determines the format of the sensor data. The app may store the data in one of two ways. The software may simply store the raw, unprocessed sensor data after collection of the data from the participant's phone—these data are termed *raw-sensor* data—or the software may process the sensor data before storage in order to make inferences about the participant's behavior—these data are termed *behavioral-inference* data. The main distinction between these two data formats is that the unprocessed raw-sensor data require an additional

processing step for researchers to obtain meaningful behavioral variables, whereas the behavioral-inference data are already processed to create variables that capture a behavior of interest. Thus, the extraction of behavioral inferences is typically the first step in processing the data for subsequent analyses because the sensor data need to be transformed into psychologically meaningful units that also lend themselves to further analyses.

Smartphone-sensing apps that store raw-sensor data typically generate large amounts of data (several gigabytes or more) that can be costly to store on the phone in terms of battery life and costly to transfer and store on a server. To illustrate, consider raw-sensor data collected from an accelerometer sensor. Raw accelerometer data consist of three values per sampled data point—an X coordinate, a Y coordinate, and a Z coordinate. These three coordinates are collected each time the sensor is sampled. In studies with continuous sampling rates (e.g., samples collected every few minutes), these data can quickly scale up and result in massive data sets that need to be processed before meaningful variables (e.g., behaviors like walking or running) can be obtained.

In contrast, apps that store behavioral-inference data do so by including the processing step within the software of the system itself. In doing so, the app runs classifiers on the phone in real time to convert the raw-sensor data to behavioral inferences before storing the behavioral-inference data. To illustrate, an app that collects behavioral-inference data from the accelerometer sensor would apply activity classifiers to the raw accelerometer data (X , Y , and Z coordinates), resulting in behavioral-inference data that might take the form of a 0 for stationary behavior, a 1 for walking, and a 2 for running. These behavioral-inference data (not the raw-sensor data) are stored and later processed further and analyzed by the researchers.

Psychologists may find behavioral-inference data to be more intuitive and easier to work with than raw-sensor data because of the interpretability and the smaller size of the behavioral-inference data sets. However, an advantage to collecting raw-sensor data is that such data contain a rich amount of sensor information, which can later be used for other behavioral inferences that are developed after the study period. Many classifiers have been developed to infer behaviors, and a review of the existing classifiers is beyond the scope of the current article. We recommend that researchers working with raw-sensor data consult with computer science collaborators and research articles published in computer science journals and conference proceedings (e.g., the Association for Computing Machinery's International Joint Conference on Pervasive and Ubiquitous Computing [UbiComp]; the International Conference on Mobile Systems, Applications, and Services [MobiSys]; and the Conference on Embedded Networked Sensor Systems [SenSys]) for guidance on selecting the appropriate classifiers to use to infer a given behavior of interest.

Combining different types of sensor data—The combination of two (or more) types of sensor data can produce more finely specified and context-specific behaviors. For instance, context-specific behavioral estimates can be obtained by binning behavioral inferences obtained from sensors (e.g., microphone, accelerometer, or Bluetooth) according to the user's physical context using GPS or WiFi data. This technique allows researchers to infer finely specified behaviors, including (among other things): talking with others in

different locations (e.g., home, campus, or work), degree of physical activity in different locations, and amount of time spent alone or with groups of people in different locations. The integration of location data in this manner paves the way for more fine-grained studies of behavior expression across situations (Harari et al., 2015).

For example, previous researchers have used combinations of sensor data to infer studying behavior among students during an academic term (Wang et al., 2015). To infer studying durations, investigators used combinations of GPS and WiFi data to determine whether the student was in a campus library or study area, microphone data to determine whether the environment was silent (not noisy or around people talking), and accelerometer data to determine whether the student's phone was stationary (and not being used; Wang et al., 2015). This combination of sensor-based behavioral estimates was used to infer the duration of time a student spent studying during the term. It is worth noting here that this approach to inferring studying behavior likely resulted in underestimation of the actual amount of studying in which the students engaged. That is, the sensor data combinations used in such a study may be a sufficient condition for inferring studying behavior, but they are clearly not a necessary condition (e.g., the students may have studied at home or at cafes or in noisy environments). Nonetheless, the average studying duration obtained from this behavior inference correlated .38 with students' academic performance at the end of the term (measured via their grade point averages; Wang et al., 2015), offering some evidence for the validity the complex behaviors inferred in this manner.

The application of algorithms to the sensor data can also produce estimates of more complex behaviors that are not easily captured from a single sensor. Complex behavioral estimates can be obtained by transforming several types of sensor data that capture lower-level behaviors into one higher-level behavioral inference via use of an algorithm designed for the task. An example of this approach is the algorithm developed to infer sleeping durations on the basis of several different types of sensor data (Chen et al., 2014); the types of data used by the algorithm included the current time (whether it is day or night time), the state of the ambient light sensor (whether the environment is light or dark), the phone logs (whether the phone is being used or not), the accelerometer (whether the phone is stationary or not), and the battery logs (whether the phone is charging or not). Using an algorithm that took into account the various states of these different types of data, the researchers were able to infer sleeping patterns that included the participant's time to bed and rise and sleep duration within ± 42 min (Chen et al., 2014). Another application for these more complex algorithms is in computing higher-level mobility patterns, such as variability of time spent in different locations, distance travelled in a given day, or the routineness of a person's mobility patterns (e.g., Farrahi & Gatica-Perez, 2008; Saeb et al., 2015). More psychological studies are needed to examine the convergent and external validities of such behavioral measures, but the initial studies are promising.

Combining sensor data with self-report data—The integration of self-report data with sensor data permits the researcher to supplement objective behavioral estimates with the participants' own reports of their experience. To illustrate this approach with an example, consider a researcher who is interested in how socializing behaviors change as a function of a person's situational context or internal state (e.g., mood or stress level). To study this

phenomenon, behavioral-inference data could be partitioned according to the participant's concurrent ecological-momentary assessment (EMA) reports (e.g., talking durations [obtained via microphone] when they report being with friends, at work, stressed out, or in a good mood).

Sensor data can also be used to trigger context-contingent EMAs (Pejovic, Lathia, Mascolo, & Musolesi, 2015). For example, when a person goes to a new place (obtained via GPS data), the app can deliver relevant EMA questions (e.g., "What is this place?", "Whom are you with?", or "What are you doing here?"). Context-contingent EMAs can also be triggered by phone use (e.g., EMAs triggered after the end of phone calls could ask about the participant's mood). Obviously, there are many possible ways to partition sensor data on the basis of participants' reported psychological experiences and many ways to deliver context-contingent EMAs. Researchers interested in deploying context-contingent designs will want to consider how this decision may affect the representation of the construct (or behavior) in the aggregated sensor data (e.g., Lathia, Rachuri, Mascolo, & Rentfrow, 2013). We expect psychological research that combines sensor data with self-reports to yield fine-grained descriptions of the behavioral antecedents and consequences of various psychological states.

Preparing the data for analysis—Once the behavioral-inference data are extracted, researchers may need to aggregate the data to the appropriate level or unit in time for their analyses. To create the aggregated variables (e.g., estimates of hourly or daily activity duration), investigators need to apply computer codes (e.g., Python or R scripts) to the processed behavioral-inference data to partition the data and aggregate them as needed. Consider, for instance, the aggregation of call log data, which might require the application of scripts to compute the duration of time spent on incoming or outgoing calls each day (by aggregating across the individual call durations in a given day). In a similar fashion, the aggregation of these data could also involve computing the frequency of incoming or outgoing calls each day. In general, the time frame selected for the data-aggregation step will be guided by the research questions and study design. For example, if the researchers are interested in how sociability is related to daily mood, the call and SMS message log data could be matched and aggregated at the daily level as well.

After the data have been aggregated to the appropriate time frame of interest (e.g., daily sociability estimates), the psychometric properties of the sensed behavioral data should be examined. For this step, we recommend using techniques that are already common in psychological methods, including measures of central tendency (e.g., mean, median, and mode), distributional qualities (e.g., standard deviation, skew, and kurtosis), and relationships among the behavioral measurements and their reliability over time (e.g., autocorrelation and test-retest correlations). Additionally, we suggest researchers compute interindividual and intraindividual estimates to examine differences in variability due to between- persons (i.e., different persons) and within-person (i.e., time) factors. These techniques are an ideal starting place because they provide descriptive information about the data (e.g., shape of the distribution or dependence among observations) that may help identify the best modeling approach based on observed properties of the sensor data (e.g., whether the data meet assumption checks).

Challenges for Smartphone Sensing in Psychological Research

Typical smartphone-sensing studies collect data over time and with great fidelity, generating huge quantities of observations and placing the approach clearly within the domain of “big data” and its associated analytic techniques (Gosling & Mason, 2015; Miller, 2012; Yarkoni, 2012). These features of the method currently require highly technical setup, meaning that researchers (or their collaborators) must have considerable technical and computational expertise (e.g., using R or Python for data management and analysis). These requirements are common to most cases of big-data research, but there are also several challenges that are unique to smartphone sensing that warrant further discussion. These challenges center on the development of behavioral measures from smart-phone data, standards for study ethics, safeguards for participant privacy, and data security.

Developing behavioral measures from smartphone data

The promise of smartphone sensing for psychologists is the possibility of converting basic sensor data (e.g., accelerometer, microphone, and GPS readings) into broader psychologically interesting variables (e.g., physical activity, sociability, and situational information; Harari et al., 2015). Psychologists are particularly well equipped to play a major role in developing such behavioral measures. To date, behavioral inferences extracted from smartphone data have varied widely across individual sensing studies, and many of the behavioral classifiers used have been validated in small, homogeneous samples. For example, researchers have examined the validity of smartphone-based accelerometers for measuring physical activity (Case, Burwick, Volpp, & Patel, 2015), as well as the validity of accelerometers and GPS for identifying modes of transportation (Reddy et al., 2010). However, a psychological approach to measurement and assessment focusing on issues of reliability, validity, and generalizability has much to contribute to the task of developing novel and meaningful behavioral measures.

Studies focused on the reliability of behavioral inferences are needed for researchers to develop the measures (e.g., certain classifiers, combinations of data, or algorithms) that are most consistent and generalizable across different smartphone devices and populations. Additionally, studies of the validity of these measures can reveal how sensor-based behavioral measures relate to self-reported behavioral measures and other objective behavioral measures (construct validity) and how sensor-based behavioral measures relate to important outcomes (external validity). Psychologists are also well equipped to identify the sampling rates (e.g., thin slices, periodic, or continuous) needed to achieve optimal predictive models with respect to behavioral measures obtained from smartphone data (criterion validity). This area is a particularly important one for future research because the behavioral inferences obtained from smartphone data may underestimate or overestimate certain behaviors. For example, social interactions could be underestimated if a person is around other people who are not speaking or do not have Bluetooth enabled on their devices because the sensing application would not register their presence. Social interactions could also be overestimated if a person is alone but watching TV with the volume turned up because the sensing application might register the presence of human speech. Such

psychometric considerations point to the roles that psychologists can play in developing behavioral measures that capture the important features of people's behavioral lifestyles.

Standards for ethics, privacy, and security

There are growing concerns regarding the extent to which mobile devices collect behavioral information, often on behalf of commercial and governmental entities (Madden & Rainie, 2015). This state of affairs raises a series of ethical issues for researchers wanting to make use of smartphone-sensing data. Here we offer some initial guidelines for meeting ethical standards, safeguarding participant privacy, and ensuring data security.

Study ethics—Smartphone-sensing research is by its nature unobtrusive, potentially continuous, and observational. Therefore, sensing studies require ethics approval from institutional review boards. As in other psychological studies, participants in sensing studies should enroll voluntarily in the study, be made aware of the data that are collected, and agree to use the sensing application on the smartphone device. Naturally, researchers need to be sensitive to the technological competence of their participants (Bakke, 2010) and should consider providing training sessions about how to use the application (and perhaps the smartphone device itself) during the consent process.

As an observational method, smartphone-sensing research demands transparency between researchers and participants as a central research practice. Transparency can be achieved in several ways. For instance, the way that the sensing application works and the data storage practices being implemented should be made clear to participants. In practice, transparency is best implemented as part of an ongoing informed-consent process, starting with a session providing participants with information about the sensing app (e.g., the types of data it collects) and ending with a debriefing session (e.g., the goals of the study and an opportunity to receive a copy of their data).

An interview approach to informed consent has proven successful in previous observational sensing studies (e.g., electronically activated recorder [EAR] studies or the StudentLife study; Mehl & Pennebaker, 2003; Mehl, Pennebaker, Crow, Dabbs, & Price, 2001; Wang et al., 2014). When both entrance and exit interviews are conducted, the researcher is able to monitor participants' reactions to the study and keep participants updated with study-relevant information. However, the delivery of the consent process may need to be adapted to alternative formats (e.g., using video chat or phone calls for communicating with participants, using short informational videos to describe the study and application), particularly for studies collecting data from larger or global samples.

Of course, research is needed to determine the degree to which participants are prepared to give such consent and whether consent rates vary according to participant characteristics (e.g., age, privacy concerns, and motivation to learn about oneself). The probability of giving consent is likely influenced by the perceived costs of doing so (e.g., compromised privacy) and the expected benefits (e.g., new insights into one's behaviors). To get an initial sense of whether participants might be willing to provide consent to participate in studies that track their behaviors, we surveyed a group of college students ($N = 1,516$) about their willingness to participate in such research. Ninety-six percent of respondents said they would be willing

to participate in research that permitted them to self-track their psychological states and behaviors over time. Of that group, percentages varied in terms of the intrusiveness of the data they were willing to provide, including responding to EMAs one or more times a day (54%) and providing access to data from sensors on their smartphone (46%); their Web-browsing history (33%); their online educational platforms, such as the Canvas Learning Management System (Instructure, Inc., Salt Lake City, Utah; 47%); their social media accounts (42%); and wearable devices, such as Fitbit (Fitbit, Inc., San Francisco, California; 47%).

Concerns about data use must be balanced against the perceived benefits of participating in smartphone-sensor studies. Our survey data suggest that students experience a variety of motivations that might serve as incentives to participate in such research. We found that students reported wanting to participate in a self-tracking program if it helped them to improve their academic performance (80%), manage their time (63%), understand when they are most productive (61%), keep track their exercise or dieting habits (59%), improve their mental health (58%), keep track of stress and its sources (58%), and improve their physical health (55%). In other domains (e.g., online personality questionnaires), personalized feedback has proven to be a powerful incentive for participation across a range of demographic groups (Gosling & Mason, 2015; Kosinski, Matz, Gosling, Popov, & Stillwell, 2015). Overall, our self-reported survey data and findings from other domains suggest that it is possible to recruit participants for smartphone studies but that some concerns about privacy clearly remain. In the coming years, research will be needed to determine the causes of these concerns and what, if anything, can be done to assuage them.

Safeguarding participant privacy—Smartphone-sensing studies also demand attention to privacy. Given the sensitive nature of the data being collected, participants should be provided with maximum control over their personal digital records to ensure participant privacy is respected. Research practices that permit participants to withdraw or retroactively redact (i.e., remove or delete) their data from the study should be the standard. Moreover, these withdrawal and redaction procedures should be relatively effortless for participants. For instance, participants could withdraw from a study at any time by simply uninstalling the sensing app from their smartphone or redact their data at any point during the study should they wish to do so by simply providing a written request (via e-mail) to the researcher. This participant-redaction approach has been used successfully with data collection in studies in which other behavioral observation methods were used (e.g., EAR studies; Mehl et al., 2001), suggesting it is an effective means of respecting participants' privacy. Tools that allow participants to view and manage their own data are another, perhaps ideal way to provide participants with control over their own personal data. As of 2016, this is not a standard feature of sensing-app systems. However, data-privacy researchers have developed and field tested a promising personal metadata management framework that allows individuals to manage their data and select third parties with whom they would like to share their data (de Montjoye, Shmueli, Wang, & Pentland, 2014).

Another important consideration is that laws vary geographically on whether audio recording is legal, so researchers who aim to collect microphone data will need to attend to local laws. One potential solution to this problem is to make use of sensing apps that process

audio data on the phone to extract behavioral inferences (e.g., stress level from voice pitch, linguistic features) without storing the microphone data. Instead, the data collected from participants would consist of the behavioral inferences made on the device (not the actual audio data).

Establishing data security—Smartphone-sensing studies also require researchers to pay special attention to data security because smartphone data are inherently personally revealing about participants' daily lives (e.g., whom they communicate with, the places they visit). Researchers need to attend to data security at the stages of collection, storage, and sharing. Examples of secure data-collection practices include using smartphone-sensing apps that transmit sensor data to their servers securely using encryption technologies. For instance, by uploading data to the servers only when a participant is connected to WiFi, the data can be transferred using secure-sockets-layer (SSL) encryption. SSL is an industry-standard technique used to ensure that data transferred between devices are encrypted and shared securely. Data storage should be done using password-protected servers. These password-protected servers should be accessible only to researchers that are central to the data collection and data processing stages of the study. Certain types of sensor data contain information that is inherently personally identifiable. For example, sensitive content (e.g., names) may be recorded via the microphone sensor (e.g., in conversation), the phone logs (e.g., calls and SMS), or the Bluetooth sensor (e.g., the name of someone else's smart-phone device). In such instances, researchers should consider ways to replace any personal names or e-mail addresses with a unique, anonymity-preserving identifier, such as a randomly generated alphanumeric code. When the data are shared with other researchers, all personally identifiable information should be removed.

Looking Forward

The last 5 years have witnessed great progress in researchers' ability to undertake smartphone-sensing studies; smartphone-sensing technology is on the verge of presenting a feasible and unobtrusive method for collecting behavioral data from people as they go about their daily lives. These methods are beginning to be used in psychological research, but their use has yet to become widespread. However, the present generation of studies is set to yield sensing systems that overcome many of the obstacles that have slowed the uptake of sensing methods. In particular, new sensing systems with point-and-click interfaces will automate many of the tasks needed in smart-phone-sensing research. These point-and-click systems will facilitate widespread use of sensing methods by facilitating the setup of the system, such as providing software for the back-end and front-end components and handling automated data-processing tasks. Moreover, conferences and workshops that facilitate interdisciplinary collaborations among psychologists and computer scientists have begun to be held (e.g., Campbell & Lane, 2013; Lee, Konrath, Himle, & Bennett, 2015; Mascolo & Rentfrow, 2011; Rentfrow & Gosling, 2012) and will continue to play a helpful role for researchers interested in integrating smart-phone sensing into their studies. As smartphone sensing becomes commonplace in psychological research, we anticipate that psychology is finally on the verge of fully realizing its promise as a truly behavioral science.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Kostadin Kushlev, Neal Lathia, Sandrine Müller, and Jason Rentfrow for their helpful feedback on earlier versions of this article.

Funding

This research was supported by National Science Foundation (NSF) Award BCS-1520288.

References

- Bakke E. A model and measure of mobile communication competence. *Human Communication Research*. 2010; 36:348–371. DOI: 10.1111/j.1468-2958.2010.01379.x
- Baumeister RF, Vohs KD, Funder DC. Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*. 2007; 2:396–403. DOI: 10.1111/j.1745-6916.2007.00051.x [PubMed: 26151975]
- Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophrenia Bulletin*. 2014; 40:1244–1253. DOI: 10.1093/schbul/sbu033 [PubMed: 24609454]
- Bernard HR, Killworth PD, Sailer L. Informant accuracy in social network data IV: A comparison of clique-level structure in behavioral and cognitive network data. *Social Networks*. 1980; 2:191–218. DOI: 10.1016/0378-8733(79)90014-5
- Bernard HR, Killworth PD, Sailer L. Informant accuracy in social-network data V. An experimental attempt to predict actual communication from recall data. *Social Science Research*. 1982; 11:30–66. DOI: 10.1016/0049-089X(82)90006-0
- Boase J. Implications of software-based mobile media for social research. *Mobile Media & Communication*. 2013; 1:57–62. DOI: 10.1177/2050157912459500
- Boase J, Ling R. Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication*. 2013; 18:508–519. DOI: 10.1111/jcc4.12021
- Brown J, Broderick AJ, Lee N. Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of Interactive Marketing*. 2007; 21(3):2–20. DOI: 10.1002/dir.20082
- Campbell, AT., Lane, ND. Smartphone sensing: A game changer for behavioral science. Workshop held at the Summer Institute for Social and Personality Psychology; Davis: The University of California; 2013 Jul.
- Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *Journal of the American Medical Association*. 2015; 313:625–626. DOI: 10.1001/jama.2014.17841 [PubMed: 25668268]
- Chen, Z., Chen, Y., Hu, L., Wang, S., Jiang, X., Ma, X., ... Campbell, AT. ContextSense: Unobtrusive discovery of incremental social context using dynamic Bluetooth data. In: Brush, AJ., Friday, A., editors. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*: Adjunct publication; New York, NY: Association for Computing Machinery; 2014. p. 23-26.
- Chen, Z., Lin, M., Chen, F., Lane, ND., Cardone, G., Wang, R., ... Campbell, AT. Unobtrusive sleep monitoring using smartphones. 2013 7th International Conference on Pervasive Computing Technologies for Healthcare: PervasiveHealth '13; Brussels, Belgium: Institute for Computer Sciences, Social-Informatics, and Telecommunications Engineering; 2013. p. 145-152.
- Chittaranjan, G., Blom, J., Gatica-Perez, D. Who's who with Big-Five: Analyzing and classifying personality traits with smartphones. In: Lyons, K.Smailagic, A., Kenn, H., editors. *Wearable Computers (ISWC): 2011 15th Annual International Symposium*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2011. p. 29-36.

- Choudhury, T., Basu, S. Modeling conversational dynamics as a mixed-memory Markov process. In: Saul, LK, Weiss, Y., Bottou, L., editors. *Advances in neural information processing systems*. Vol. 17. Cambridge, MA: MIT Press; 2004. p. 281-288.
- Craik, KH. The lived day of an individual: A person-environment perspective. In: Walsh, WB, Craik, KH., Price, RH., editors. *Person-environment psychology: New directions and perspectives*. Mahwah, NJ: Erlbaum; 2000. p. 233-266.
- de Montjoye, YA., Quoidbach, J., Robic, F., Pentland, AS. Predicting personality using novel mobile phone-based metrics. In: Greenberg, AM, Kennedy, WG., Bos, ND., editors. *SBP '13: Proceedings of the 6th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction*. Berlin, Germany: Springer; 2013. p. 48-55.
- de Montjoye YA, Shmueli E, Wang SS, Pentland AS. openPDS: Protecting the privacy of metadata through SafeAnswers. *PLoS ONE*. 2014; 9:e98790. doi: 10.1371/journal.pone.0098790 [PubMed: 25007320]
- Eagle N, Pentland AS, Lazer D. Inferring friendship network structure by using mobile phone data. *PNAS: Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:15274–15278. DOI: 10.1073/pnas.0900282106
- Farrahi, K., Gatica-Perez, D. What did you do today? Discovering daily routines from large-scale mobile data. In: El Saddik, A., Vuong, S., editors. *Proceedings of the 16th ACM International Conference on Multimedia*; New York, NY: Association for Computing Machinery; 2008. p. 849-852.
- Funder DC. Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*. 2006; 40:21–34. DOI: 10.1016/j.jrp.2005.08.003
- Furr RM. Personality psychology as a truly behavioural science. *European Journal of Personality*. 2009; 23:369–401. DOI: 10.1002/per.724
- Gardy JL, Johnston JC, Sui SJH, Cook VJ, Shah L, Brodtkin E, Varhol R. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *New England Journal of Medicine*. 2011; 364:730–739. [PubMed: 21345102]
- Gosling SD, John OP, Craik KH, Robins RW. Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology*. 1998; 74:1337–1349. [PubMed: 9599447]
- Gosling SD, Mason W. Internet research in psychology. *Annual Review of Psychology*. 2015; 66:877–902. DOI: 10.1146/annurev-psych-010814-015321
- Harari GM, Gosling SD, Wang R, Campbell A. Capturing situational information with smartphones and mobile sensing methods. *European Journal of Personality*. 2015; 29:509–511. DOI: 10.1002/per.2032
- Henrich J, Heine SJ, Norenzayan A. The weird-est people in the world? *Behavioral & Brain Sciences*. 2010; 33:61–83. DOI: 10.1017/S0140525X0999152X [PubMed: 20550733]
- International Data Corp. Smartphone OS market share, Q1 2015. 2015. Retrieved from <http://www.idc.com/prodserv/smartphone-os-market-share.jsp>
- Kobayashi T, Boase J. No such effect? The implications of measurement error in self-report measures of mobile communication use. *Communication Methods and Measures*. 2012; 6(2):126–143. DOI: 10.1080/19312458.2012.679243
- Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*. 2015; 70:543–556. DOI: 10.1037/a0039210 [PubMed: 26348336]
- Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT. A survey of mobile phone sensing. *IEEE Communications Magazine*. 2010; 48(9):140–150. DOI: 10.1109/MCOM.2010.5560598
- Lane, ND., Mohammad, M., Lin, M., Yang, X., Lu, H., Ali, S., ... Campbell, A. Bewell: A smartphone application to monitor, model and promote wellbeing. In: Maitland, J, Augusto, JC., Caulfield, B., Chairs, editors. *5th International ICST Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*; Gent, Belgium: Institute for Computer Sciences, Social Informatics and Telecommunications Engineering; 2011. p. 23-26.

- Lathia N, Pejovic V, Rachuri KK, Mascolo C, Musolesi M, Rentfrow PJ. Smartphones for large-scale behavior change interventions. *IEEE Pervasive Computing*. 2013; 12:66–73. DOI: 10.1109/MPRV.2013.56
- Lathia, N., Rachuri, KK., Mascolo, C., Rentfrow, PJ. Contextual dissonance: Design bias in sensor-based experience sampling methods. In: Mattern, F., Santini, S., editors. *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*; New York, NY: Association for Computing Machinery; 2013. p. 183-192.
- Lee, S., Konrath, S., Himle, J., Bennett, D. Positive technology: Utilizing mobile devices for psychosocial intervention. Conference at the University of Michigan Institute for Social Research; 2015 Mar. Retrieved from <https://sites.google.com/a/umich.edu/positive-tech/home>
- LiKamWa, R., Liu, Y., Lane, ND., Zhong, L. Moodscope: Building a mood sensor from smartphone usage patterns. In: Chu, HH.Huang, P.Choudhury, RR., Zhao, F., Chairs, editors. *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*; New York, NY: Association for Computing Machinery; 2013. p. 389-402.
- Lu, H., Frauendorfer, D., Rabbi, M., Mast, MS., Chittaranjan, GT., Campbell, AT., ... Choudhury, T. Stress-Sense: Detecting stress in unconstrained acoustic environments using smartphones. In: Dey, AL.Chu, HH., Hayes, G., Chairs, editors. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*; New York, NY: Association for Computing Machinery; 2012. p. 351-360.
- Madden, M., Rainie, L. Americans' attitudes about privacy, security, and surveillance. 2015. Retrieved from Pew Research Center website: <http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/>
- Mascolo, C., Rentfrow, PJ. Social sensing: Mobile sensing meets social science. Workshop held; Cambridge, England: University of Cambridge; 2011.
- Mast MS, Gatica-Perez D, Frauendorfer D, Nguyen L, Choudhury T. Social sensing for psychology: Automated interpersonal behavior assessment. *Current Directions in Psychological Science*. 2015; 24:154–160. DOI: 10.1177/0963721414560811
- Mehl MR, Gosling SD, Pennebaker JW. Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*. 2006; 90:862–877. DOI: 10.1037/0022-3514.90.5.862 [PubMed: 16737378]
- Mehl MR, Pennebaker JW. The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*. 2003; 84:857–870. DOI: 10.1037/0022-3514.84.4.857 [PubMed: 12703653]
- Mehl MR, Pennebaker JW, Crow DM, Dabbs J, Price JH. The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, & Computers*. 2001; 33:517–523.
- Miller G. The smartphone psychology manifesto. *Perspectives on Psychological Science*. 2012; 7:221–237. DOI: 10.1177/1745691612441215 [PubMed: 26168460]
- Miluzzo, E., Lane, ND., Fodor, K., Peterson, R., Lu, H., Musolesi, M., ... Campbell, AT. Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application. In: Abdelzahar, T.Martonosi, M., Wolisz, A., editors. *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*; New York, NY: Association for Computing Machinery; 2008. p. 337-350.
- Morris ME, Aguilera A. Mobile, social, and wearable computing and the evolution of psychological practice. *Professional Psychology: Research and Practice*. 2012; 43:622–626. DOI: 10.1037/a0029041
- Paulhus, DL., Vazire, S. The self-report method. In: Robins, RW.Fraley, RC., Kruger, RF., editors. *Handbook of research methods in personality psychology*. New York, NY: Guilford; 2007. p. 224-239.
- Pejovic, V., Lathia, N., Mascolo, C., Musolesi, M. Mobile-based experience sampling for behaviour research. 2015. Retrieved from <https://arxiv.org/pdf/1508.03725.pdf>
- Pew Research Center. Cell phone and smartphone ownership demographics. Washington, DC: Author; 2014. Retrieved from <http://www.pewinternet.org/data-trend/mobile/cell-phone-and-smartphone-ownership-demographics/>

- Pew Research Center. Internet seen as positive influence on education but negative influence on morality in emerging and developing nations. Washington, DC: Author; 2015. Retrieved from <http://www.pewglobal.org/2015/03/19/internet-seen-as-positive-influence-on-education-but-negative-influence-on-morality-in-emerging-and-developing-nations/>
- Rabbi, M., Ali, S., Choudhury, T., Berke, E. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In: Landay, J., Shi, Y., Patterson, D., Rogers, Y., Xie, X., Chairs, editors. Proceedings of the 13th International Conference on Ubiquitous Computing; New York, NY: Association for Computing Machinery; 2011. p. 385-394.
- Rachuri, K.K., Musolesi, M., Mascolo, C., Rentfrow, P.J., Longworth, C., Aucinas, A. EmotionSense: A mobile phones based adaptive platform for experimental social psychology research. In: Bardram, J., Langheinrich, M., Chairs, editors. Proceedings of the 12th ACM International Conference on Ubiquitous Computing; New York, NY: Association for Computing Machinery; 2010. p. 281-290.
- Reddy S, Mun M, Burke J, Estrin D, Hansen M, Srivastava M. Using mobile phones to determine transportation modes. ACM Transactions on Sensor Networks (TOSN). 2010; 6(2) Article 13. doi: 10.1145/1689239.1689243
- Reis, H.T., Gosling, S.D. Social psychological methods outside the laboratory. In: Fiske, S.T., Gilbert, D.T., Lindzey, G., editors. Handbook of social psychology. 5. Vol. 1. New York, NY: Wiley; 2010. p. 82-114.
- Rentfrow, P.J., Gosling, S.D. Using smart-phones as mobile sensing devices: A practical guide for psychologists to current and potential capabilities. Preconference for the annual meeting of the Society for Personality and Social Psychology; San Diego, CA. 2012 Jan.
- Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, Mohr DC. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. Journal of Medical Internet Research. 2015; 17(7):e175. doi: 10.2196/jmir.4273 [PubMed: 26180009]
- Sandstrom GM, Lathia N, Mascolo C, Rentfrow PJ. Putting mood in context: Using smartphones to examine how people feel in different locations. Journal of Research in Personality. in press.
- Smith, A. Smartphone ownership: 2013 update. Washington, DC: Pew Research Center; 2013. Retrieved from <http://pewinternet.org/Reports/2013/Smartphone-Ownership-2013.aspx>
- Stisen, A., Blunck, H., Bhattacharya, S., Prentow, T.S., Kjærgaard, M.B., Dey, A., Jensen, M.M. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In: Song, J., Abdelzahar, T., Mascolo, C., Chairs, editors. Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (SenSys); New York, NY: Association for Computing Machinery; 2015. p. 127-140.
- Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Campbell, A.T. Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Brush, A., Friday, A., Kientz, J., Scott, J., Song, J., Chairs, editors. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; New York, NY: Association for Computing Machinery; 2014. p. 3-14.
- Wang, R., Harari, G.M., Hao, P., Zhou, X., Campbell, A. SmartGPA: How smartphones can assess and predict academic performance of college students. In: Mase, K., Langheinrich, M., Gatica-Perez, D., Gellersen, H., Choudhury, T., Yatani, K., Chairs, editors. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; New York, NY: Association for Computing Machinery; 2015. p. 295-306.
- Wrzus C, Brandmaier AM, von Oertzen T, Müller V, Wagner GG, Riediger M. A new approach for assessing sleep duration and postures from ambulatory accelerometry. PLoS ONE. 2012; 7:e48089. doi: 10.1371/journal.pone.0048089 [PubMed: 23110178]
- Wrzus C, Mehl MR. Lab and/or field? Measuring personality processes and their social consequences. European Journal of Personality. 2015; 29:250-271. DOI: 10.1002/per.1986
- Yarkoni T. Psychoinformatics: New horizons at the interface of the psychological and computing sciences. Current Directions in Psychological Science. 2012; 21:391-397.

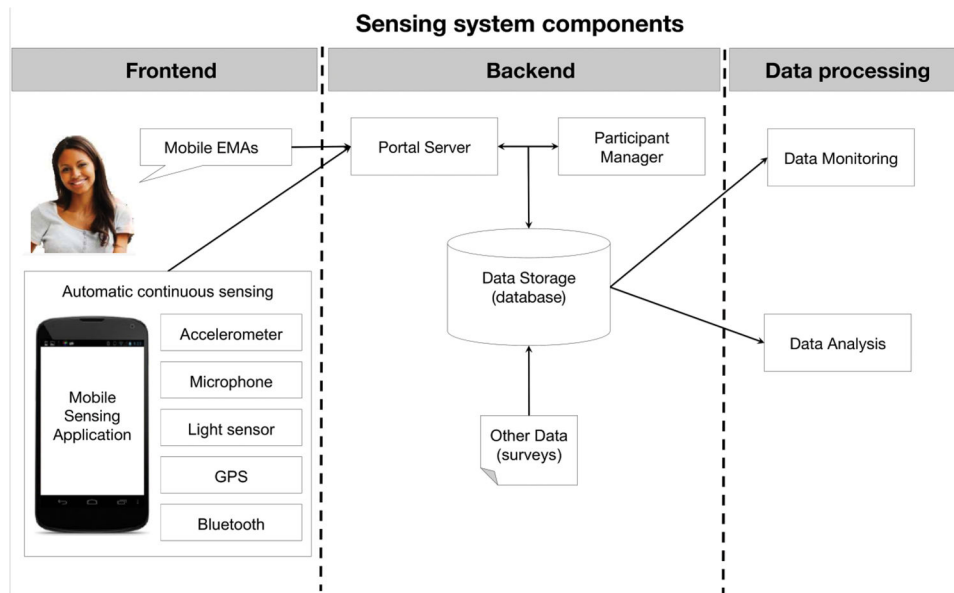


Fig. 1. Example set up of a smartphone-sensing system. EMA = ecological-momentary assessments.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Overview of Types of Smartphone Data: Functions, Features, and the Behaviors They Capture

Type of data	Function in the device	Features of the data	Behaviors captured from smartphone data	
			Social interactions	Daily activities
Mobile sensor data				
Accelerometer sensor	Orients the phone display horizontally or vertically	Coordinates X, Y, and Z; duration and degree of movement vs. stationary periods	✓	✓
BT radio	Allows the phone to exchange data with other BT-enabled devices	Number of unique scans; number of repeated scans	✓	
GPS scans	Obtains the phone location from satellites	Latitude and longitude coordinates; course (100–500 m) or fine-grained (100 m)	✓	✓
Light sensor	Monitors brightness of the environment to adjust phone display	Information about ambient light in the environment	✓	✓
Microphone sensor	Permits audio for calls	Audio recordings in the acoustic environment	✓	
Proximity sensor	Indexes when the phone is near the user's face to put display to sleep	Measurement of the proximity of an object to the screen (e.g., in centimeters)	✓	
WiFi scans	Permits the phone to connect to a wireless network	Number of unique WiFi scans; locations of WiFi networks		✓
Other phone data				
Call log	Records calls made and received	Incoming and outgoing calls; number of unique contacts	✓	
SMS log	Records text messages made and received	Incoming and outgoing text messages	✓	
App use log	Records phone applications used and installed	Number of apps; frequency and duration of app use	✓	✓
Battery status log	Records battery status	Battery charge times; low/medium/high battery status	✓	✓

Note: These are the most commonly used types of smartphone data at the time of this writing. This list is bound to grow as more sensors are embedded in the devices. BT = Bluetooth; GPS = global positioning system; m = meters; SMS = short message service; app = application.

Table 2

Summary of Areas of Opportunity for Psychological Research Using Smartphone Sensing Methods

Research objective	Types of research questions	Suggested analytic techniques
Describing behavioral patterns over time	<ul style="list-style-type: none"> • How does the behavior manifest across different units of time? • What are the normative behavior trajectories? • What individual difference factors predict the behavior trajectories? • What are the main types or profiles of behavioral patterns? • Do certain individuals or groups have a signature behavioral pattern? • What are the behavioral signatures associated with psychological constructs (e.g., personality, well-being)? 	<ul style="list-style-type: none"> • Psychometric analyses • SEM-based longitudinal models (GCM, LGCM) • Time series models • Change models (pre- vs. postevent) • Unsupervised machine learning techniques (e.g., <i>K</i>-means clustering, mixture models, hierarchical clustering) • Longitudinal profile and class analyses (LPA, LGCA)
Predicting life outcomes and implementing mobile interventions	<ul style="list-style-type: none"> • What are the key behavioral predictors of a given outcome (e.g., physical health, mental health, subjective well-being, performance)? • How does behavior change pre and post a significant life event/intervention? • When is the best time to intervene to promote positive behavior change? 	<ul style="list-style-type: none"> • Supervised machine learning techniques (e.g., DTA, CART, random forests) • Unsupervised machine learning techniques (e.g., <i>K</i>-means clustering, mixture models, hierarchical clustering) • Change models (pre- vs. postevent)
Examining social network systems	<ul style="list-style-type: none"> • How do social relationships manifest in the network? • How do social behaviors vary by other psychological factors (e.g., personality, status)? • What outcomes do features of the social network predict? 	<ul style="list-style-type: none"> • Social network analysis (SNA) • Dyadic analyses • Social relations model (SRM)

Note: SEM = structural equation modeling; GCM = growth curve modeling; LGCM = latent growth curve modeling; LPA = latent profile analysis; LGCA = latent growth curve analysis; DTA = decision tree analysis; CART = classification and regression tree.

Table 3

Overview of Key Design Decisions for Smartphone Sensing Studies

Key decisions	Description	Considerations	Implications
How long is the study duration?	The study duration will depend in part on the research questions of the study (e.g., interested in hourly, daily, weekly, monthly behavioral trends)	<ul style="list-style-type: none"> Participant incentives must be considered for encouraging use of the sensing app for long periods of time 	<ul style="list-style-type: none"> Attrition rates are bound to increase as study designs get longer in duration The duration of the study influences the types of generalization that can be made from the behavioral results
What is the sampling rate?	Sensing apps vary in how frequently they sample from the mobile sensors and phone logs (e.g., continuous, semi-continuous, periodic)	<ul style="list-style-type: none"> Longer study durations with high frequency samples tend to result in larger datasets. (e.g., hundreds of gigabytes of smartphone data) The best sampling rates for predicting various outcomes have yet to be determined 	<ul style="list-style-type: none"> Working with big datasets requires some technical and computational skills (e.g., in R or Python) Researchers may need to aggregate the smartphone data to the appropriate unit of analysis for their research questions
What smartphone device will participants use?	Participants may either be given devices to use for the duration of the study, or they may use their own devices	<ul style="list-style-type: none"> Which smartphone operating system (if any) is preferred by the researchers (e.g., Android or iOS)? How frequently do the smartphone data need to be sampled? 	<ul style="list-style-type: none"> The operating systems may be more or less popular in a given area (e.g., country), or with a given demographic group (e.g., socioeconomic status) The operating systems have different sampling constraints built into their systems (for a description of the differences, consult Online Materials, Appendix A)
What sensing application will be used?	Researchers may decide to design a sensing app, or use a pre-existing app (e.g., a commercial, open-source, or prototype app)	<ul style="list-style-type: none"> Sensing apps can be built with the proper resources and technical skills, but at this stage this requires computer scientist collaboration What types of smartphone data are needed? 	<ul style="list-style-type: none"> There are technical challenges that will be faced when designing smartphone software (e.g., bugs or crashes in the app) The app selection impacts the types of smartphone data collected, the sampling rates, and the setup of the sensing system (for example references see Table 4)

Table 4

Summary of Features and Functions of Smartphone Sensing Study Design

Design feature	Function	Examples
Sensing device and sensing application software	Front-end of the sensing system <ul style="list-style-type: none"> The smartphone user-interface that people use to respond to surveys and participate in the study Determines the types of sensor data collected and the sampling rate 	Sensing apps can be: <ul style="list-style-type: none"> Commercial (e.g., Easy M, MetricWire) Open-source (e.g., AWARE, Emotion Sense, Funf, Purple Robot; Sensus) Prototypes (e.g., BeWell, StudentLife, StressSense)
Server storage space	Back-end of the sensing system <ul style="list-style-type: none"> Communicates with the front-end to run the sensing software Can be either physical servers (hardware) or virtual servers (cloud-based) Stores the data in databases, in various file formats (e.g., CSV, JSON) 	Servers can be hosted by: <ul style="list-style-type: none"> Commercial platforms (e.g., Amazon Web Services) University or company-based computing and information technology services Databases: <ul style="list-style-type: none"> MongoDB MySQL
Data management	Data processing component of sensing system <ul style="list-style-type: none"> Monitor data collection to identify potential problems Extract behavioral inferences from the smartphone data (e.g., applying classifiers, algorithms, combining data) 	Programming languages <ul style="list-style-type: none"> R Python
Data analyses	<ul style="list-style-type: none"> Aggregate the sensor data to appropriate units of analysis (e.g., hourly, daily, weekly units) Run more formal analyses of the given research questions of interest 	Analytic software <ul style="list-style-type: none"> MATLAB Mplus R SPSS