

A simple location tracking app for psychological research

Kristoffer Geyer and David A. Ellis

Lancaster University

Lukasz Piwek

University of Bath

Authors Note

Kristoffer Geyer and David A. Ellis, Department of Psychology, Lancaster University, Bailrigg, Lancaster, United Kingdom.

Lukasz Piwek, School of Management, University of Bath, Bath, UK

Corresponding author: David Ellis, Department of Psychology, Lancaster University, Bailrigg, LA1 4YW, Lancaster, United Kingdom. E-mail: d.a.ellis@lancaster.ac.uk

A simple location tracking app for psychological research

Abstract

Location data gathered from a variety of sources is particularly valuable when it comes to understanding individuals and groups. However, much of this work relies on participants' active engagement to regularly report their location. More recently, smartphones have been used to assist with this process, but while commercial smartphone applications are available, these are often expensive and not designed with researchers in mind. In order to overcome these and other related issues, we have developed a freely available Android application that logs **location accurately, stores** data securely, and ensures participants can provide consent or withdraw from a study at any time. Further recommendations and R code are provided to assist with subsequent data analysis.

Keywords: digital traces, GPS, location semantics, ecological momentary assessment

Introduction

Where a person spends their time can provide numerous insights into their behaviour, personality and mood (Chorely, Whitaker, and Allen, 2015). For example, location measures can be predictive of depressive symptoms, and levels of social anxiety (Huang, et al., 2016; Palmius et al., 2017; Saeb, et al., 2016). Other research has shown that individuals with comparable personalities often access similar locations (Noë, et al., 2016). While these studies remain important, critics have argued that comparatively little research has been conducted when it comes to understanding what is psychologically important about the locations people choose to occupy in real-time (e.g., Rauthmann et al., 2014). Often, designs have relied on location databases harvested from social media websites (Chorley, Whitaker and Allen, 2015). However, this method presents new limitations because using social media to sample multiple locations is likely to only include the reporting of socially desirable locations (Schwartz, and Halegoua, 2015). This effect may be magnified further as social media users are motivated to selectively report their location in order to maintain or boost their social status (Fitzpatric, Birnholtz, and Gergle, 2016; Guha, and Birnholtz, 2013; Schwartz, and Halegoua, 2015). Similar approaches have involved self-report derived from experience sampling smartphone applications (e.g., Sandstrom, Lathia, Mascolo, and Rentfrow, 2017). However, like social media capture, the reporting of every location that an individual visits requires an extensive amount of effort. As a result, data generated from either method provides a patchy account of where a person spends their time.

Related research in medicine has also sought to understand how environmental factors influence a variety of other health outcomes (James et al., 2016). GPS data specifically, can provide highly accurate, time-stamped geographic co-ordinates, which link locations with environment (Müller et al., 2017). For example, trips between location points can then help quantify general levels of physical activity (Carlson et al., 2015; Jankowska, Schipperijn, and

Kerr, 2015). Unfortunately, much of this research relies on the use of stand-alone GPS trackers, which are often expensive and may not work correctly in some buildings (Piazarro et al., 2017). **In addition, stand-alone trackers may place a significant burden on participants who may not want to wear additional devices for extended periods of time (Schmidt et al., 2018).**

Smartphones, in contrast, are readily available and used frequently by the majority of the general population (Wilcockson et al., 2018). **Advances in battery development, power management systems and location triangulation have also ensured that GPS data derived from mobile devices has become a realistic prospect (Gadziński, 2018).**

However, despite almost every device containing a GPS sensor, there remains a lack of suitable software that is freely available for those working within psychology and the social sciences more generally (Harari et al., 2017; Piwek, Ellis, and Andrews, 2016). Researchers will struggle to find appropriate alternatives from commercial application repositories, e.g., via Google Play or App stores (Apple, 2017; Google, 2017a). This is largely because these applications have not been developed with social research in mind (Table 1). Many other commercial applications often struggle to strike a suitable balance between high levels of accuracy and duration of logging, which are methodologically important for location-based research (Palmius et al., 2017). Alternative ‘out of the box’ solutions include OpenPaths (2012) and Google Timeline (**Google, 2018**). While functional, OpenPaths relies on drawing data from other applications, which request location updates. Therefore, data collection becomes completely under the jurisdiction of another application and beyond a researcher’s control. Similarly, Google Timeline operates by documenting changes in location. Location is not mapped after a specific length of time but only when a pre-defined distance has been covered, in order to conserve both battery and memory. **Secondary data analysis derived**

from these systems also makes it easier for participants to omit location data from their records at any time.

In order to overcome previous methodological limitations, we have developed a freely available application (*PEG LOG*) that records the location of an Android smartphone. This is an attempt to enhance the quality and quantity of data that is available to researchers when studying the significance of individual and group movements. Additionally, we wish to prompt transparency and replication by making the source code and supplementary materials freely available. Finally, the application requires minimal effort from participants, while ensuring that associated data remains encrypted and secure throughout.

[Table 1. A comparison of features offered by current methods, which aim to track location from smartphones.]

	P	M	FTU	Features			Signal			Extra		
				AS	CR	C	GPS	Wi-fi	Ce	PI	PM	OS
PEG LOG	Android	*	*	*	*	*	*	*	*	*	*	*
AWARE (Ferreira et al., 2015)	Android/iOS	*	*		*	*	*	*	*	*	*	*
Device Analyser (Wagner et al., 2014)	Android	*	*	*	*				*		*	
EmotionSense (Lathia et al., 2013)	Android		*		*	*			*		*	*
Funf (Aharony et al., 2011)	Android		*		*	*	*	*	*	*	*	*
Lifedata (Runyan et al., 2013)	Android/iOS	*		*		*	*	*	*	*		
Google Timeline (Google, 2018)	Android/iOS	*	*	*			*	*	*	*	*	
MetricWire	Android/iOS	*		*	*	*	*	*	*	*	*	
Momento (Carter et al., 2007)	Android		*			*			*			
MovisensXS	Android	*		*	*	*	*	*	*	*	*	
Ohmage (Ramanathan et al., 2012)	Android/iOS		*	*			*					*
Open Paths (Open Paths, 2017)	Android/iOS		*	*	*		*	*	*	*	*	
ResearchKit (Apple, 2016)	iOS	*	*			*	*	*	*	*		*
SystemSens (Falaki et al., 2011)	Android		*		*				*		*	*

Notes: P = Platform, M = Actively maintained, FTU = Free to use, AS = Available in app store, CR = Continuous Recording, C= Customisable, GPS = Extracts single from Global Positioning System, Wi-fi = Computes location from wi-fi connection, Ce = Cellular location tracking, PI = Provides point of interest analysis (increases accuracy), PM = Allows for path mapping of movements, OS = Open source code

Summary of application architecture

The application runs on Android devices and is available from the Google Play store (see **supplementary materials**). It was designed in order to provide regular updates that circumvent limitations associated with stand-alone location trackers. For example, GPS signals are typically inaccessible from inside a building, but the application can switch to rely on other available sources that report location, e.g., Wi-Fi and Network signals. However, it should be noted that both these signals are generally less accurate than GPS alone (Android, 2018a; Canzian, and Musolesi, 2015).

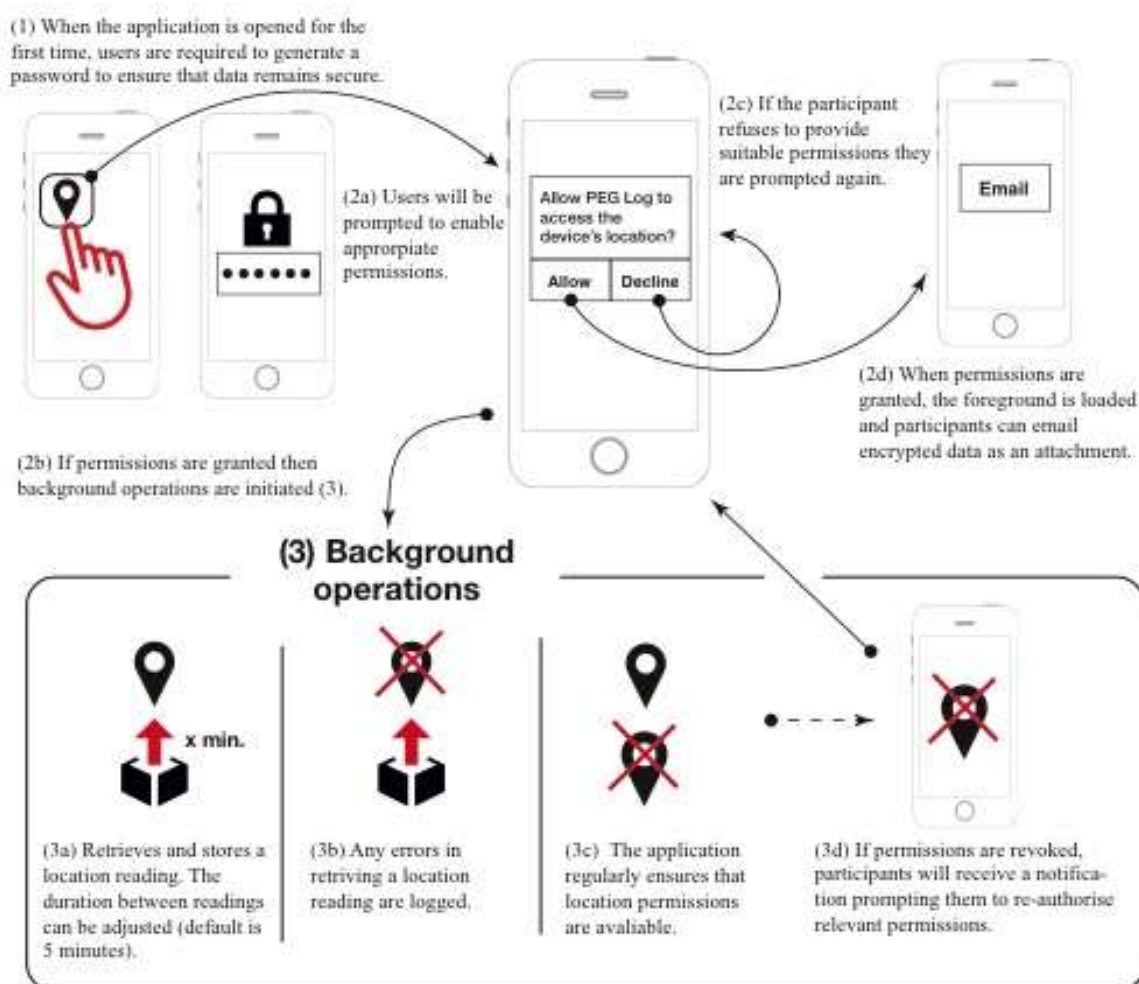


Figure 1: Infographic demonstrating foreground (1, 2a-2d) and background (3a-3d) operations of the PEG LOG application. The only aspect of the application accessible by participants is the main

activity page that requests the relevant location permissions (2a) and allows participants to read documentation, change their password, view their location data and email files (2d).

Installation

The installation process is intended to be straightforward and requires almost no time or commitment from participants. The application must first be downloaded from the Google Play store **and will require less than 30MB of space**. Once installed, participants simply have to open the application. This will allow users to view how the application works, set a password, authorise appropriate permissions, and confirm that data collection can commence (Figure 1). Participants will typically be asked to provide permissions relating to the use of location and call data. The latter permission is required so the application can record errors if GPS data, for example, is not available via a standard cellular connection. It is advisable that all participants send some pilot data to a researcher at the beginning of any study to ensure location tracking is proceeding as expected.

Foreground operations

Consent and data security

Location and related behavioural data are sensitive measures and therefore protocols must ensure that the privacy of participants' information is protected during data collection (James et al., 2016). On first launching the application, participants are presented with a brief information screen that specifies what information is being collected and how to stop data collection. **This information and the Terms of Service and Privacy Policy can be recalled**

at any time from within the application¹. The **app** will then instruct participants to provide a 6-digit password to secure their data. This password has to be communicated to the researcher in due course to allow for any encrypted data to be accessed. It is not possible to start data collection without first defining a password. If a participant wishes to withdraw they can choose to not submit their data or password. A participant can also delete files from their device by simply uninstalling the application.

To ensure participants remain informed and aware of their active participation at all times, *PEG LOG* provides two visual reminders. First, the application displays a small icon in the top corner of the screen at all times. Second, a permanent text reminder will appear in the notification drawer, which explicitly states that the application is collecting location data. This has the added benefit of improving the reliability of the application as the Android operating system allocates more processing power for applications, which declare that operations are running in the background (Android, 2018b).

During the collection phase, all location data is stored in a 256-bit SQL cypher database. This ensures that even if the source code of the application was compromised, no data can be retrieved without the original password. Only *PEG LOG* can access this database. All exported data and associated error logs are encrypted with a 128-bit key. However, while the original password for the SQL database is fixed and cannot be changed, if the original 6-digit code is forgotten, participants can modify their password via the main screen. This change will only apply to data after it has been prepared for export.

¹ See supplementary materials for a link to the Terms of Service and Privacy Policy.

Data storage and export

PEG LOG allows for the **passive** recording of location data, which is stored locally on the device. This can be exported on demand. We have opted to avoid the use of a central server to maximise usability i.e., researchers who want to use this app do not need to set up a cloud-based storage system, which assists with increased reliability and longevity. This also ensures that participants are in complete control of their own data throughout the entire collection process.

Data is stored in a manner that ensures only *PEG LOG* can access this information. Any active email account can be used to send the file with encrypted data attached. To export data, participants can select 'email' and then provide permission for the application to write to external memory if required (Figure 1). Data is then retrieved from the SQL cypher database and placed into an encrypted attachment. A separate encrypted error log is also provided. **The application places no limit on how much data can be collected. While it may not be possible to send larger files via email, this is unlikely to be an issue for the vast majority of studies. For example, collecting a location reading every minute for a period of 2 weeks generates approximately 3MB of data.**

All data is exported in PDF format and the included R-code allows this data to be converted and unencrypted to a text file quickly (see supplementary materials). The PDF format was chosen because it can be encrypted while also allowing participants to view their own data on almost any computer or device (including smartphones). **While this may inadvertently allow participants to edit their own data, the nature and format of location files mean that such alterations would require considerable effort. Participants who become uncomfortable with data collection or who no longer want to take part are more likely to simply uninstall the application.**

Background operations

Recording location

The application relies on the FuseLocationProvider (Google, 2017b). This provides access to GPS, Wi-Fi, and Network analysis in order to retrieve latitude, longitude, accuracy levels in meters measured by a radius of confidence, and a UNIX timestamp. The application is considered high priority, which means the most accurate reading available is provided regardless of battery expenditure. The order of favourability of trace (in relation to accuracy) is therefore: GPS, Wi-Fi, followed by network analysis (Canzian, and Musolesi, 2015). A location update is requested by default every 5 minutes. The file returned is a lengthy string, which is stored in an SQL database (see supplementary materials for an example of raw data). When location data cannot be collected, the application attempts to diagnose the source of the problem. For example, if the phone is restarted, this is recorded. A list of potential errors identified and their associated codes are documented in the supplementary materials. Beyond these tasks, **very little processing of location data is carried out within PEG LOG itself. As a result, the application has a minimal impact on battery performance, even when the gap between each location reading is comparatively small (e.g., one-minute intervals). This compares favourably with many other popular applications, which run a large a number of background processes and data sharing mechanisms by default, which are rarely made clear to the end user (Van Kleek et al., 2017). Final decisions regarding specific data processing and analysis operations are therefore** left open to researchers after data has been collected and exported from the device.

Resilience of the application

We have identified seven potential ways that the background operations of the application could be prevented from functioning. Participants could inadvertently stop data collection by: (1) turning off their phone, (2) closing the application, (3) closing all tasks running in the foreground, (4) forcing the closure of all active applications (5) disabling location services, (6) enabling power saving modes, or (7) uninstalling the application. Addressing these issues in order, if the phone is turned off, upon restarting, the application will automatically resume and continue collecting data. This will be documented in internal memory and mark an interruption of data collection due to a restart event. Similarly, if the foreground section of the application is closed then the background service will continue to run. Even if all foreground applications are cleared, background services will not be interrupted. However, if a force closure of all applications occurs then the participant will be required to open the application again in order to continue with data collection. If a participant does not have location permissions enabled or these are turned off, the application will send the participant a notification. This will remind them that location permissions should be enabled.

Participants can click on the notification, which will point them to relevant settings where relevant permissions can be re-enabled. In addition, power saving modes present in some Android devices may limit the number of location points recorded by a device if it has not been used for a lengthy period of time (Android, 2018b). However, this can be partly mitigated by ensuring that participant's manually whitelist the application. This reduces the impact of battery optimization techniques (see supplementary materials for more information). Finally, uninstalling the application is interpreted as a desire to withdraw from the study and this will stop the collection of data and delete all associated files.

Customization

Which location data source (GPS, Wi-Fi, etc.) is used by default, and the frequency of location updates can be customised by following a simple modification to the original source code. This is outlined within one non-expert friendly file: Constants (this file explaining the project structure is available via the associated GitHub account). Following customisation, the application can then be redistributed on the Google Play store.

PEG LOG will never share data with other applications however, location information collected could be analysed alongside other streams of data obtained from other applications and devices. This might include methodologies, which capture time-stamped objective measures of behaviour (e.g., physical activity from an accelerometer) or survey response items over longer periods of time (e.g., mood assessment from an experience sampling application) (Carlson et al., 2015; Jankowska, Schipperijn, and Kerr, 2015; Pizarro et al., 2017).

Storing Location Data

Following standard data protection guidelines, all data should be removed from email servers following transmission and stored in line with standard ethical and data protection procedures. While data will always remain encrypted when stored on an email server, passwords should not be sent in the same email as raw data. In addition, while the application presented here remains open and freely available, location data should be treated as particularly sensitive. Researchers should keep in mind that raw and processed location data may reveal activity patterns, which participants may want to keep private (James et al., 2016). While data can be anonymised, location co-ordinates are likely to reveal a persons' place of work and home address with very little pre-processing. If this data were to be shared openly

with additional anonymisation, one option could involve the removal or masking of spatial data in sensitive locations (e.g., the home). Ensuring that participants understand the granularity of data collected will help guide subsequent sharing decisions however, more work is required as it is now possible to generate even larger datasets from a variety of smartphone metrics (Harari et al., 2017; Piwek, Ellis and Andrews, 2016).

Data Processing and Analysis

A complete review concerning how location data can be analysed is beyond the scope of this paper however, broadly speaking; there are three key ways of **analysing location data**. First, location points can be placed into space-based topologies such as: cafés, university campus buildings, nightclubs, etc. Locations via this method can be further characterised based on how they are clustered or relate to other geographic databases, e.g., census records, crime statistics, foursquare database (Canzian, and Musolesi, 2015; Chorley, Whitaker, and Allen, 2015; Jankowska, Schipperijn, and Kerr, 2015; Rauthmann, et al., 2014). Second, movements as a form of behaviour can be characterised in a number of ways (Canzian, and Musolesi, 2015). This can provide information relating to: distance travelled, radius of gyration, etc. For example, recent psychological research has shown that an analysis including information relating to both journey and destination is incrementally more valuable (Huang, et al., 2016). Finally, a consideration of time can provide information regarding when an individual is engaged in specific activities or behaviours. For example, it is possible to separate indoor time from outdoor time (Jankowska, Schipperijn, and Kerr, 2015). Research can, of course, combine all of these approaches, however, there remains potential for these analyses to develop further **as location** data becomes easier to collect. We have therefore included additional supplementary R-code to assist with these developments. This marked-up code

will process raw location data, prepare data for analysis, and generate some basic visualisations (Figure 2).

A**B**

[Figure 2: A simple visualisation derived from a short period of location tracking. This includes a point map (A) which plots individual location points (darker points demonstrate repeated observations) and a path map (B) showing the directionality of the same participants movement.]

Conclusion

Previous research that has involved the collection and analysis **of location** data from smartphones and other digital devices have found this digital trace to be both predicative of future **behavior** and a variety of other individual differences (Chorely, Whitaker, and Allen, 2015). However, conclusions are often based on incomplete recordings of location from systems and devices, which are not transparent in their functionality or freely available to other researchers. Overcoming these limitations for social science remains important in order to pre-empt the well-documented issues with self-reported data, especially when recording location information **over days, weeks or even months** (Rauthmann et al., 2014). In summary, here we have presented a freely available location-tracking application and associated analysis code, which will allow researchers across a variety of disciplines to conduct rigorous research into individual and group movements.

Declarations

Availability of data and material

Supplementary materials including links to the application, source code, example data and associated R-code are available here:

<https://github.com/kris-geyer/PEGlog>

Competing interests

The authors report no conflicts of interest.

Funding

This work was part funded by the Centre for Research and Evidence on Security Threats [ESRC Award: ES/N009614/1].

Authors' contributions

KG developed and tested the application and wrote the first draft of the manuscript. DAE contributed to the writing of the manuscript and supplementary materials. LP contributed to the writing of the manuscript.

References

Aharony, N., Pan, W., Ip, C., Khayal, I., and Pentland, A. (2011). Social fMRI: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*, 7(6), 643-659.

Android. (2018a). Android Features and APIs Retrieved from URL: <https://developer.android.com/about/versions/pie/android-9.0>

Android (2018b). Services Overview Retrieved from URL: <https://developer.android.com/guide/components/services>

Apple (2016). Apple Researchkit Retrieved from URL: <http://researchkit.org>

Apple (2017) *App store*. Retrieved from URL: <https://www.appstore.com/>

Carlson, J. A., Jankowska, M. M., Meseck, K., Godbole, S., Natarajan, L., Raab, F., ... and Kerr, J. (2015). Validity of PALMS GPS scoring of active and passive travel compared to SenseCam. *Medicine and Science in Sports and Exercise*, 47(3), 662-667.

Carter, S., Mankoff, J, and Heer, J. (2007). Memento: support for situated ubicomp experimentation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 125-134). ACM.

Canzian, L., and Musolesi, M. (2015). 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* (pp. 1293-1304). ACM.

Chorley, M. J., Whitaker, R. M., and Allen, S. M. (2015). Personality and location-based social networks. *Computers in Human Behavior*, 46, 45-56.

Falaki, H., Mahajan, R., and Estrin, D. (2011). SystemSens: a tool for monitoring usage in smartphone research deployments. In *Proceedings of the sixth international workshop on MobiArch* (pp. 25-30). ACM.

Ferreira, D., Kostakos, V., and Dey, A. K. (2015). AWARE: mobile context instrumentation framework. *Frontiers in ICT*, 2, 6.

Fitzpatrick, C., Birnholtz, J., and Gergle, D. (2016). People, places, and perceptions: effects of location check-in awareness on impressions of strangers. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 295-305). ACM.

Gadziński, J. (2018). Perspectives of the use of smartphones in travel behaviour studies: Findings from a literature review and a pilot study. *Transportation Research Part C: Emerging Technologies*, 88, 74-86.

Google (2017a) *Play store*. Retrieved from URL: <https://play.google.com/store?hl=en>

Google [Computer software] (2017b) *FusedLocationProviderAPI* Retrieved from URL: <https://developers.google.com/android/reference/com/google/android/gms/location/FusedLocationProviderApi>

Google (2018). Google Timeline. Retrieved from: <https://www.google.com/maps/timeline>

Guha, S., and Birnholtz, J. (2013). Can you see me now?: location, visibility and the management of impressions on foursquare. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services* (pp. 183-192). ACM.

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

Huang, Y., Xiong, H., Leach, K., Zhang, Y., Chow, P., Fua, K., ... and Barnes, L. E. (2016). Assessing social anxiety using GPS trajectories and point-of-interest data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 898-903). ACM.

- James, P., Jankowska, M., Marx, C., Hart, J. E., Berrigan, D., Kerr, J., ... and Laden, F. (2016). "Spatial energetics": integrating data from GPS, accelerometry, and GIS to address obesity and inactivity. *American Journal of Preventive Medicine*, 51(5), 792-800.
- Jankowska, M. M., Schipperijn, J., and Kerr, J. (2015). A framework for using GPS data in physical activity and sedentary behavior studies. *Exercise and Sport Sciences Reviews*, 43(1), 48-56.
- Lathia, K.K. Rachuri, C. Mascolo, P.J., and Rentfrow Contextual dissonance: Design bias in sensor-based experience sampling methods. In *Proceedings of the 2013 ACM international joint conference on pervasive and ubiquitous computing* (pp. 183-192), ACM.
- Müller, S. R., Harari, G. M., Mehrotra, A., Matz, S., Khambatta, P., Musolesi, M., ... and Rentfrow, P. J. (2017, September). Using human raters to characterize the psychological characteristics of GPS-based places. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (pp. 157-160). ACM.
- Noë, N., Whitaker, R. M., Chorley, M. J., and Pollet, T. V. (2016). Birds of a feather locate together? Foursquare checkins and personality homophily. *Computers in Human Behavior*, 58, 343-353.
- OpenPaths (2017). About OpenPaths. Retrieved from: <https://openpaths.cc/about>
- Palmius, N., Tsanas, A., Saunders, K. E. A., Bilderbeck, A. C., Geddes, J. R., Goodwin, G. M., and De Vos, M. (2017). Detecting bipolar depression from geographic location data. *IEEE Transactions on Biomedical Engineering*, 64(8), 1761-1771.
- Pizarro, A. N., Schipperijn, J., Ribeiro, J. C., Figueiredo, A., Mota, J., and Santos, M. P. (2017). Gender Differences in the Domain-Specific Contributions to Moderate-to-Vigorous Physical Activity, Accessed by GPS. *Journal of Physical Activity and Health*, 14(6), 474-478.
- Piwek, L., Ellis, D. A., and Andrews, S. (2016). Can programming frameworks bring smartphones into the mainstream of psychological science?. *Frontiers in psychology*, 7, 1252.
- Ramanathan, N., Alquaddoomi, F., Falaki, H., George, D., Hsieh, C. K., Jenkins, J., ... and Tangmunarunkit, H. (2012). Ohmage: an open mobile system for activity and experience

sampling. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, 2012 6th International Conference on (pp. 203-204). IEEE.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., ... and Funder, D. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, 107(4), 677.

Runyan, J. D., Steenbergh, T. A., Bainbridge, C., Daugherty, D. A., Oke, L., and Fry, B. N. (2013). A smartphone ecological momentary assessment/intervention “app” for collecting real-time data and promoting self-awareness. *PLOS ONE*, 8(8), e71325.

Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., and Mohr, D. C. (2016). The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ*, 4, e2537.

Sandstrom, G. M., Lathia, N., Mascolo, C., and Rentfrow, P. J. (2017). Putting mood in context: Using smartphones to examine how people feel in different locations. *Journal of Research in Personality*, 69, 96-101.

Schmidt, T., Kerr, J., Kestens, Y., and Schipperijn, J. (2018). Challenges in using wearable GPS devices in low-income older adults: Can map-based interviews help with assessments of mobility?. *Translational Behavioral Medicine*, iby009.

Schwartz, R., and Halegoua, G. R. (2015). The spatial self: Location-based identity performance on social media. *New Media & Society*, 17(10), 1643-1660.

Van Kleek, M., Liccardi, I., Binns, R., Zhao, J., Weitzner, D. J., and Shadbolt, N. (2017, May). Better the devil you know: Exposing the data sharing practices of smartphone apps. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5208-5220). ACM.

Wagner, D. T., Rice, A., and Beresford, A. R. (2014). Device Analyzer: Large-scale mobile data collection. *ACM SIGMETRICS Performance Evaluation Review*, 41(4), 53-56.

Wilcockson, T. D., Ellis, D. A., and Shaw, H. (2018). Determining typical smartphone usage: What data do we need?. *Cyberpsychology, Behavior, and Social Networking*, 21(6), 395-398.