J Multidisciplinary Scientific Journal

MDPI

Article

Concept, Possibilities and Pilot-Testing of a New Smartphone Application for the Social and Life Sciences to Study Human Behavior Including Validation Data from Personality Psychology

Christian Montag [1,*], Harald Baumeister [1], Christopher Kannen [2], Rayna Sariyska [1], Eva-Maria Meßner [1] and Matthias Brand [3]

[1] Institute of Psychology and Education, Ulm University, 89081 Ulm, Germany; harald.baumeister@uni-ulm.de (H.B.); rayna.sariyska@uni-ulm.de (R.S.); eva-maria.messner@uni-ulm.de (E.-M.M.)
[2] Independent Researcher; info@ckannen.com
[3] General Psychology: Cognition and Center for Behavioral Addiction Research (CeBAR), University of Duisburg-Essen, 47048 Duisburg, Germany; matthias.brand@uni-due.de
* Correspondence: christian.montag@uni-ulm.de; Tel.: +49-731-502-6550

Received: 27 December 2018; Accepted: 18 March 2019; Published: 27 March 2019

Abstract: With the advent of the World Wide Web, the smartphone and the Internet of Things, not only society but also the sciences are rapidly changing. In particular, the social sciences can profit from these digital developments, because now scientists have the power to study real-life human behavior via smartphones and other devices connected to the Internet of Things on a large-scale level. Although this sounds easy, scientists often face the problem that no practicable solution exists to participate in such a new scientific movement, due to a lack of an interdisciplinary network. If so, the development time of a new product, such as a smartphone application to get insights into human behavior takes an enormous amount of time and resources. Given this problem, the present work presents an easy way to use a smartphone application, which can be applied by social scientists to study a large range of scientific questions. The application provides measurements of variables via tracking smartphone–use patterns, such as call behavior, application use (e.g., social media), GPS and many others. In addition, the presented Android-based smartphone application, called *Insights*, can also be used to administer self-report questionnaires for conducting experience sampling and to search for co-variations between smartphone usage/smartphone data and self-report data. Of importance, the present work gives a detailed overview on how to conduct a study using an application such as *Insights*, starting from designing the study, installing the application to analyzing the data. In the present work, server requirements and privacy issues are also discussed. Furthermore, first validation data from personality psychology are presented. Such validation data are important in establishing trust in the applied technology to track behavior. In sum, the aim of the present work is (i) to provide interested scientists a short overview on how to conduct a study with smartphone app tracking technology, (ii) to present the features of the designed smartphone application and (iii) to demonstrate its validity with a proof of concept study, hence correlating smartphone usage with personality measures.

Keywords: smartphone use; digital phenotyping; mobile sensing; psychoinformatics; personality; psychology

1. Background

Scientists around the globe currently discuss if a paradigm shift is happening in the social and life sciences. With the rapid distribution of the World Wide Web since the early 1990s, and the advent of the iPhone in 2007, dramatic changes have occurred in societies around the globe [1]. Digital technologies have drastically changed how humans communicate (faster and at lower costs via long distance), find their way in unknown territory, and how they entertain themselves.

In December 2017, already 54.4% of the world's population was online [2] and 2.3 billion smartphone users have been estimated for the year 2017 with an estimate of nearly three billion users for 2020 [3]. This means that nowadays a vast part of the global population has access to such powerful technological devices. These developments are supposed to develop further and even stronger in the next years with the rise of the Internet of Things (IoT), where basically everything is connected to the Internet ranging from household machines to the car a person is driving (see also Stachl & Bühner [4]). Notably, every interaction with a digital device leaves a digital trace, which can be used to study human behavior ([5]; see also the area of micro-targeting in politics and economy in Matz et al. [6]).

This research area is of tremendous interest for scientists, because these digital traces can give genuine new insights into human behavior. In contrast to behavior in classic experiments, human behavior can be studied (i) in real life, (ii) on a longitudinal level, and (iii) at relatively low costs (compared to costly experiments conducted in laboratories). Therefore, it is not surprising that the recent past has seen tremendous efforts to use mobile phone technologies to record behavior via the different in-built sensors of smartphones (see also the smartphone manifesto by Miller [7]). The possibilities arising from recording human-(digital)-machine-interaction is unprecedented, because smartphones are ubiquitously available and used in nearly all situations of modern life. This led to a new movement in the scientific community, which has been coined *Psychoinformatics* [5,8,9]. In short, methods from computer science are applied to do diagnostics and interventions in psychology (such as predicting the personality of a person (example for a *trait*) or a depressed state of mind from the gathered digital data sources; example for a *state*). Note, that a trait describes an enduring or rather stable characteristic of a person, such as personality or intelligence. A state describes the current mood or well-being of a person, to give an example. Clearly, traits and states are entwined. Therefore, a person being anxious (trait) is more often in an anxious state in everyday life. Thereby state variables over a period of time can be seen as an approximate for a trait, but associations between these variables are far from perfect [10].

Other researchers prefer the term *digital phenotyping* over *Psychoinformatics* [11–16]. Note, that also the term *personal sensing* has been proposed [17]. It describes that the ubiquitously available sensors in diverse technological devices can be used to get insights into characteristics of a person. Clearly the terms *Psychoinformatics* and *digital phenotyping* belong together, because *digital phenotyping*, but also *mobile/personal sensing*, can be achieved via methods from *Psychoinformatics*. A growing body of studies have been already published in the realm of *Psychoinformatics,* predicting psychological traits, such as personality from Facebook usage [18,19], Twitter usage [20], WhatsApp usage [21] and also call variables from smartphones [22,23]. New work demonstrates the feasibility to predict cognitive functions from smartphone data [24] and mood from typing style on keyboards [25].

The first empirical evidence is also available for studies bringing together neuroscientific data (MRI scans) and smartphone tracked variables, such as Facebook: Here, it can be demonstrated that lower gray matter volumes of the nucleus accumbens can be linked to longer and more frequent Facebook use on smartphones [26]. In this study, participants' Facebook use was tracked over five weeks and appeared to be stable. Therefore, this behavior might be a manifestation of a trait (such as extraversion) and taken for itself might perhaps best be characterized as a habit (note that a single behavior such as Facebook behavior is probably too narrow to be defined as a trait; but see discussions in Baumert et al. [27]).

Other researchers demonstrated the feasibility to do depression and anxiety tracking via smartphones (states), by observing which kind of smartphone tracked variables co-vary with depressive or anxious states of a person (e.g., [28–30]) or well-being [31]. Note, that depressive tendencies can be also assessed via the study of Facebook profiles ([32]; see also the review on depression and mining of social media data by Guntuku et al. [33]). Aside from recording the interaction of a person with a smartphone, self-reported data can be collected via these devices, when programmed into an application.

Although the field of *mobile sensing* is developing rapidly, to our knowledge, easy-to-use smartphone applications are mostly lacking for the social sciences to give a large number of scientists access to this kind of new powerful research tool. Some existing solutions need to be mentioned: Among others, *Mindstrong Health* currently uses smartphone data (derived via their application) in clinical trials to test its usefulness for psychiatry/neurology [34]. Aside from this, an application called *BiAffect* for smartphones exist to enable mobile sensing in the field of bipolar disorders [35]. In addition, Schueller et al. [36] proposed their *Purple* platform to enable scientists to develop behavioral intervention tools. As one can see from these few existing examples, this research field is still new and many scientists interested in this area cannot participate, because it usually takes a long time to establish interdisciplinary networks and come up with a digital research solution, e.g., with a custom made designed smartphone application to study human behavior.

Therefore, the aim of the present work is to present a new smartphone application which can track a large number of relevant smartphone variables, including the frequency and length of usage of all installed applications, GPS tracking, call behavior, and so forth. Going beyond this, the presented application with the name *Insights* enables researchers to program questionnaires which can be administered to the participants of a study at any given time point. Therefore, it is possible to not only ask participants to fill in questionnaires at one or two time points, but also to do experience sampling with this technology, an important part of gathering longitudinal data in ecologically valid environments [37].

In the following, we aim at presenting deeper insights into the new *Insights* application for Android smartphones and explain how a study using smartphones can be designed and programmed. Moreover, we present further steps of the scientific process to demonstrate how recorded smartphone and questionnaire variables can be downloaded from the server and implemented in statistical analysis software (such as SPSS) to conduct the analyses for the scientific question at hand. In addition, important issues, such as server requirements and privacy protection issues will be discussed. The present work also provides first validation data from personality psychology demonstrating the validity of the tracked variables (see Appendix A). In the following, we discuss three steps, which provide an overview over a typical study using *Insights*.

Note that the present smartphone application "only" runs on Android, because Apple's iOS is sealed and does not allow the development of an application as the present one. This might cause a problem from a researcher's point of view, because iOS and Android users have been reported to differ in some psychological variables [38]; although differences might be in fact rather small [39].

2. Step 1: Designing a Smartphone-Based Study Using *Insights*

In the following, we will discuss how to set-up the smartphone application *Insights* that tracks smartphone variables and how to implement self-report questionnaires in the application.

After logging into the backend software, the researcher can add a new project. Here, the researcher names the project and creates a project code, which is unique for every study and will be used later for participants' registration (participants will insert this project code in the app to enter the study at a later time point). Afterwards, the researcher has the opportunity to insert information about the study, which will later be presented to the participants. Next, the researcher specifies how often the recorded data will be transferred from a participant's smartphone to the server (e.g., every 15 min). At this point the researcher also specifies which smartphone variables (e.g., calls, SMS, GPS) will be tracked and how often data will be actualized and saved to the phone of participants (e.g., every 15 min). Table 1 shows a list of all variables that can be recorded by the *Insights* app.

Table 1. All smartphone variables and data that can be tracked with *Insights*.

Recordable Data Types	Details
Contact List	Names, phone numbers, changes in list over time
Calls	Phone number, contact name, call type (incoming, outgoing or missed call), call duration, timestamp
SMS	Phone number, contact name, SMS type (incoming or outgoing SMS), message length, timestamp, text mining
User Sessions	Time of screen-on-event, time of phone-unlock-event, time of screen-off-event, session duration, elapsed time since last session
Battery Status	Battery level, timestamp, battery health information
Ringtone Settings	Ringtone type (loud, vibration or mute), ringtone volume
Installed Apps	App title, app package name, changes in app installations over time
App Sessions	App title, app package name, beginning of app session, end of app session, session duration
Aggregated App Sessions	Aggregated information from app sessions with total usage time and app entry count
Locations	GPS-, mobile- or WiFi-locations (latitude, longitude and altitude), precision, speed, record timestamp
Network Traffic	Number of bytes and network-packets received and transmitted, record timestamp
Phone Events	Events such as start, restart and shutdown of the device, plug-in for charging events, enabling/disabling the airplane-mode event
Questionnaires	Custom questionnaires programmed with *SurveyCoder* can be shown by a time schedule or can be manually accessed from a list.
Experiments	Custom HTML pages to bind experiments or web applications to the *Insights* data. This includes common HTML, Javascript and CSS and an optional connection to backend servers.

Researchers also have the option of choosing the "optional data recording", which lets participants decide which variables should be tracked and those that should not. In this case, in the beginning of the study, participants use the app menu to selectively choose which variables will be recorded on their phones (e.g., participants can decide to allow recording calls but not GPS signal). Regarding the tracking of SMS data researchers can choose from different additional settings, such as recording word count, numbers, length of words, and emoticons. After specifying which variables will be tracked, researchers have the opportunity to activate the feedback function. This means that data will be saved to participants' phones so that they can receive feedback on their digital consumption later, for example when the study is finished.

To sum up, at this point researchers had set up their project and defined which smartphone variables will be recorded. Next, the questionnaires, experiments and feedback, included in the study, were defined. Questionnaires can be programmed using a tool named *SurveyCoder*. After programming a questionnaire, the researchers can easily export it from the *SurveyCoder* software and import it to *Insights*. Experiments and feedback pages are developed as small HTML applications that can be imported to *Insights* as well. These three types of user interaction pages are called *popups* in *Insights*. Researchers can add *popups* to their project and specify whether they appear in the app menu as a list, or whether they are presented automatically to the participants at previously specified time intervals (e.g., every two days, every Monday at 2 p.m., every day randomly between 8 and 10 o'clock and so on).

3. Step 2: (De-)Installation of the Smartphone Application

In order to install the *Insights* application, the researcher asks the participants of his/her study to visit *www.insightsapp.org* and follow the link to download the *Insights* app. On a more general note: The *Insights* app is available for other scientists. Given the high development costs of the *Insights* application and the further support needed in the future to enable the service on future Android versions, a reasonable subscription fee will be charged.

After completing this first step, the general terms and conditions are presented on the screen of the phone and need to be accepted by the participant of the research project before proceeding further with registration. After doing so, the researcher guiding the installation process, inserts the project code, which is project-specific (meaning that it is different for every research project). Notably, this step and the other presented steps, could in principle, also be followed by a study's participant alone, thereby overcoming the boundaries of locally bound studies. In order to reduce potential errors, however, we recommend installing and initiating *together* with the participant.

On the following screen, the participant information is presented, describing which activities on the phones will be recorded during participation (e.g., calls, SMS, app usage). Here, the participants are again asked to give their consent. After doing so, the app installing person (i.e., researcher and/or participant) needs to grant all required permissions to the app.

Permissions for optional tracking functions are not mandatory. At this point, the app is ready to track data. On the following screen the installing person uses the "Start" button, in order to begin with the study. On the same screen, a personalized participation code (16-digit) is presented, which is an option that can be used if the *Insights* application needs to be reinstalled (e.g., if the participant switches to a new phone). With this 16-digit code the installing person can retrieve older data, when this is needed. Moreover, on the same screen another four menus are presented: For the questionnaires and for the experiments of the study (if applicable), for the statistics (this menu shows a feedback that participants receive during or after their participation in the study; if appropriate and in line with the research question) and for the settings of the *Insights* application. Last but not least, using the "Settings" menu the installing person, guided by the researcher, determines whether the data should be transferred to the server only when he/she has WiFi Internet access (by clicking on "Only WiFi") or also to use mobile data for sending the data (by clicking "Mobile Data and WiFi"). Notably, the study can be terminated at any time point and the tracked data can be handled and analyzed (see Figures 1 and 2 for further illustrations).

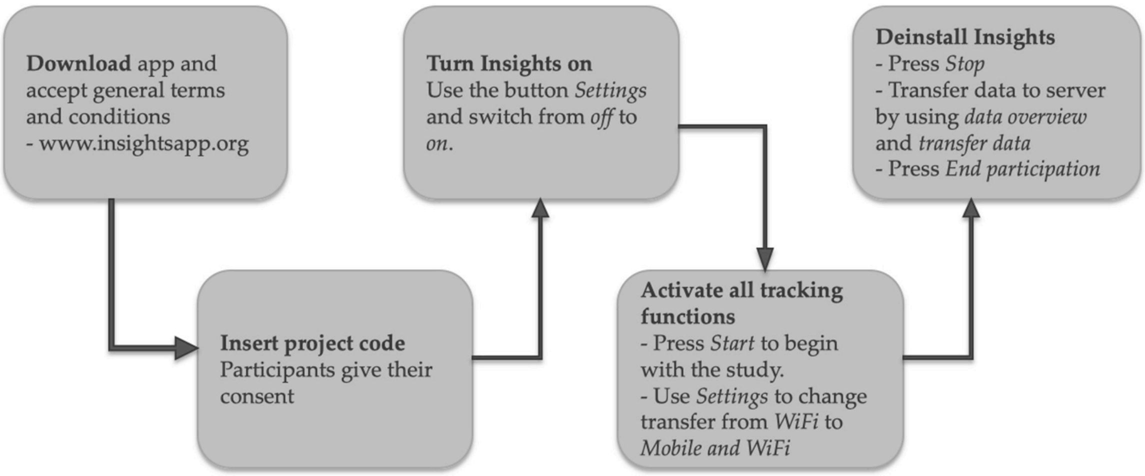

Figure 1. Tracking the smartphone behavior from a person: From installing the application on a smartphone of a study's participants to deinstallation.

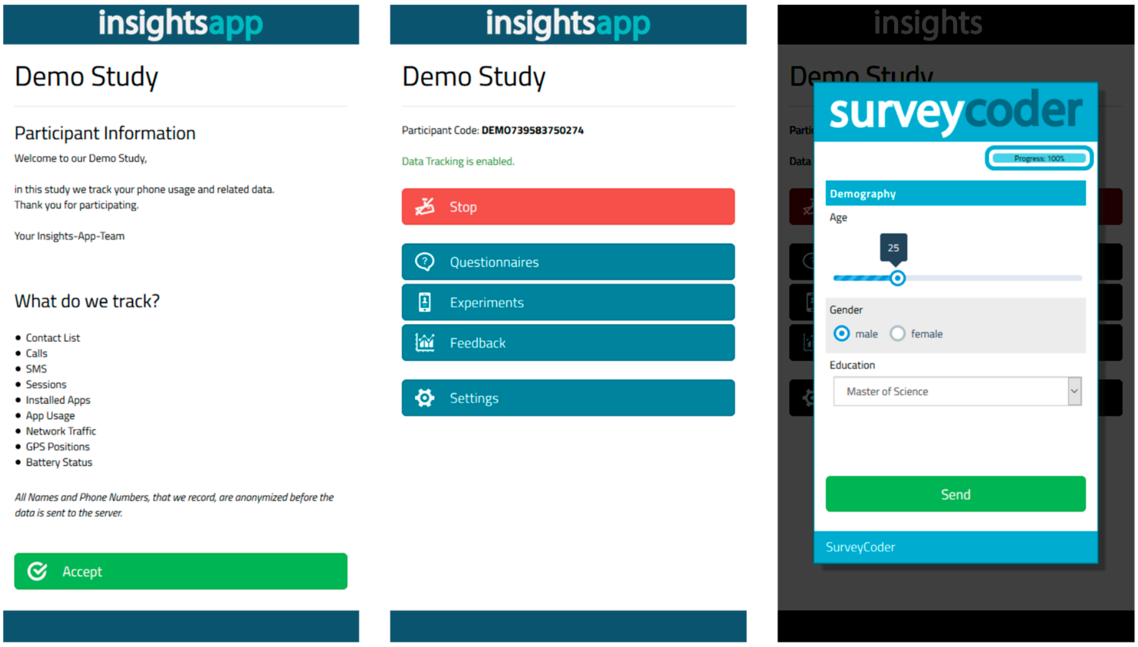

Figure 2. Guiding the participant of a study through the installation process: The left part of the figure informs the participant of a study supervised by the researcher what will be tracked on his/her phone before the study begins. The middle part of the figure shows the main menu with the "to do list" for the participants—such as the option to answer questionnaires. The right side of the figure shows a questionnaire that was chosen from the list.

4. Step 3: Data Handling and Analyses

Insights data can be analyzed in multiple ways, directly in the backend or with another software such as Excel or SPSS. For analyses in the *Insights* backend, the software offers a programming platform to build and run own scripts on the data. These scripts are written in PHP and support MySQL queries to collect and preprocess the database records. Data extraction and data preparation scripts allow the results to be printed directly on the screen or let the user export the data into files and download them (in long or wide format). If the servers support the installation of third-party software, external libraries can be imported and used in the scripts, too. It might be, for example, useful to create maps out of the GPS data or draw charts. In the pilot-phase, the authors of the present work developed and tested several scripts in order to export all database tables and preprocess the data (i.e., scripts for the aggregation of calls, SMS, app sessions, and user sessions to daily data). The results of these scripts are stored in CSV files and can be easily downloaded from the server. Afterwards they can be processed by third party software like SPSS. All of this together allowed a very easy and fast, but also very detailed and custom-tailored way for data analysis.

In sum, no programming knowledge is necessary. But users need to learn to handle some available scripts. Nevertheless, knowledge in PHP programming is helpful in writing custom-made scripts. Customized questionnaires are written in the Survey Coder language and tailored feedback is coded in HTML.

5. Server Requirements

Due to the collection of Big Data and the essential level of security, *Insights* has special requirements in relation to the server architecture. The app collects large amount of data, in direct proportion to the selected tracking settings. Therefore, the server that runs the *Insights* backend needs to be able to handle this amount of data.

With the default settings, the app stores around 250 KB of data per participant and day. During a study of 90 days and with 1000 participants in the study, this makes a total of 22.5 GB raw data.

But this is only a rough estimation, the real data amount bases on the settings and the individual participant's smartphone usage. It can be easily increased, when the settings are changed to a higher tracking rate or when many large questionnaires are displayed to the participants during the study. I.e., when the tracking rate for locations is set from the default of 15 min to one minute, 15 times more GPS positions will be recorded.

The server needs to be able to handle this data, especially at two points in time: The data logging and the data analysis. When the participants' devices collected some data, it is sent to the server. This typically happens every 15 min, but the transfer rate can be changed in the settings. A faster rate means more, but smaller server requests. A slower rate leads to less, but larger server requests.

However, after the settings have been established, the server needs to take the incoming data and write it into the database. For most servers this will be hassle-free possible. But for large studies with many thousands of participants, the webserver software should be chosen wisely. Even for smaller studies, the special requirement at this point will be the database storage, because the storage needs to be large enough to store the gigabytes of incoming data.

The second big task for the server is data analysis. During this process, the server needs to compute results on the collected data. This is done with MySQL and PHP or other installed third-party software. Processing a large amount of data takes quite a bit of CPU time and the processes might allocate a lot of memory. For faster progress, dedicated CPUs and enough RAM are useful. However, in most use cases a single virtual server or dedicated server will do the job.

6. Privacy Issues

The *Insights* app collects sensitive data. Hence it is very important to protect that data. The app tracks information about many aspects of smartphone usage, such as phone calls, SMS, Internet usage, or geographical locations. Even if only quantitative data is stored for most categories, it must be protected to ensure the privacy of each participant. *Insights* implements several layers of privacy protection (i) on the phones and (ii) on the server.

(i). Privacy Protection on the Phone

Android, the operating system for which the *Insights* app is designed for, already has strong privacy protection methods. All apps run capsuled inside their own sandbox, which means that no other app can access the data of another app.

The same applies of course for the *Insights* app. In addition, all sensitive information collected by the app is stored inside a private database, which is not accessible by any third-party software. Although this is already strong data protection, *Insights* brings its own privacy protection mechanisms. The most important of these privacy protection mechanisms is that the data collection is anonymous. This does not only guard the participant against other apps, but also guarantees his/her privacy even in the study. This is done by anonymizing the identity of the participant. Neither the name, the phone number, nor the phone ID is stored. Instead, *Insights* assigns a 16 character long unique identifier, called "participant code", to each user, in order to relate data to him/her. The first 8 digits of the participant code represent the project code, which is unique for each project. The second 8 digits are randomly created and identify the single participant. Furthermore, but not less important, is the privacy of all persons whose contact information are stored on the participant's phone or who are in contact with the participant during the study. Therefore, all phone numbers and contact names, that are collected during a study, are also anonymized directly on the phone. In this case, the anonymization is done by the cryptographic hash function "SHA-256". Cryptographic hash functions, generate a unique alphanumeric output string for any given input string. They are easy to compute, but the reverse way, reproducing the original information out of a hashed value, is extremely difficult. The big benefit of this approach is that, it is still possible to inspect social graphs across participants, since the same contact on two participants' phones will result in the same hash value.

Another privacy protection mechanism is that *Insights* collects only the data that is necessary for the current study. For that the director of studies defines which data is supposed to be tracked, when creating the project. He/she can disable sections i.e., disable the collection of calls or he/she can make sections optional, i.e., the location tracking. In the latter case, it is up to the participant to decide whether he/she wants to share the data. He/she will still be able to be part of the study anyway. For the named example of optional location tracking, he/she will still participate and provide all the other phone usage data, even if he/she does not share his/her locations.

Text messages can be fully read by *Insights*, which creates another possible privacy issue. To defuse this issue, the content of text messages will not leave the phone, but will be analyzed directly on the device. *Insights* has powerful functions to do so. On the one hand, these functions compute the text length, number of words, uppercase and lowercase characters. One the other hand, word lists can be set-up, then *Insights* will search for occurrences inside the message text and return the number of found words per list as result.

(ii). Privacy Protection on the Server

Each workgroup operates their own *Insights* backend installation. This allows workgroups to have full control on the server location and the data security. The backend software can be downloaded from *www.insightsapp.org* and installed on any server or web space that conform to the requirements in regard to the server software and processor power. All data that is tracked by the *Insights* app on the phone is only send to this server. The *Insights* backend software also has its own access permission system. Only the creator of a study or users that received permissions from him/her, have access to the study and the stored data. This assures that only members of the work-group can gain access to the study information.

The last level of privacy protection on the server is actually more directly related to protection of data transmission between smartphones and the server. Both, the *Insights* app and the *Insights* backend are designed to use SSL (https), even with self-signed certificates. SSL encrypts data when it is sent over the Internet and it is the de facto standard for all websites that needs a high level of security, like banking or shopping websites. The correct setup of a SSL certificate is in the responsibility of the server admin. We highly recommend installing a SSL certificate for data transportation, since this ensures that nobody can read the collected data during the transmission.

Finally and again, we want to mention in this privacy section that every researcher is free to choose his/her server and can therefore secure data privacy to his/her own standards, hence also respect the recently proposed GDPR. Moreover, telephone numbers and names of persons are irrevocably anonymized by a cryptographic hash function.

7. Conclusions and Future Directions

With *Insights* we provide the framework for examining all kind of research questions in the fields of behavior science, moving the (psychological) lab to where humans are and interact with each other. This approach probably will allow for gathering and analyzing large datasets in order to provide, not only standard statistics on people's smartphone behavior, which can be associated with all kinds of psycho-social assessments, but to also examine individual trajectories, examining prototypic usage patterns, and develop prediction models on for example, an individual's mental health course based on *Insights'* passive mobile sensing. Analyzing individuals' behavior on smartphones has immense potential for several research areas and also for clinical practice, given that, for example, *digital phenotyping* has been considered a key future research area in the field of understanding excessive and uncontrolled or problematic usage of the Internet [40,41]. Please note that other similar platforms have been designed, e.g., to foster responsible gaming [42] and/or use *mobile sensing* in the social sciences [43,44]. We also want to point to the interesting AWARE framework (http://www.awareframework.com) and the Beiwe framework (https://www.beiwe.org/). Both systems are open-source and have their own strength. But it is also important to note that the set of sensors used for data recordings differ between these projects and also with *Insights*: As a rule of thumb, AWARE and *Insights* currently seem to support more types of measurements compared to Beiwe. *Insights* additionally implements the feature of using HTML-based pages in the app, which allows to run experiments or give feedback information to the participants.

Furthermore, next development steps will enlarge the scope of *Insights* by integrating e.g., voice sensing features, implementing interfaces for integrating bio-psychophysiological (smart-)devices (e.g., heart rate and other stress and activity indicators) and extending *Insights* to e.g., provide just in time interventions (e.g., a mindfulness based intervention) based on the respective participants' assessed needs (e.g., stressful morning) and the contextual window of chance (e.g., GPS tracking shows a 10 min waiting time at a bus station).

Author Contributions: Conceptualization, C.M., H.B., C.K. and M.B.; Data curation, R.S. and E.-M.M.; Formal analysis, R.S.; Investigation, C.M., C.K., R.S. and E.-M.M.; Methodology, C.M. and C.K.; Project administration, C.K., R.S. and E.-M.M.; Software, C.K.; Supervision, C.M.

Funding: The position of Christian Montag was funded by Deutsche Forschungsgemeinschaft: MO 2363/3-2.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A.

Appendix A.1. First Validation Data from Personality Psychology

As a first (external) validation study of the *Insights* application (app) we revisited the question how call variables on a smartphone relate to the Big Five of Personality/Five Factor Model of Personality [22,45]. Such associations will make it possible to predict trait characteristics of a person from digital sources such as the interaction with the smartphone to call a person. In an earlier work, we already reported that smartphone-related call behavior is moderately linked to extraversion [22]. The smartphone data were recorded with a custom-built designed application called *Menthal* [46]. Of note, another work by Stachl et al. [23] also observed with their own built application a positive link between extraversion and smartphone-based call variables. This strongly supports the idea, that calling behavior on smartphones can be linked to extraversion. Extraverted persons are socially outgoing, vivid, and assertive. Therefore, it is convincing to assume that their extraverted behavior also manifests in their calling behavior. It was predicted that we would be able to replicate these findings using *Insights*.

Appendix A.1.1. Participants

For validation reasons, we report the data from n = 106 participants (37 males; mean-age: 23.34, median = 21, (SD = 7.59, range 18–63)). Only participants at the age of 18 or older were included in the analyses of the study. In total 0.9% of participants reported having secondary school leaving certificate, 6.6% a vocational baccalaureate diploma and 74.5% A-level. Another 1.9% of participants reported having polytechnic degree, and 16% a university degree. Participants were recruited mostly at Ulm University and gained university credits or monetary compensation for their participation. All participants owned a smartphone with an Android operating system. Data was recorded for the period of eight weeks, however, due to missing data and based on our efforts to include as many participants as possible at the time point of data analysis, 12 days of recordings were used in the present study. Thus, for all smartphone variables means over 12 days were computed. Participants were provided with detailed information about the study prior to their participation. The participant information included information on the aim of the study, on all variables tracked with *Insights* in the current study, and how data was recorded and where it was saved and analyzed. Moreover, participants were informed on who has access to the recorded data and on their right to request the deletion of their data at any time during the study. All participants gave digital consent prior their participation in the study. The present study was approved by the local ethic committee of Ulm University, Ulm, Germany.

Appendix A.1.2. Measures

All participants brought their smartphones to the study and installed the *Insights* application. Moreover, all participants filled in demographic information and a short version of the *Trait Self Description Inventory* (TSDI, the administered German version has been described in Olaru et al. [47]) on their smartphones. The short version consists of 42 items. Each item was answered via a seven-point Likert scale ranging from strongly disagree (1) to strongly agree (7). Internal consistencies were satisfying with Openness to Experience ($\alpha = 0.80$), Conscientiousness ($\alpha = 0.71$), Extraversion ($\alpha = 0.75$), Agreeableness ($\alpha = 0.76$) and Neuroticism ($\alpha = 0.84$).

For the present work we recorded the following variables as depicted in Tables A1 and A2. Means for each variable were built over the period of 12 days (not necessarily 12 consecutive days, when days in between had missing data).

Appendix A.1.3. Statistical Analyses

First, we report the means, medians and standard deviations in Table A1. Second, correlation analyses were conducted in order to assess the associations between the recorded smartphone variables and the TSDI scales. These are reported in Table A2. All smartphone variables were non-normally distributed. The variables agreeableness and conscientiousness also deviated from the normal distribution. Thus, for consistency reasons only Spearman's correlations were conducted and the Bias corrected and accelerated bootstrap CIs are reported. Please note, that we also assessed in the present project self-reports on "smartphone addiction" [48], molecular genetic samples (see a first work published with this data [49]), and tracked other variables such as Facebook or WhatsApp usage on the smartphone. Presenting all these findings would go far beyond the main aim of this research piece and therefore we only present the personality associations in this paper. For the current study, all measures and data exclusions concerning the number of recorded days are reported. Please note that the here presented data will be made accessible in a completely anonymized form for other scientists upon reasonable request (as SPSS or excel file).

Appendix A.1.4. Results

In Table A1, the descriptive statistics of the call variables, recorded in the present study, using *Insights*, are presented against the descriptive statistics, reported by Montag et al. [22] using the *Menthal* app.

Table A1. Descriptive statistics (including the mean, standard deviation and the median) for the recorded call variables on the smartphones using the *Menthal* app (in Montag et al. [22]) and the *Insights* app (in the current study).

Recorded Variables	Montag et al. [22] ($n = 49$)	Present Work ($n = 106$)
Total number of calls per day (incoming + outgoing + missed calls)	M = 3.34 SD = 2.46 Median = 2.71	M = 2.25 SD = 2.40 Median = 1.46
Number of incoming calls per day	M = 1.07 SD = 1.03 Median = 0.79	M = 0.45 SD = 0.49 Median = 0.33
Number of outgoing calls per day	M = 2.29 SD = 1.67 Median = 1.68	M = 1.42 SD = 1.77 Median 0.75
Number of missed calls per day	M = 0.71 SD = 1.02 Median = 0.50	M = 0.37 SD = 0.35 Median = 0.33
Total call duration (min.) per day (incoming + outgoing calls)	M = 9.02 SD = 8.09 Median = 6.93	M = 7.71 SD = 10.08 Median = 4.15

Note: M = mean, SD = standard deviation.

We refrain from directly comparing the results from both studies using inferential statistics since the samples differ both in size and demographic characteristics. Moreover, due to specific usage patterns in both samples, we did not expect the exact same results. Overall the values in the study by Montag et al. [22] are slightly higher than the results from the present study. However, the differences seem to be very small and the patterns are rather consistent (e.g., the mean number of outgoing calls is higher than the mean number of incoming calls in both studies).

Please find all correlation patterns between the assessed call variables and the Big Five of Personality depicted in Table A2. Mostly in line with previous works [22,23] we were able to robustly link extraversion with the call variables measured on the smartphone in the present study. Notably, the correlations are about the same as in the older cited works ([22]: Association between extraversion and all call variables (mean over all parametric correlations; see further information beyond Table A2): 0.35; [23]: Extraversion and frequency of call behavior: 0.33).

The negative association between neuroticism and incoming calls in Table A2 is interesting as it is not available in earlier works. This might be due to the different measures used to assess the Big Five Personality (here the 42-item Big-Five scale basing on the TSDI, in Montag et al.'s earlier work the NEO-FFI [22] and in Stachl et al.'s work the BFSI [23]). The different inventories used to assess the Big Five Personality speak on the other hand for a robust effect concerning the investigated call behavior and extraversion, although it needs to be mentioned that duration of calls was not associated with extraversion in Stachl et al.'s work [23].

The inverse link between neuroticism and incoming calls is also plausible, because persons who are anxious, tense, and moody, might refrain from this kind of communication. In light of the lack of evidence in earlier works, this link clearly needs to be further investigated. For reasons of completeness, we also like to point out a new piece of interesting work investigating the link between the personality trait sensation seeking and tracked smartphone variables [50]. Further work from Denmark presents additional data linking extraversion to several tracked smartphone variables [51].

Table A2. Associations between the Big Five of Personality and several call variables tracked on the smartphones (Spearman's correlations are presented).

	Openness	**Conscientiousness**	**Agreeableness**	**Extraversion**	**Neuroticism**
Total number of calls	0.043	−0.028	−0.002	0.360 **	−0.179
Sig. (2-tailed)	0.659	0.779	0.988	<0.001	0.066
BCa 95% CI Lower	−0.157	−0.201	−0.211	0.164	−0.342
Upper	0.240	0.152	0.199	0.540	0.002
Incoming calls	−0.012	0.032	−0.119	0.364 **	−0.274 **
Sig. (2-tailed)	0.905	0.748	0.225	<0.001	0.005
BCa 95% CI Lower	−0.210	−0.147	−0.336	0.194	−0.446
Upper	0.173	0.204	0.099	0.527	−0.087
Outgoing calls	0.039	−0.054	0.023	0.315 **	−0.115
Sig. (2-tailed)	0.694	0.584	0.814	0.001	0.242
BCa 95% CI Lower	−0.157	−0.219	−0.175	0.103	−0.295
Upper	0.223	0.127	0.215	0.498	0.069
Missed calls	−0.047	−0.040	−0.016	0.212 *	−0.105
Sig. (2-tailed)	0.632	0.688	0.871	0.030	0.284
BCa 95% CI Lower	−0.243	−0.216	−0.225	0.000	−0.277
Upper	0.149	0.142	0.185	0.412	0.071
Total call duration	0.151	0.017	−0.080	0.288 **	−0.108
Sig. (2-tailed)	0.122	0.866	0.413	0.003	0.270
BCa 95% CI Lower	−0.047	−0.149	−0.286	0.094	−0.283
Upper	0.346	0.194	0.127	0.466	0.083

Note. * $p < 0.05$, ** $p < 0.01$; $n = 106$. All correlations are presented on two-sided test level. The here presented correlations were computed on the daily average over the period of 12 days. BCa 95% CI = bias-corrected and accelerated 95% confidence intervals. For reasons of better comparability, we also provide the Spearman correlations between extraversion and the following variables from the work by Montag et al. [22] beyond what has been reported on a parametric level in this earlier work: total number of calls: 0.407, $p = 0.004$; incoming calls: 0.337, $p = 0.018$; outgoing calls: 0.407, $p = 0.004$; missed calls: 0.464, $p = 0.001$; calls duration: 0.429, $p = 0.002$.

In sum, both the similar numbers with respect to the mean/standard deviations of the call variables, and comparable associations between extraversion and the call variables, provide a first indication of *Insights*, being a valid application for psycho-informatic research approaches.

References

1. Montag, C.; Diefenbach, S. Towards Homo Digitalis: Important research issues for psychology and the neurosciences at the dawn of the internet of things and the digital society. *Sustainability* **2018**, *10*, 415. [CrossRef]
2. Internet World Stats. Available online: https://www.internetworldstats.com/stats.htm (accessed on 14 November 2017).
3. Statista. Available online: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ (accessed on 14 November 2017).
4. Stachl, C.; Bühner, M. Show me how you drive and I'll tell you who you are. Recognizing gender using automotive driving parameters. *Procedia Manuf.* **2015**, *3*, 5587–5594. [CrossRef]
5. Montag, C.; Duke, É.; Markowetz, A. Toward psychoinformatics: Computer science meets psychology. *Comput. Math. Methods Med.* **2016**, *2016*, 2983685. [CrossRef] [PubMed]
6. Matz, S.C.; Kosinski, M.; Nave, G.; Stillwell, D.J. Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12714–12719. [CrossRef]
7. Miller, G. The smartphone psychology manifesto. *Perspect. Psychol. Sci.* **2012**, *7*, 221–237. [CrossRef] [PubMed]
8. Yarkoni, T. Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Curr. Dir. Psychol. Sci.* **2012**, *21*, 391–397. [CrossRef]
9. Markowetz, A.; Błaszkiewicz, K.; Montag, C.; Switala, C.; Schlaepfer, T.E. Psycho-informatics: Big data shaping modern psychometrics. *Med. Hypotheses* **2014**, *82*, 405–411. [CrossRef]
10. Augustine, A.A.; Larsen, R.J. Is a trait really the mean of states? *J. Individ. Differ.* **2012**, *33*, 131–137. [CrossRef]
11. Onnela, J.P.; Rauch, S.L. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* **2016**, *41*, 1691–1696. [CrossRef]
12. Insel, T.R. Digital phenotyping: Technology for a new science of behavior. *JAMA* **2017**, *318*, 1215–1216. [CrossRef]
13. Raballo, A. Digital phenotyping: An overarching framework to capture our extended mental states. *Lancet Psychiatry* **2018**, *5*, 194–195. [CrossRef]
14. Torous, J.; Onnela, J.P.; Keshavan, M. New dimensions and new tools to realize the potential of RDoC: Digital phenotyping via smartphones and connected devices. *Transl. Psychiatry* **2017**, *7*, e1053. [CrossRef] [PubMed]
15. Torous, J.; Staples, P.; Barnett, I.; Sandoval, L.R.; Keshavan, M.; Onnela, J.P. Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. *NPJ Digit. Med.* **2018**, *1*, 15. [CrossRef]
16. Torous, J.; Keshavan, M. A new window into psychosis: The rise digital phenotyping, smartphone assessment, and mobile monitoring. *Schizophr. Res.* **2018**, *197*, 67–68. [CrossRef] [PubMed]
17. Mohr, D.C.; Zhang, M.; Schueller, S.M. Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annu. Rev. Clin. Psychol.* **2017**, *13*, 23–47. [CrossRef]
18. Kosinski, M.; Stillwell, D.; Graepel, T. Private traits and attributes are predictable from digital records of human behavior. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5802–5805. [CrossRef] [PubMed]
19. YouYou, W.; Kosinski, M.; Stillwell, D. Computer-based personality judgments are more accurate than those made by humans. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 1036–1040. [CrossRef]
20. Quercia, D.; Kosinski, M.; Stillwell, D.; Crowcroft, J. Our twitter profiles, our selves: Predicting personality with twitter. In Proceedings of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, Boston, MA, USA, 9–11 October 2011.
21. Montag, C.; Błaszkiewicz, K.; Sariyska, R.; Lachmann, B.; Andone, I.; Trendafilov, B.; Eibes, M.; Markowetz, A. Smartphone usage in the 21st century: Who is active on WhatsApp? *BMC Res. Notes* **2015**, *8*, 331. [CrossRef]

22. Montag, C.; Błaszkiewicz, K.; Lachmann, B.; Andone, I.; Sariyska, R.; Trendafilov, B.; Reuter, M.; Markowetz, A. Correlating personality and actual phone usage: Evidence from psychoinformatics. *J. Individ. Differ.* **2014**, *3*, 158–165. [CrossRef]
23. Stachl, C.; Hilbert, S.; Au, J.Q.; Buschek, D.; De Luca, A.; Bischl, B.; Hussmann, H.; Bühner, M. Personality traits predict smartphone usage. *Eur. J. Pers.* **2017**, *31*, 701–722. [CrossRef]
24. Dagum, P. Digital biomarkers of cognitive function. *NPJ Digit. Med.* **2018**, *1*, 10. [CrossRef]
25. Stange, J.P.; Zulueta, J.; Langenecker, S.A.; Ryan, K.A.; Piscitello, A.; Duffecy, J.; McInnis, M.G.; Nelson, P.; Ajilore, O.; Leow, A. Let your fingers do the talking: Passive typing instability predicts future mood outcomes. *Bipolar Disord.* **2018**, *20*, 285–288. [CrossRef] [PubMed]
26. Montag, C.; Markowetz, A.; Blaszkiewicz, K.; Andone, I.; Lachmann, B.; Sariyska, R.; Trendafilov, B.; Eibes, M.; Kolb, J.; Reuter, M.; et al. Facebook usage on smartphones and gray matter volume of the nucleus accumbens. *Behav. Brain Res.* **2017**, *329*, 221–228. [CrossRef]
27. Baumert, A.; Schmitt, M.; Perugini, M.; Johnson, W.; Blum, G.; Borkenau, P.; Costantini, G.; Denissen, J.J.; Feleeson, W.; Grafton, B.; et al. Integrating personality structure, personality process, and personality development. *Eur. J. Pers.* **2017**, *31*, 503–528. [CrossRef]
28. Canzian, L.; Musolesi, M. Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, Osaka, Japan, 7–11 September 2015; ACM: New York, NY, USA, 2015.
29. Chow, P.I.; Fua, K.; Huang, Y.; Bonelli, W.; Xiong, H.; Barnes, L.E.; Teachman, B.A. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. *J. Med. Internet Res.* **2017**, *19*, e62. [CrossRef] [PubMed]
30. Elhai, J.D.; Tiamiyu, M.F.; Weeks, J.W.; Levine, J.C.; Picard, K.J.; Hall, B.J. Depression and emotion regulation predict objective smartphone use measured over one week. *Pers. Individ. Differ.* **2018**, *133*, 21–28. [CrossRef]
31. DeMasi, O.; Feygin, S.; Dembo, A.; Aguilera, A.; Recht, B. Well-Being tracking via smartphone-measured activity and sleep: Cohort study. *JMIR mHealth uHealth* **2017**, *5*, e137. [CrossRef] [PubMed]
32. Schwartz, H.A.; Eichstaedt, J.; Kern, M.L.; Park, G.; Sap, M.; Stillwell, D.; Ksinski, M.; Ungar, L. Towards assessing changes in degree of depression through facebook. In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, Baltimore, MD, USA, 27 June 2014; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014.
33. Guntuku, S.C.; Yaden, D.B.; Kern, M.L.; Ungar, L.H.; Eichstaedt, J.C. Detecting depression and mental illness on social media: An integrative review. *Curr. Opin. Behav. Sci.* **2017**, *18*, 43–49. [CrossRef]
34. Mindstrong. Available online: https://mindstronghealth.com/ (accessed on 25 May 2018).
35. BiAffect. Available online: www.biaffect.com (accessed on 25 May 2018).
36. Schueller, S.M.; Begale, M.; Penedo, F.J.; Mohr, D.C. Purple: A modular system for developing and deploying behavioral intervention technologies. *J. Med. Internet Res.* **2014**, *16*, e181. [CrossRef]
37. Hofmann, W.; Patel, P.V. SurveySignal: A convenient solution for experience sampling research using participants' own smartphones. *Soc. Sci. Comput. Rev.* **2015**, *33*, 235–253. [CrossRef]
38. Shaw, H.; Ellis, D.A.; Kendrick, L.R.; Ziegler, F.; Wiseman, R. Predicting smartphone operating system from personality and individual differences. *Cyberpsychol. Behav. Soc. Netw.* **2016**, *19*, 727–732. [CrossRef]
39. Götz, F.M.; Stieger, S.; Reips, U.D. Users of the main smartphone operating systems (iOS, Android) differ only little in personality. *PLoS ONE* **2017**, *12*, e0176921. [CrossRef]
40. Fineberg, N.A.; Demetrovics, Z.; Stein, D.J.; Ioannidis, K.; Potenza, M.N.; Grünblatt, E.; Brand, M.; Billieux, J.; Carmi, L.; King, D.L.; et al. Manifesto for a European research network into Problematic Usage of the Internet. *Eur. Neuropsychopharmacol.* **2018**, *28*, 1232–1246. [CrossRef] [PubMed]
41. Montag, C.; Reuter, M.; Markowetz, A. The impact of psychoinformatics on Internet addiction including new evidence. In *Internet Addiction*; Springer: Cham, Switzerland, 2017; pp. 221–229.
42. Drosatos, G.; Nalbadis, F.; Arden-Close, E.; Baines, V.; Bolat, E.; Vuillier, L.; Kostoulas, T.; Budka, M.; Wasowska, S.; Bonello, M.; et al. Enabling Responsible Online Gambling by Real-time Persuasive Technologies. *Complex Syst. Inform. Model. Q.* **2018**, *17*, 44–68. [CrossRef]
43. Harari, G.M.; Lane, N.D.; Wang, R.; Crosier, B.S.; Campbell, A.T.; Gosling, S.D. Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspect. Psychol. Sci.* **2016**, *11*, 838–854. [CrossRef]

44. Harari, G.M.; Gosling, S.D.; Wang, R.; Campbell, A.T. Capturing Situational Information with Smartphones and Mobile Sensing Methods. *Eur. J. Pers.* **2015**, *29*, 509–511. [CrossRef]
45. McCrae, R.R.; John, O.P. An introduction to the five-factor model and its applications. *J. Pers.* **1992**, *60*, 175–215. [CrossRef]
46. Google Commerce Ltd. Mental. Available online: https://play.google.com/store/apps/details?id=open.menthal (accessed on 28 February 2018).
47. Olaru, G.; Witthöft, M.; Wilhelm, O. Methods matter: Testing competing models for designing short-scale big-five assessments. *J. Res. Pers.* **2015**, *59*, 56–68. [CrossRef]
48. Kwon, M.; Lee, J.Y.; Won, W.Y.; Park, J.W.; Min, J.A.; Hahn, C.; Gu, X.; Choi, J.; Kim, D.J. Development and validation of a smartphone addiction scale (SAS). *PLoS ONE* **2013**, *8*, e56936. [CrossRef] [PubMed]
49. Sariyska, R.; Rathner, E.M.; Baumeister, H.; Montag, C. Feasibility of linking molecular genetic markers to real-world social network size tracked on smartphones. *Front. Neurosci.* **2018**, *12*, 945. [CrossRef] [PubMed]
50. Schoedel, R.; Au, Q.; Völkel, S.T.; Lehmann, F.; Becker, D.; Bühner, M.; Bisschl, B.; Hussmanna, H.; Stachl, C. Digital Footprints of Sensation Seeking. *Z. Psychol.* **2018**, *226*, 232–245. [CrossRef]
51. Mønsted, B.; Mollgaard, A.; Mathiesen, J. Phone-based metric as a predictor for basic personality traits. *J. Res. Pers.* **2018**, *74*, 16–22. [CrossRef]