Heliyon 9 (2023) e15751

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Evaluation of a smartphone-based methodology that integrates long-term tracking of mobility, place experiences, heart rate variability, and subjective well-being

Matteo Giusti *, Karl Samuelsson

Department of Building Engineering, Energy Systems and Sustainability Science, University of Gävle, Sweden

ARTICLE INFO

Keywords:
Smartphone methodology
Longitudinal research
GPS
Heart rate variability
PPGIS
Wellbeing

ABSTRACT

This study presents MyGävle, a smartphone application that merge long-term tracking of mobility data, heart rate variability and subjective and objective well-being records. Developed to address the challenges faced in researching healthy and sustainable lifestyles, this app serves as a pioneering implementation of Real-life Long-term Methodology (ReaLM). After eight months' use by 257 participants from Gävle (Sweden), we evaluate the completeness, accuracy, validity, and consistency of all data collected. MyGävle produced remarkable results as a ReaLM method. On average, it precisely tracked participants daily locations for approximately 8 h and accurately collected heart-rate variability values throughout the day (12 h) and night (6 h). Participants reported 5115 subjective place experiences (ranging from 160 to 120 per week) and seasonal participation, although declining, is accurate. Our findings indicate that the amount of data collected through smartphone sensors, fitness wristbands and in-app questionnaires is consistent enough to be leveraged for integrated assessments of habits, environmental exposure, and subjective and physiological well-being. Yet, considerable variation exists across individuals; thus diagnostic analysis must precede use of these datasets in any particular research endeavors. By doing so we can maximise the potential of ReaLM research to delve into real life conditions conducive to healthy living habits while also considering broader sustainability goals.

1. Introduction

Understanding and promoting healthy and sustainable lifestyles is a key challenge for humanity. This goal requires transparent, replicable, and unbiased data that can simultaneously explore the drivers of healthy and of sustainable lifestyles. These data thus need to link socio-economic factors and real-life environmental exposures (e.g., to pollutants and nature) with salutogenic and pro-environmental behaviours [1–3]. Today, ubiquitous data generation offers opportunities to achieve this, but meeting academic requirements for transparency, privacy, and longitudinal validity remains a challenge.

Research has benefited immensely from the Big Data revolution. Studies on sustainable urban environments using commercially generated Big Data have increased rapidly in recent years [4]. However, secondary Big Data poses challenges in terms of representativeness and can reinforce existing cultural and societal biases [5]. For example, Twitter data is increasingly used as a tool for health research [6], despite being used by an unrepresentative group of people (11% in the US – the country with the highest usage rate - with

* Corresponding author.
E-mail address: Matteo.giusti@hig.se (M. Giusti).

https://doi.org/10.1016/j.heliyon.2023.e15751
Received 21 November 2022; Received in revised form 6 April 2023; Accepted 20 April 2023
Available online 27 April 2023

38.5% of users aged 25–34 and 33% with at least a university degree, only 18% of users from rural areas) [7,8]. Commercial Big Data also brings challenges in terms of validity [8,9]. For example, Google Flu Trends cannot distinguish between people performing flu-related Google searches due to illness and other reasons. These results also rely on undisclosed and proprietary algorithms that are at odds with scientific transparency and replicability [10]. Research with sensitive data also requires protection of participants' privacy, a common problem with commercial data [9]. By comparison, primary Big Data is rare in academia but has the potential to transform research in health [9,11] and sustainable living [12,13], while helping to create the policy-making and governance structures needed for sustainable and healthy improvements [14]. For these reasons, scientific research into healthy and sustainable lifestyles requires Big Data methods and tools different from those available on the market.

Smartphone applications based on *Ecological Momentary Assessments* (EMA) or *Geographically-explicit* EMA (GEMA) methods (together hereafter referred to as [G]EMA) are increasingly used in science to temporarily record behaviour or well-being [15]. [G] EMA aims to assess people's everyday thoughts, experiences or behaviours in a real-life setting (i.e., *ecological*), as close to real-time as possible (i.e., *momentary*), and possibly with the associated geographic location (i.e., the G for *geographically-explicit*). [G]EMA offer several advantages when applied in the health sciences [16]. First, momentary assessments reduce recall errors, such as people being more likely to retrieve negative attributes, memories, or opinions when they are in a negative mood [17]. Second, [G]EMA recognise that many behaviours and emotions are influenced by - or emerge from – a physical and social context that cannot be adequately replicated in the laboratory. For the assessed experience to be an ecologically valid observation, it must be collected in the contexts in which it naturally occurs [16,18]. Finally, [G]EMA allow for fine temporal resolution in longitudinal studies. Repeated ecological measurements allow researchers to detect variations within a subject that occur over time or in different contexts. However, [G]EMA have also well-known limitations. Collecting data as close as possible to real-time events places a significant burden on participants as they go about their daily lives, and many studies struggle to achieve acceptable compliance rates [19]. In one case, reported compliance was as high as 90%, but actual compliance was estimated at 11% [18]. Participant adherence and response accuracy increase in places where people can be inactive (e.g., home, restaurant, friends' house) or when participants are already using the smartphone, and even then response latency can reach 25 min [20]. Despite relatively short studies (2–42 days), problems with compliance and delays limit analysis of real-time experiences that take place on the move, outside familiar locations, or in situations where participants are unable or unwilling to use their phones [20]. Therefore, [G]EMA studies must be carefully designed to minimise participant burden while ensuring temporal and ecological validity, especially if the study spans over extended periods of time [21].

In this paper, we build on previous efforts [22] and present a research-driven smartphone application capable of generating valid data over long and continuous periods of time to study healthy and sustainable lifestyles. This application is called MyGävle (as our case study is the municipality of Gävle in Sweden) and it is the first implementation of a methodological approach that we term Real-life Long-term Methodology (ReaLM). MyGävle integrates spatio-temporal tracking (via smartphone's sensors), biometrics and heart rate variability data (via in-app connection to Garmin fitness wristbands) and subjective well-being (via in-app surveys) while being minimally invasive for the participants. Below we provide an overview of the methodology behind MyGävle and the validation analysis after using the app for eight months with participants living in Gävle. We report on the quality of the data collected before discussing improvements for further development of ReaLM methods.

2. Overview of MyGävle

MyGävle is a native software application for smartphones that participants voluntarily install on their smartphones (see screenshots in fig. A.1 and A.2). MyGävle was developed as part of the research project BIG (Better quality of life with Integrated Geospatial data; see https://fpx.se/en/big-project-eng/for more details). The case study of our data collection is the municipality of Gävle (Sweden).

MyGävle integrates four different methods of data collection. First, it uses GPS and activity sensors in smartphones to track participants' movement and visits to places. Second, MyGävle connects to fitness wristbands (Garmin vìvosmart 4) that are distributed to participants. The optical heart rate sensor in the fitness wristband is used to monitor personal biometric data and heart rate variability (HRV), an important physiological marker of health [23]. Thirdly, a Public Participation Geographical Information System (PPGIS) survey is used to capture subjective experiences of the visited landscape (hereafter place experiences). Finally, in-app seasonal questionnaires are used to monitor changes in indicators of interest to the study over long periods of time. For the project BIG, these surveys relate to human flourishing, human-nature relationships, social environmental norms and social capital.

By combining these methods, each participant generates a large amount of data. The spatio-temporally coded GPS measurements are the fulcrum of this heterogeneous dataset, to which all other data are linked and analysed together. All biometric and HRV data are temporally defined, while the PPGIS questionnaires are geo-coded so that both types of data can be linked to specific locations, activities, or experiences. Seasonal variations can also be evaluated in relation to geographical exposures or biometric information. All these data can then be linked to the ecological and geographical characteristics of a landscape, which is the main objective of the overarching research project BIG, although this cannot be addressed here.

A guiding principle in the development of MyGävle was to combine ecological and temporal validity with non-invasiveness to overcome known shortcomings of [G]EMA (see Introduction). To achieve this, we asked participants to report PPGIS questionnaires when it was convenient for them, using survey alerts (Figure A.1a), we selected a small and appealing fitness wristband with considerable battery life, we limited the recording of GPS locations to minimise smartphone battery consumption, we provided in-app maps with participants' data (Figure A.1c) and gamified the user experience (Fig. A.1d).

3. Methods of data collection in MyGävle

3.1. Smartphone GPS and activity sensors

We use the GPS sensors built into participants' smartphones to continuously track their location. To limit battery consumption, we use the phone's accelerometer to determine when the phone is in motion. GPS records are stored when the transition from still to moving occurs, every 10 s or 50 m when the phone is in motion (whichever comes first), and when the transition from moving to still occurs. Each GPS record contains a margin of error (MOE) in metres. We use the Activity Recognition Application Programming Interface (API) (for Android) and the CoreMotion API (for iOS) for activity recognition. These use the phone's accelerometer and gyroscope to distinguish between six activity types: stationary, walking, running, cycling, in a vehicle or unknown.

3.2. Fitness wristbands

A subset of all participants (see section 4.1) agreed to use a Garmin vìvosmart 4 fitness wristband that was offered at a discounted price, thus contributing to the study with additional biometric data. Participants who already owned a Garmin fitness wristband could connect it to MyGävle. The biometric data is collected via the Garmin API. As HRV data is not available through the Garmin API, participants used a special module developed using the Garmin Companion SDK to connect the wristband directly to MyGävle via Bluetooth. Participants with a Garmin fitness wristband received an email and two reminders asking them to manually set up this connection in MyGävle. The app connects to the fitness wristband every few hours (depending on phone model, operating system, and activity), whereupon all newly recorded heart interbeat intervals since the last connection are sent to University of Gävle's servers.

3.3. Seasonal questionnaire

At the beginning of each season, participants are asked to complete a questionnaire (see Table B.2 for the full questionnaire). This seasonal questionnaire was sent to participants in September 2021, January 2022, April 2022, July 2022, and November 2022. This seasonal questionnaire contains 23 items: six for human flourishing based on [24], nine for human-nature relationships based on [25], three for social environmental norms based on [26] and five for social capital based on [27,28]. Each item is a statement to which the participant responds using a colour coded 7-point Likert scale from "strongly disagree" to "strongly agree". Participants have the option to defer answering the seasonal questionnaire a maximum of three times while continuing to use the app.

3.4. Public participatory GIS survey

In MyGävle there is a fixed prompt ("mark a place where you had a recent experience") at the top of the home screen. This is the first question of the PPGIS questionnaire, which asks about subjective place experiences (Fig. A.1a). The rest of the screen shows an interactive map centred around the participants' current location, with an overlaid heat map of the participant's most recent GPS locations (i.e., the last seven days). Participants report on their experiences by sliding the map to place the pin icon at the location of the experience. At the bottom of the home screen is a red button labelled "Bad" and a green button labelled "Good", which participants can use to start reporting the experience. Tapping one of these buttons takes you to a PPGIS questionnaire with 15 questions (see Table B.3 for the full questionnaire). Participants were allowed to report a maximum of three experiences per week.

4. Study procedure

4.1. Participant onboarding, marketing, and incentives

As of 25 July 2021, participants could download MyGävle for free from Google Play and Apple Store. When MyGävle was first opened, the potential participant was guided through a series of introductory screens (see Fig. A.2). The criteria for participation in the study were that the person lived in Gävle, spoke Swedish or English, was between 15 and 79 years old, and did not suffer from dementia or cardiological diseases or was otherwise impaired from accessing the outside environment of Gävle municipality. By giving consent to participate, the participant was enrolled in the *public study*. After giving consent and granting the necessary permissions (permission to track location and activity sensors and to receive notification prompts; Fig. A.2d), the participant answered a questionnaire to collect socio-demographic data (e.g. gender, age, education and income levels) (see Table B.1 for more details) (Fig. A.2e).

After enrolling in the *public study,* participants were invited to take part in the *health study,* where they provided their biometric data by wearing a Garmin fitness wristband. Participants were given additional information about the collection and use of biometric data and were able to give consent and enter the health study through MyGävle. Participants were asked to wear the fitness wristband regularly, including during sleep.

In late July and early August in 2021, information about the research project and MyGävle was disseminated by University of Gävle via the university website and social media. From late August to early December, sponsored Facebook posts, articles in local newspapers, printed ads on the university campus, and digital ads in the city's ice hockey arena were used to reach different groups of participants. This marketing campaign capitalised on the intrinsic motivation to participate in the project as a citizen scientist. In addition, financial incentives included participation in a raffle of five electric bicycles that took place at the end of the study in December 2022 and receiving a Garmin vìvosmart 4 for 99 SEK instead of 1.199 SEK, which participants could keep after completing the study.

In short, MyGävle allowed for a multi-layered engagement with the research study so that each participant could choose the level of engagement they felt is non-invasive in their daily lives.

4.2. Study timeline

Data collection began on September 1st' 2021, and ended on November 30th' 2022 (a total of 15 months or five seasons). During this period, participants provided data continuously and at weekly and seasonal intervals (see Fig. 1). The HRV recording function was added to the app on February 10th' 2022.

4.3. Ethical considerations

All data generated from MyGävle was transferred and encrypted on servers at the University of Gävle. The app was pilot tested among a group of university staff over a period of about two months in the spring of 2021. The entire research design was reviewed and approved by the Swedish Ethical Review Authority (application number 2021–02212).

5. Data assessment

5.1. Data quality attributes

To assess data quality, we select data from the first eight months of data collection (September 2021 to April 2022). We assess four main characteristics that define data quality: Completeness, accuracy, validity and consistency [29,30]. Completeness refers to the extent of missing data. Assessing data completeness is critical for ReaLM because researchers want to have minimal impact on participants' daily lives, which presents challenges for compliance and long-term retention. Data accuracy refers to the Margin of Error (MOE) or signal-to-noise ratio. This attribute refers to how accurately the data reflect the event or item that they are intended to measure or describe. Validity of dataset can vary depending on the context in which they are applied. For example, the HRV data from consumer devices, such fitness wristbands used here may be suitable environmental epidemiology [31,32], yet may not fulfil clinical healthcare requirements [33]. Consistency of information refers to the absence of internal contradictions and corroboration by other data sources. The ReaLM project serves as an example, whereby potential lapses in precision could occur due to the real-life conditions of data collection. However, this may be counteracted by combining a variety of sources during the analysis process which helps strengthen each dataset's validity.

5.2. Participant samples

Because MyGävle allowed multilevel participation in the research study, we analyse both the *public sample,* i.e., everyone who used MyGävle, and the *health sample*, i.e., those who also used the fitness wristband (see 4.1). In addition to analysing data quality, we assess

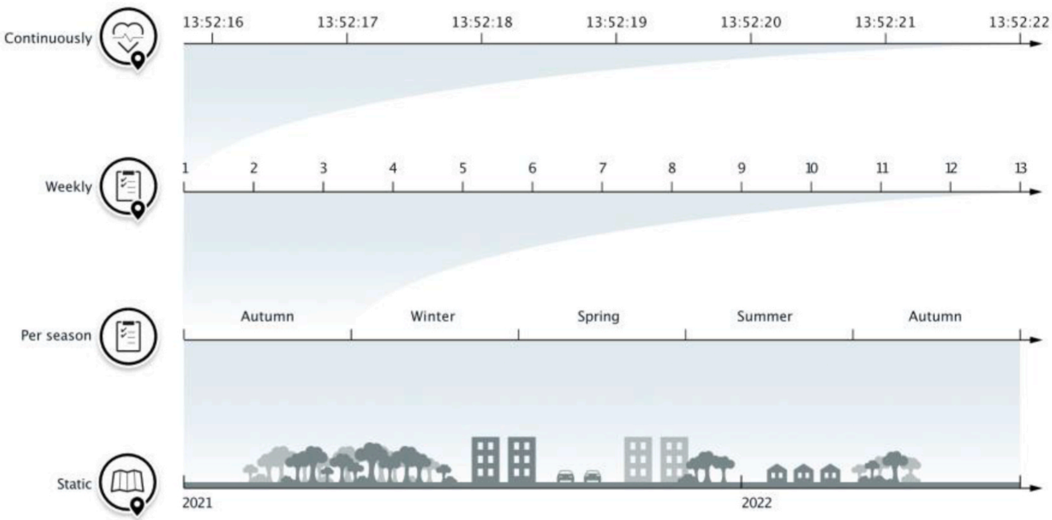

Fig. 1. Timing of data collection and temporal resolution of the various data sources collected. From top to bottom, the figure shows that GPS and heart rate sensors collect data continuously, subjective place experiences are collected almost weekly via PPGIS questionnaires, while seasonal questionnaires are answered once per season. The bottom row of the figure shows that all of this data can be connected to geospatial data about physical and social context and analysed together.

the retention rates (i.e., the extent to which participants remain in the study) and sociodemographic representativeness for the population of Gävle.

5.3. Spatio-temporal tracking

Completeness and accuracy of spatio-temporal data are closely related: a dataset may be complete at a certain level of resolution, while incomplete at a higher level of resolution. For spatio-temporal tracking, we therefore perform the completeness and accuracy analysis in conjunction. We consider GPS measurements with a MOE of 25 m or less as "high accuracy" and those with a larger MOE as "low accuracy" and calculate the number of high and low accuracy GPS points per participant per day. Based on this classification, we divide each participant's daytime hours (6:00–23:00) into four classes of temporal coverage: high accuracy, low accuracy, likely missing, and missing (see Appendix E for the details of this estimation). As a validity analysis, we calculate the spatio-temporal coverage for each day period of each participant and for different activity types (as recorded by the phone's sensors). These analyses provide an indication of appropriate applications for the data. For the health sample, we evaluate the consistency between the phone's location tracking and the fitness wristbands by comparing the summed deviation between the daily records from GPS and the total step count recorded by the fitness wristband for the same day. To obtain a measure of the distance traveled by steps, only GPS records with walking and running activity types and those immediately preceding or following were included. To exclude days when either the phone or fitness wristband was not used, we only consider days when 1) at least one GPS record and 2) fitness wristband measurements were taken on at least 16 of the 24 h of the day (to ensure that they overlapped with most hours of the day). We calculate Pearson correlation coefficients between step counts and distance between GPS records for each day and each participant, and separately for each participant who had at least 14 days of shared step counts and GPS records. Kruskal-Wallis and chi-square tests and beta regression are used to examine whether consistency is predicted by distance between GPS records, fitness wristband step count, phone operating system, and participant being in the health sample.

5.4. Heart rate variability records

We assess HRV data completeness as the number of participants in the health sample who connected their fitness wristband to MyGävle and the average temporal coverage of this data collection. Given the known limitations in accuracy of commercial heart rate monitors for non-stationary conditions (Georgiou et al., 2018; Tedesco et al., 2019) we separate the analysis for daytime (6:00–23:00) and night-time (23:00–6:00). For the accuracy analysis, we select a random day and a random night for each participant and calculate the proportion of observations that are classified as errors for each of them following a procedure described by Lipponen and Tarvainen [34] (see Appendix F for details). As a validity analysis, we calculate the temporal coverage for each day and night for each participant. Similar to the validity analysis of the spatio-temporal tracking, this provides an indication of appropriate applications for the data. Because we use step the fitness wristbands step counts to analyse consistency with location data, we do not perform additional consistency analysis.

5.5. Seasonal questionnaires

We use the number of participants responding the seasonal questionnaire in the public and health samples for each season as a measure of completeness. For accuracy, we measure differences in results between successive surveys for the same questionnaire item and participant, assuming that an accurate survey instrument produces many identical values or small changes, and few large changes. Ensuring questionnaire validity and consistency is critical for survey measurements but is beyond the scope of this paper because we adopted questionnaires that have already been published and validated.

5.6. Subjective place experiences

We calculated the number of completed PPGIS questionnaires per week during the study period and use the weekly proportion of participants reporting at least one weekly experience as a measure of completeness. As a measure of accuracy, we calculate the average time to complete the survey per week, which is equivalent to what Brown [35] refers to as coverage effort, an important aspect of PPGIS data quality. As a measure of validity, we compare the words that participants used to describe their place experiences in autumn (Sep–Nov) and winter (Dec–Feb). For all words used by at least five participants in either season, we calculate differences in occurrence between seasons and plot particularly seasonal words in word clouds. Consistency is assessed using location data from GPS by calculating the proportion of reported experiences that have spatially and temporally adjacent GPS records. For each experience and for a set of spatially circular buffer areas (25- to 150-m radius in 25-m increments), we calculate, first, whether there was at least one GPS record in the week preceding taking the PPGIS questionnaire, and second, for how many days in the preceding week there was at least one GPS record.

6. Results

6.1. Participants samples

From July 25, 2021 to April 30' 2022, 271 Gävle residents downloaded MyGävle and accepted the consent form. Of these, 14 left the study, leaving 257 participants in the *public sample*. Of these, 149 agreed to wear a fitness wristband; these comprise the *health sample*. Recruitment was fastest from early August to mid-October, while most dropouts occurred in the second half of October and November (Fig. C.1). The samples did not differ significantly on demographic and socioeconomic variables (Table 1). Both samples had a slight overrepresentation of women and persons aged 30–59 years. Individuals with higher education, high income, and employment were disproportionately represented in both samples.

6.2. Spatio-temporal tracking

A total of 25.9 million GPS records were collected during the eight-month period. The daily number of GPS measurements recorded varied considerably among participants, from 3.75 (mean for the seven participants with the fewest measurements) to 2670 (mean for the five participants with the most measurements) (Fig. 2A). On average, 89.3% of a participant's GPS recordings were highly accurate (MOE $<$ 25 m). The diurnal hours covered by the GPS records (6:00–23:00) ranged from 3 h to 29 min (20.5%; average of the seven participants with the lowest coverage) to 12 h and 42 min (74.7%; average of the five participants with the highest coverage) (Fig. 2B), with a mean of 7 h and 41 min (45.2%). 65.7% of temporal coverage is highly accurate. Daily coverage is higher among participants in the health sample than among participants outside of the sample (see Appendix E for more details). Daily coverage varies widely across individual days (Fig. 2C); for example, 221 participants (86%) have at least 14 days with at least 2 h of coverage, whereas only 152 participants (59%) have at least 14 days with at least 12 h of daily coverage. Coverage also varies widely across activities (Fig. 2D). Most of the time, the phone is stationary (5 h and 56 min) or the activity is unknown (1 h and 6 min). In motion, people are more likely to be in a vehicle (27 min) or walking (23 min) than bicycling (3 min and 7 s) or running (1 min and 13 s). The proportion of time covered by highly accurate measurements is greater for moving activities (from 83% for being in a vehicle to 97% for running) than for stationary and unknown activities (67% and 57%, respectively).

There is a moderate correlation between the daily number of steps measured with the fitness wristband and the daily sum of distances between GPS records ($\rho = 0.524$; Fig. 2E). Consistency at the individual level varies from excellent for some participants to nonexistent for others and likely depends more on behavioural factors (e.g., how often the participant is indoors or outdoors, or carries the phone) than on fitness wristband use or phone operating system (see Appendix E).

6.3. Heart rate variability

72 out of 149 (48.3%) from the health sample connected their fitness wristband directly to MyGävle to enable HRV data collection. Among these participants, HRV data were collected on an average of 50 days and 45 nights (out of a maximum of 79 possible), for 11.8 h per day and 5.7 h per night (Fig. 3A and B). On average, 11.8% of daytime observations and 9.0% of night-time observations were classified as errors (Fig. 3C and D). Participants who wore fitness wristbands tended to wear them during most hours of the day

Table 1
Descriptive statistics of the samples and comparison with Gävle population.

Participant characteristic	Count (%)	Count (%)	% in Gävle
	Public sample	*Health sample*	
	257 (100)	149 (100)	XXX
Gender			*
Man	114 (44.4)	68 (45.6)	50.3
Woman	143 (55.6)	81 (54.4)	49.7
Age			*
15–29	37 (14.4)	19 (12.8)	23.8
30–39	59 (23.0)	29 (19.5)	16.9
40–49	54 (21.0)	31 (20.8)	15.7
50–59	61 (23.7)	41 (27.5)	16.7
60–69	31 (12.1)	19 (12.8)	13.8
70–79	13 (5.06)	10 (6.71)	13.0
Education			†
Elementary or secondary	84 (32.7)	54 (36.3)	63.7
Bachelor's degree	126 (49.0)	68 (45.6)	NA
Master's or doctoral degree	47 (18.3)	27 (18.1)	NA
Monthly income			†
Less than 20 000 SEK	50 (19.4)	27 (18.1)	41.3
20 000–29 999 SEK	38 (14.8)	27 (18.1)	23.3
30 000–49 999 SEK	136 (52.9)	80 (53.7)	28.9
50 000 or more	33 (12.9)	15 (10.1)	6.52
Employment			‡
Employed or self-employed	202 (78.6)	119 (79.9)	54.6
Retired	23 (8.95)	17 (11.4)	29.5
Other	32 (12.5)	13 (8.72)	NA

* in 2021, population aged 15-79.
† in 2020, population aged 16 and older.
‡ in 2021, population aged 15–74, in Gävleborg region.

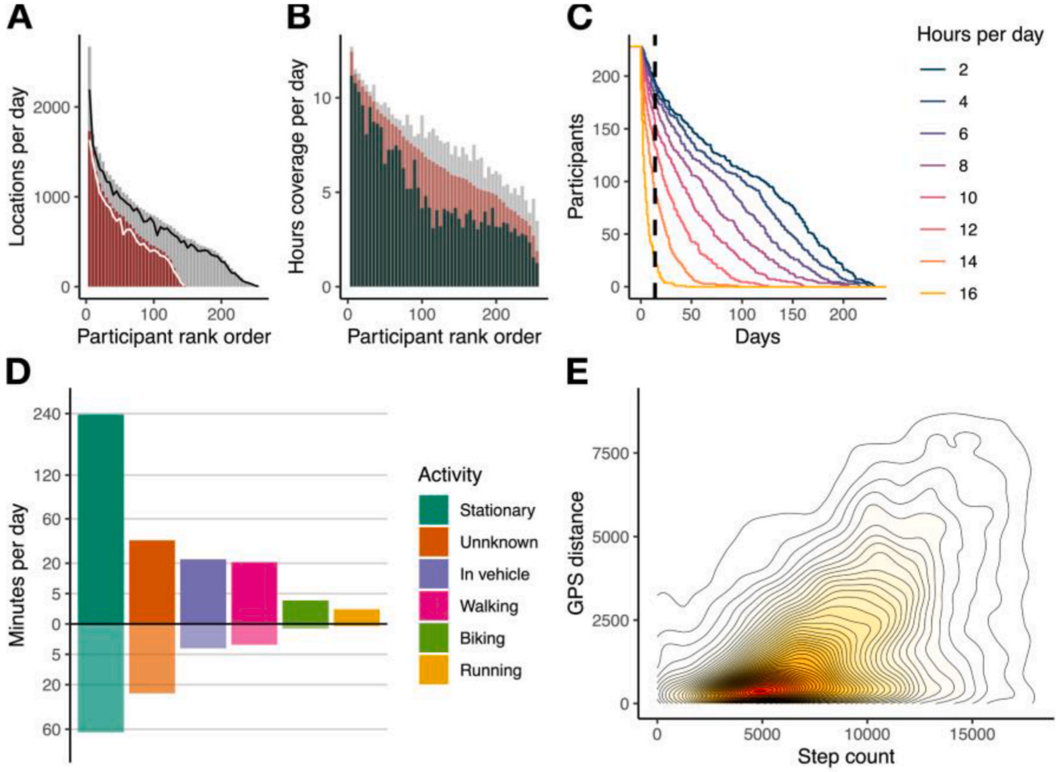

Fig. 2. (A) GPS measurements per day in the public sample (grey bars) and the health sample (red bars). Each bar represents five participants, ordered, and subdivided by the number of daily locations. The lines show the number of highly accurate records (public sample: black, health sample: white). (B) Hours of daily GPS coverage per day in the public sample (dark green = high accuracy coverage, red = low accuracy coverage, grey = likely missing). Each bar represents five participants, ordered, and subdivided by combined daily high accuracy and low accuracy coverage. (C) Cumulative frequency distribution of participants with days with daytime coverage above a range of thresholds (2–16 h). The dashed line is placed at 14 days, which is sufficient for activity area mapping (Zenk et al., 2018). (D) Mean daily distribution of daily activity recorded by GPS. Bars running up from 0 represent high accuracy measurements, while semi-transparent bars running down represent low accuracy measurements. (E) Consistency plot between individual-level daily fitness wristband step count (x axis) and daily calculated sum of distances between GPS records (y axis). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(Fig. 3E) and virtually all hours of the night (Fig. 3F). Only very high thresholds for HRV data during the day significantly reduced the number of days (Fig. 3E).

6.4. Seasonal questionnaire

The number of participants responding the seasonal questionnaire decreased from 178 in autumn (69.3%) to 146 in winter (56.8%) and further to 113 in spring (44.0%) (Fig. 4A). Response rates were lower and declined more dramatically among those outside the health sample (from 46.3% to 21.3%–7.4%) than among those who were in the sample (from 85.9% to 82.6%–70.5%). The questionnaire shows signs of good accuracy, with mostly identical values or little change between successive surveys (Fig. 4B and C).

6.5. Subjective place experiences

The PPGIS questionnaire recorded 5115 subjective experiences of place during the study period, with the number of weekly responses decreasing from about 160 per week at the beginning of the study to about 120 per week at the end of the spring (Fig. 5A). The weekly response rate was about 40% at the beginning and decreased over the study period to about 30% for participants in the health sample and to about 5% for participants outside this sample (Fig. 5B). The average questionnaire completion time decreased from about 120 s to about 90 s (Fig. 5C). Participants used particularly seasonal words to describe place experiences point. In autumn, common words were “vatten” (water), “natur” (nature), “lugn” (calm), “höst” (autumn), “skog” (forest), “träning” (exercise), “motion” (exercise) and “cykelväg” (bike road) (Fig. 5D). In winter, common words were “snö” (snow), “promenad” (walk), “is” (ice), “familj” (family), “halt” (slippery), and “vinter” (winter) (Fig. 5E). The proportion of place experiences with adjacent GPS records within the previous week ranged from 65% (within a 25-m circular buffer) to 76% (within a 150-m circular buffer) (Fig. 5F). The number of days within the previous week that participants were adjacent to the location of the place experience ranged from 1.5 (on average within a

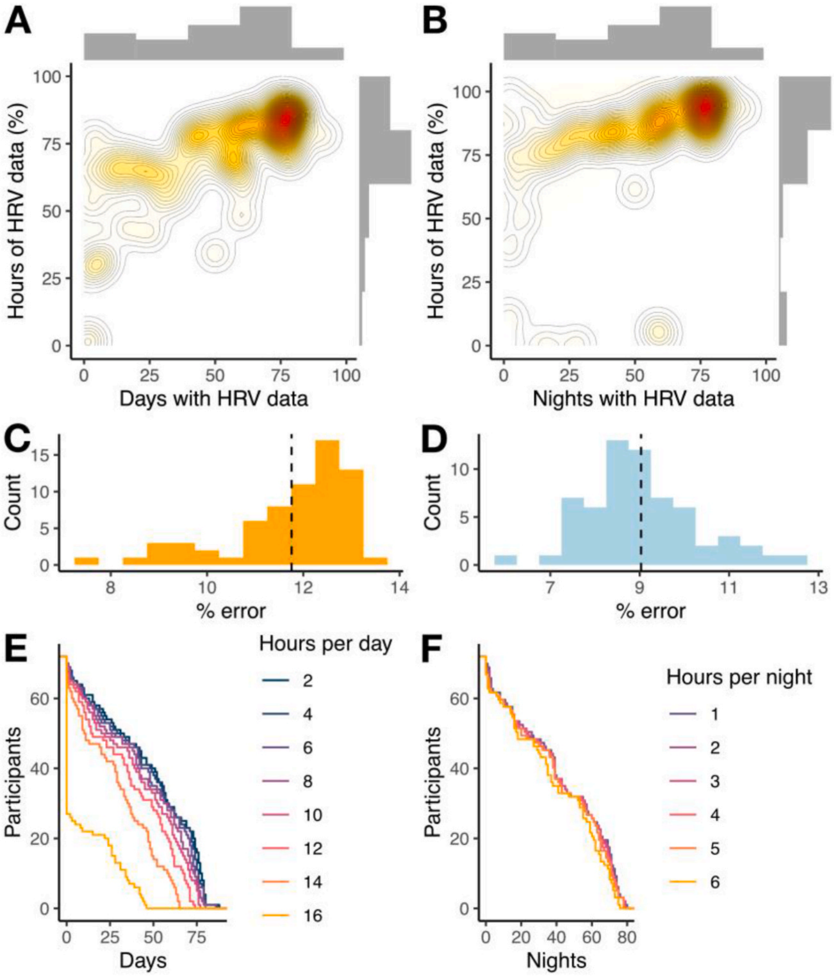

Fig. 3. (A) Density plot showing marginal histogram of days with HRV data and average number of hours of with HRV data during the day per participant, excluding participants without HRV data. (C) Percentage of observations classified as erroneous (one random day sampled for each participant with HRV data), following Lipponen and Tarvainen (2019). Dashed line is at the mean. (E) Cumulative frequency distribution of participants with days with HRV data measured during the day above thresholds (2–16 h). (B), (D), (F) like (A), (C), (E) but for HRV data collected at night.

25-m circular buffer) to 2.1 (on average within a 150-m circular buffer).

7. Discussion

7.1. Participation

Despite a diverse marketing strategy aimed to at reaching different demographic groups, both samples were overrepresented in terms of gender (women), age (30–59), high education, high income, and employment. This is not entirely surprising given that the less affluent and elderly are underrepresented among smartphone and fitness wristband users [36]. This study attempted to overcome such selection biases by offering fitness wristbands at a greatly reduced price while remaining open to participants who already owned one. However, the difficulty in reaching specific populations and motivating them to participate proved to be an obstacle to achieving a higher degree of representativeness.

Over an eight-month period, there was a decline in participation in both the seasonal questionnaire (approximately 37%) and the PPGIS questionnaire (approximately 25%), possibly due to a decrease of the novelty of the study, intrinsic motivation to participate, and appreciation for receiving the fitness wristband. Nevertheless, participants were not trained, monitored, or assisted, except for some initial notifications (which participants could also decline), so the fact that most of them were still actively participating after 8 months without being reminded is encouraging. Of note, half of the participants in the health sample did not connect their Garmin to MyGävle, suggesting that the process was perceived by many as either too complex or too demanding. Once again, this demonstrates that any requirement for active participant engagement comes at a significant compliance cost. Nonetheless, using passive monitoring

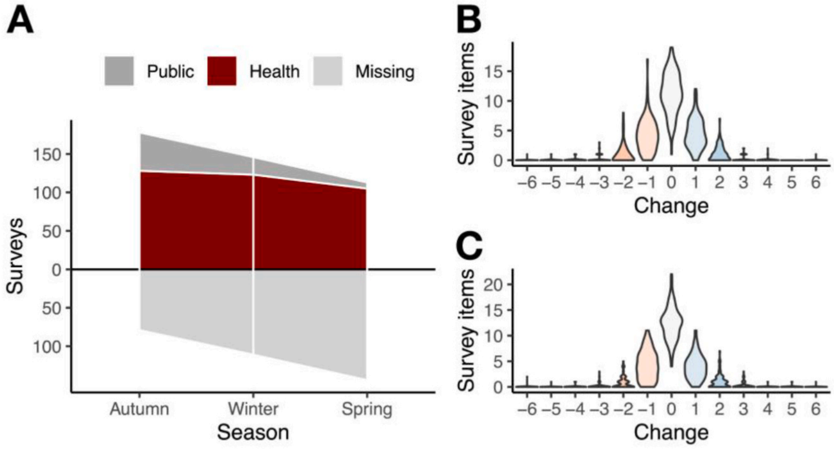

Fig. 4. (A) Number of participants who responded to answering the seasonal questionnaires, broken down by sample, and number of missing surveys. (B) Number of survey items per participant by magnitude of change in Likert scale scores (1–7), from autumn to winter. Violin plots show the density distribution of participants. For example, many participants answered identically on 10–15 items, and each had about five items that increased or decreased by 1. (C) Like (B) but from winter to spring.

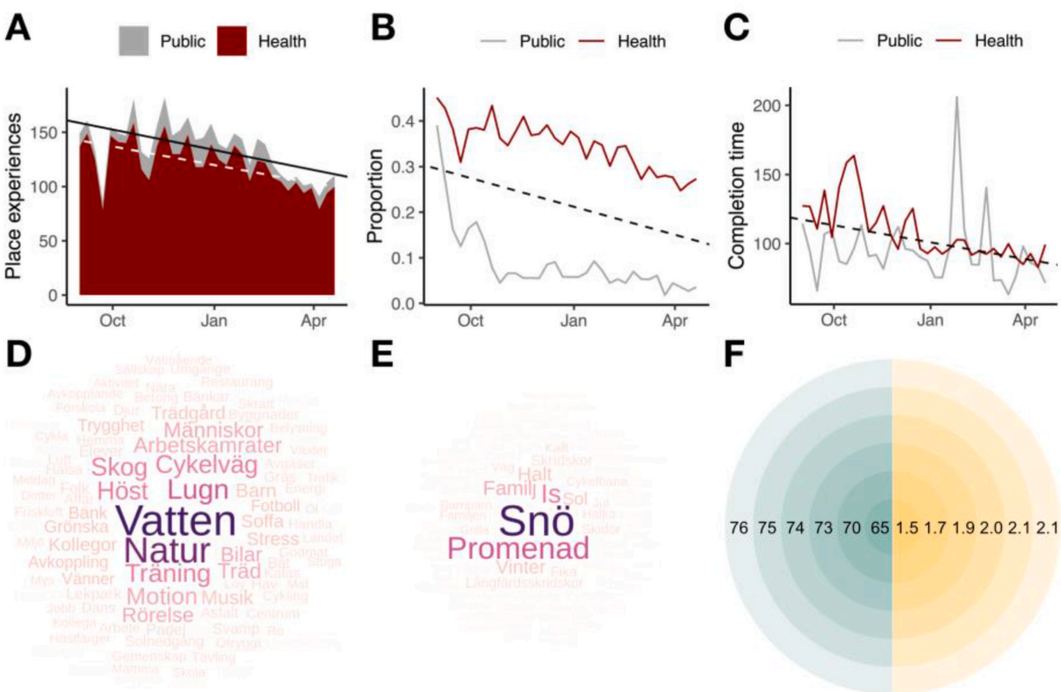

Fig. 5. (A) Number of place experiences recorded by the PPGIS questionnaire divided by sample and week. The lines show fitted linear trends (black for the public sample, white dashed for the health sample). (B) Weekly proportion of participants recording at least one place experience. The line shows the fitted linear trend for the public sample. (C) Time (seconds) to complete of PPGIS questionnaires for each week. Line shows fitted linear trend for the public sample. (D) Words that participants used to describe their place experience during autumn (Sep–Nov) and during winter (Dec–Feb). (F) Consistency between GPS records and PPGIS questionnaires. Each concentric circle represents a spatial buffer of 25 m (total of 150 m). On the left side, the numbers represent the percentage of place experiences which have at least one GPS point in the previous seven days from the same participant. On the right side, the numbers represent the average number of days with at least one GPS point during the previous seven days from the same participant.

we were able to collect valuable mobility and HRV data without interfering with participants' daily lives. Our results suggest to that this type of integrated data collection may be at the heart of future human-centred, non-invasive, longitudinal tracking of mobility and physiological health.

In short, ReaLM methods such as MyGävle need to consider the socio-demographic differences of the participants in order to arouse and maintain their motivation and enthusiasm. The frequency of questionnaires impacts participant compliance and needs to be matched with the level of disruption in people lives. Technological solutions that enable for passive data collection may be invaluable for future ReaLM methods.

7.2. Data quality and research applications

Even with the preliminary analytical procedures presented here, we argue that the quality of the data collected is generally sufficient to map activity spaces [37], identify spatial habits [38], conduct landscape exposure assessments [39], and contribute to understanding health-enhancing lifestyles [40].

The quality of the spatiotemporal data collected varies considerably. In both samples, the accuracy of GPS locations is high (89.3%), especially when in motion. However, the amount of daytime coverage varies widely among participants (from about 3.5 to about 12.5 h), and although coverage spanned eight months, about 60% of participants had at least 14 days with temporal coverage exceeding 12 h. These large differences in completeness are likely due to a combination of factors. One important factor appears to relate to how different operating systems record GPS locations, with iOS phones having lower temporal coverage than Android phones, despite recording significantly more locations (see Fig. E.1B). However, the large differences at the individual level suggest additional influencing factors, such as the level of physical activity or the quality of GPS signals in the environments to which participants are exposed (e.g., GPS signal quality may decrease inside buildings or even outdoors when surrounded by tall buildings). Further analysis is needed to better understand the sources of the missing data. However, despite these data gaps, we found general agreement between the different data sources. We found consistency between distances traveled, calculated using measurements from GPS records, and step counts from fitness wristbands. Analysis of place experiences also confirmed that 70% of the reported experiences were within 50 m of GPS location of the previous week, suggesting ecological and temporal validity of the reported experiences.

The quality of the HRV data collected is consistent with the existing literature [31]. We have no coverage for 30% of daytime (12% mean error) and 20% of night-time (9% mean error). While measurements during high-intensity exercise may be inaccurate [31,33], it has been argued that night-time data from commercial fitness trackers are sufficiently reliable to accurately estimate HRV at rest [31]. By leveraging the longitudinal and multidimensional nature of ReaLM methods, it would be possible to control for individual and device-specific differences as well as activity-levels and environmental exposures for short- or long-term HRV deviations.

Overall, these results suggest the need for thorough diagnostic analysis before applying ReaLM data to specific research questions. For example, exploring spatial habits [38] may require limiting data to a number of days of particularly high coverage for each participant within a given time frame, whereas exploring meaningful place experiences allows for much broader data criteria.

7.3. Future development

To achieve insight into the development of healthy and sustainable lifestyles, researchers must leverage longitudinal data that examines environmental exposures, health practices and sustainable habits at a systemic level. MyGävle innovates this process by creating an interactive platform for individuals to contribute their data while living in accordance with their daily routines. To be effective this methodology must understand people's long-term realities without requiring intrusive measures. We term this Real-life in the Long-term with a non-invasive Methodology (ReaLM). The first operationalization of ReaLM done here reveals its promising potential and certain limitations. Overall, our data analysis revealed that quality criteria need be considered within disciplinary boundaries and appropriate diagnosis tools, yet simultaneously opened up fresh frontiers for interdisciplinary and transdisciplinary Big Data science [41].

Self-reporting in daily life has the potential to transform public health, clinical psychology, and in general human-centred research [21]. Our study suggests that engaging participants' motivation over extended periods of time is a major challenge for these domains. To address this, ReaLM techniques must draw inspiration from traditions of social media. As social media increasingly captivates participants, ReaLM projects must look beyond volunteer motivation to drive adoption and ensure the continued use of these methods in the future. Non-intrusivity, multi-level engagement, user experience and performance should be major focus points when designing such tools [42]. To further motivate users, gaming elements can be used to provide them with meaningful feedback, gratification, and direct rewards in return for their contribution. By forging a tangible link between the data participants supplied and their contributions to research, we experienced an increased enthusiasm that sustained throughout our process. To elevate ReaLM methods, passive data collection should be further explored as well as case-specific sources of GPS signals. Through the implementation of automated consistency protocols and minimized active user participation, there is potential to enhance these techniques.

Investing in ReaLM research is key to investigate the systematic roles that social environments and infrastructure have on fostering healthy and sustainable standards of living. ReaLM's mission to simultaneously prioritize personal health and environmental sustainability is one that resonates with many people; so much so that they eagerly embrace the opportunity to participate in the project. Interpreting the effects of personal health and environmental sustainability at a local level requires joining forces between academics and decision-makers. ReaLM initiatives provide an instrumental platform for crafting mutually beneficial collaborations to achieve this. With a concerted effort in the academic field, Big Data can be generated to provide creative and impactful solutions for real-world issues facing modern societies. This could place academia on an equal footing with commercial interests as leaders of future innovation.

The potential applications of ReaLM can be escalated to a new level with the integration of blockchain technology. This revolutionary form of data registry would empower participants by granting them complete autonomy and control over personal, biometric or any other sensitive information they provide [43]. By making sure ethical regulations align with enhancing privacy laws, we can ensure that trust in this process remains intact. The use of smart contracts combined with blockchain technology has remarkable potential for democratic participation [44], as demonstrated in successful studies that mapped plastic pollution [45] and analysed lake water quality [46]. This pioneering arrangement of incentives empowers citizens to actively shape their environments, with automatic rewards for data and engagement in projects that enhance their own health or sustainability. By doing so, ReaLM has the potential to revolutionize transdisciplinary science by transcending the boundaries of a single research project and by feeding directly to the promotion of impactful solutions.

In the future, ReaLM technologies aimed at creating healthy and sustainable lifestyles could be powered by various non-invasive health monitoring devices. Furthering this endeavour are Blockchain methods which can form a foundation for a citizen science platform that permits citizens to actively contribute in research activities as well as contributing to local decision-making processes.

Author contribution statement

Matteo Giusti: Karl Samuelsson: Conceived and designed the experiments; Performed the experiments; Analysed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Data availability statement

The authors do not have permission to share data.

Acknowledgements

This work was supported by the Vinnova project GeoLife Region and the University of Gävle. We thank the whole team of the BIG project, the Urban Studio and Future Position X for the support and great work.

10. Appendices

A. Screenshots from MyGävle

Figure A.1. Series on screenshots of all tabs and characteristics of MyGävle. a) home page tab with PPGIS guiding question; b) notification tab with received messages; c) feedback tab with aggregated positive and negative experiences of all participants; d) seasonal leaderboard with number of reported experiences for gamification; e) Settings tab; f) page of the manual procedure to connect Garmin watch via Bluetooth for HRV.

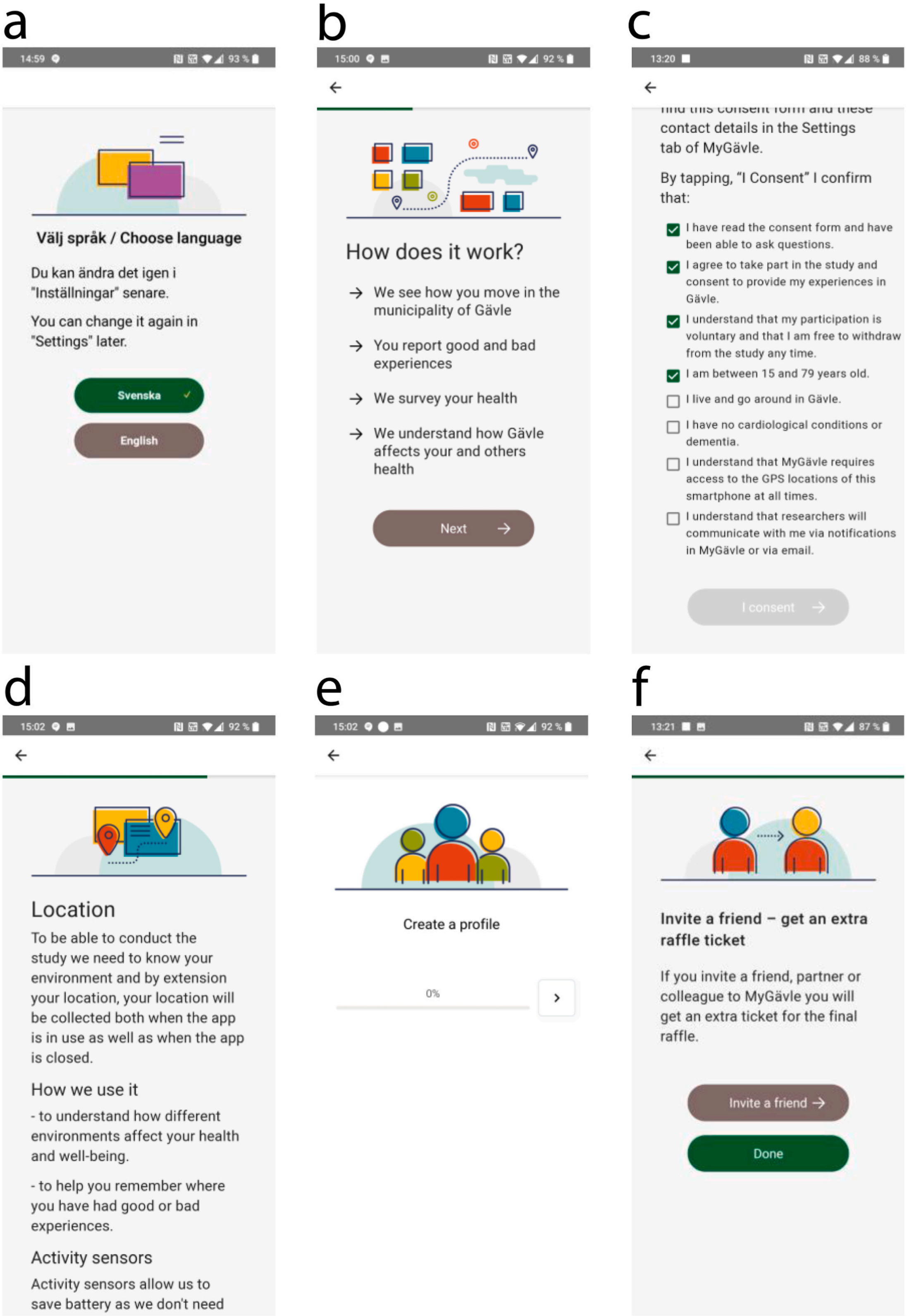

Figure A.2. Series on screenshots from the onboarding of a participant: a) the first screen allowing participants to choose between English and Swedish as the operating language of MyGävle, b) an introductory text to briefly clarify the purpose of MyGävle, c) the consent form with button and statements, d) the request to use the GPS sensor of the smartphone to access participants' locations, e) the beginning of the survey to collect demographic data, f) the screen that offered the possibility to participants to invite friends to the study.

B. Questionnaires

Table B.1
Profile questionnaire

Question	Answer/options	
Name	Text	
Personal number[1]	Twelve digits	
Email address	Text	
What is your employment status?	Employed	Student
	Self-employed	Retired
	Unemployed	Other
How much do you work or study from home?	0%	41–60%
	1–20%	61–80%
	21–40%	81–100%
What is your monthly income before taxes?[2]	Less than SEK 10 000	SEK 50 000 to 75 000
	SEK 10 000 to 20 000	SEK 75 000 to 100 000
	SEK 20 000 to 30 000	More than SEK 100 000
	SEK 30 000 to 50 000	
What level of education have you completed?	Elementary school	Master's degree or equivalent
	Secondary school	Doctoral degree
	Bachelor's degree or equivalent	
How long have you lived in your current neighbourhood?	Your whole life	2–5 years
	0–6 months	5–10 years
	7 months - 1 year	10–20 years
	1–2 years	More than 20 years but not your whole life
How long have you lived in Gävle?	Your whole life	2–5 years
	0–6 months	5–10 years
	7 months - 1 year	10–20 years
	1–2 years	More than 20 years but not your whole life
How did you get to know about MyGävle?	Social media	Organized event with researchers
	Articles on newspapers	Somebody invited me via MyGävle
	Physical posters at HiG	Somebody told me about it
	Physical posters in public spaces	Other

[1] Swedish personal numbers contain information about birthdate and gender
[2] SEK 10 000 was in autumn 2021 roughly equal to EUR 1000

Table B.2
Seasonal questionnaire.

Area	Statement
Human flourishing	Overall, I am satisfied with my life as a whole.
	I am usually happy.
	In general, I am physically healthy.
	In general, I am mentally healthy.
	Overall, the things I do in life give purpose to my life.
	I worry about meeting monthly living expenses.
Human-nature relationship	I am always comfortable being in natural outdoors, even in unpleasant weather.
	I am curious about how different plants, animals, and ecosystems look and work.
	In nature, I can find something to do everywhere and in every season.
	There are many natural places that I feel particularly attached to.
	I can tell if plants, animals, or ecosystems surrounding me are healthy or not.
	I have vivid memories in natural places or with plants and animals that shaped who I am.
	I know how to take care of plants, animals, and ecosystems around me.
	I am concerned, care profoundly about, and respect plants, animals, and ecosystems around me.
	I feel a deep love for plants, animals, and ecosystems around me.
Environmental social norms	Family, friends, or people around me actively protect nature.
	Family, friends, or people around me want me to actively protect nature.
	Family, friends, or people around me praise me for protecting nature.
Social capital	I have strong relationships with family and friends.

(*continued on next page*)

Table B.2 (*continued*)

Area	Statement
	I have strong relationships with people that are different from me in terms of religion, ethnicity, political opinions, education, or income levels.
	People in my everyday life could help me with money, babysitting, small house jobs, or in case of illness.
	I trust local politicians and governmental bodies.
	I trust my neighbors and the community where I live.

Table B.3
PPGIS questionnaire for subjective place experiences.

Request/question	Answer/options
1. Describe this experience in three words	Three text boxes
2. How natural is this place?	Not at all (1) to Very natural (7)
3. How did this level of nature contribute to your experience?	Very negatively (1) via Neither (4) to Very positively (7)
4. How many people were around you in this place?	None at all (1) to Very many (7)
5. How did this amount of people contribute to your experience?	Very negatively (1) via Neither (4) to Very positively (7)
6. How much built infrastructure is in this place (e.g. facilities, services, buildings)?	None at all (1) to Very much (7)
7. How did this level of nature contribute to your experience?	Very negatively (1) via Neither (4) to Very positively (7)
8. How did this experience affect your happiness or life satisfaction at the time?	Very negatively (1) via Neither (4) to Very positively (7)
9. How did this experience affect your sense of meaning or purpose in life at the time?	Very negatively (1) via Neither (4) to Very positively (7)
10. How did this experience affect your mental wellbeing at the time?	Very negatively (1) via Neither (4) to Very positively (7)
11. How did this experience affect your physical wellbeing at the time?	Very negatively (1) via Neither (4) to Very positively (7)
12. How did this experience affect your relationships with family and friends at the time?	Very negatively (1) via Neither (4) to Very positively (7)
13. How did this experience affect your relationships with the local community at the time?	Very negatively (1) via Neither (4) to Very positively (7)
14. How did this experience affect your relationship with nature at the time?	Very negatively (1) via Neither (4) to Very positively (7)
15. How did this experience affect your sense of safety at the time?	Very negatively (1) via Neither (4) to Very positively (7)

C. Participant recruitment.

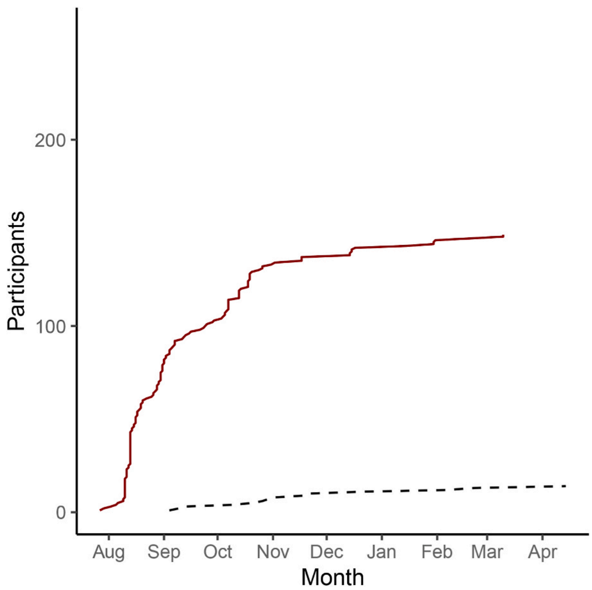

Figure C.1. Time series of participant recruitment. The grey line represents the public sample, the dark red line represents the health sample, and the dashed line represents exits from the study.

D. Temporal coverage estimation

We consider GPS measures with a margin of error (MOE) of 25 m or lower to be “high accuracy” and those with larger MOE “low accuracy”. Based on this division, we divide each participant’s day time (6:00–23:00) into four classes of temporal coverage: high-accuracy, low-accuracy, likely missing, and missing.

Because GPS points are recorded only in motion or when transitioning between in motion and being stationary, we use different procedures to estimate stationary and mobile coverage, respectively.

The most straight-forward case is coverage while in motion. We assign a GPS record with activity type walking, running, biking or in vehicle as covering the entire timespan to the next measurement if these are within 11 s of each other (adding a 10% margin of error to the 10 s interval used by the location tracking app). This coverage is set to either high-accuracy or low-accuracy depending on the

MOE of the GPS point.

In case of a GPS record with stationary or unknown activity type, we assign high-accuracy coverage during the entire timespan to the next measurement if the measured distance between them plus their combined MOE is less than 100 m. This limits the maximum actual distance between them to 100 m and the likely distance to considerably less than that in most cases. We assign low-accuracy coverage if the measured distance between the measurements is less than 100 m but the measured distance between them plus their combined MOE is greater than 100 m. This makes it likely but not certain that the actual distance between them is less than 100 m. We assign likely missing coverage if the measured distance between the measurements is more than 100 m but the measured distance between them minus their combined MOE is less than 100 m. This makes it likely but not certain that the actual distance between them is more than 100 m. Lastly, we assign missing coverage if the measured distance between the measurements minus their combined MOE is more than 100 m. This limits the minimum actual distance between them to at least 100 m. See Fig. D.4 for a schematic illustration of this classification.

In the case of missing coverage, we treat the first 10 s of the timespan until the next measurement as either high-accuracy or low-accuracy (depending on the MOE of the first GPS point) and the remaining time as missing.

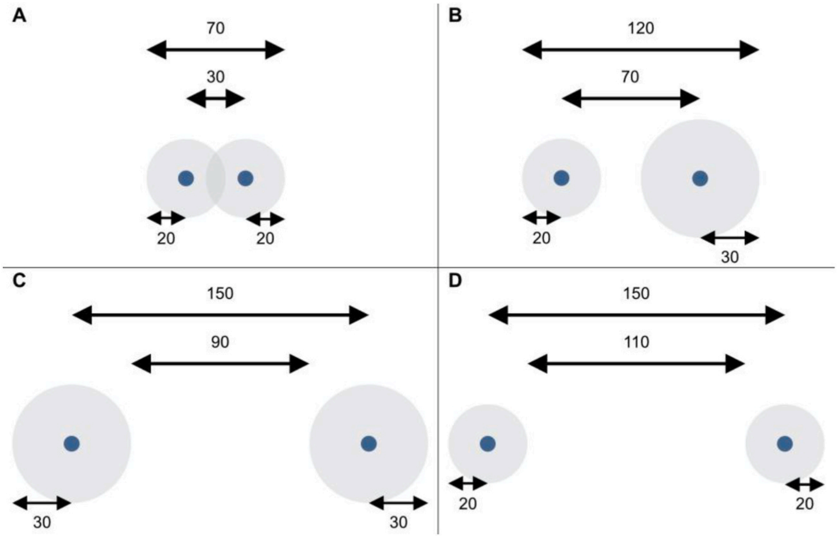

Figure D.1. Schematic illustration of the temporal coverage classification. The blue dots represent consecutive GPS points and the grey circles around them their accuracy margin of error (MOE). (A) High-accuracy coverage: the measured distance between them plus their combined MOE is less than 100 m. (B) Low-accuracy coverage: the measured distance between them is less than 100 m but the measured distance between them plus their combined MOE is greater than 100 m. (C) Likely missing coverage: the measured distance between them is more than 100 m but the measured distance between them minus their combined MOE is less than 100 m. (D) Missing coverage: the measured distance between them minus their combined MOE is more than 100 m.

E. Spatio-temporal tracking: supplementary analysis of completeness and consistency

Beta regression revealed that hours of daily GPS data coverage is associated with the number of daily GPS locations recorded ($z = 3.38$, $p < 0.001$), as well as with whether the participant belongs to the health sample ($z = 3.99$, $p < 0.001$) (Fig E1A). This latter difference might be related to the phones used, as participants in the health sample are more likely to use an Android phone (X2(1,257) = 10.2, $p = 0.001$) and hours of daily coverage is greater for phones with Android than iOS ($H(1) = 21.3$, $p < 0.001$; Fig. E.1B).

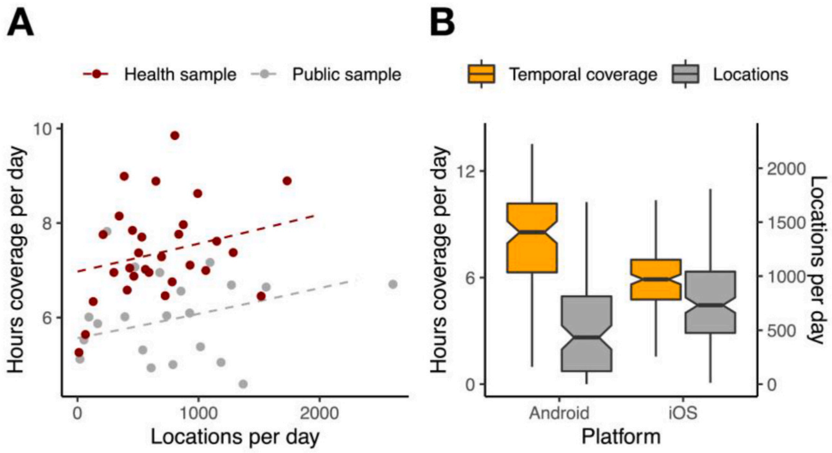

Figure E.1. (A) Hours of GPS record coverage per day predicted by locations per day and whether participants are in the health sample or not. Each dot represents five participants, ranked and binned by locations per day. Red dots represent participants in the health sample while grey dots represent those outside it. (B) Differences in hours of GPS coverage and number of locations per day between phones using Android and iOS. Android phones have greater temporal coverage despite not recording more locations.

Individual-level consistency between GPS point distances and fitness wristband step counts varies from excellent with coefficients around 0.9 for some participants to non-existent for others, and notably shows no association with daily step count (pseudo-$R^2<0.001$, $p = 0.763$; Fig. E.2A) or phone operating system (pseudo-$R^2 = 0.012$, $p = 0.272$), but a positive association with distance between daily GPS records (pseudo-$R^2 = 0.135$, $p < 0.001$; Fig. E.2B). This indicates that consistency between phone and fitness wristband data relates more to factors related the phone (such as how often the participant carries the phone with her or him) than the fitness wristband use or phone operating system.

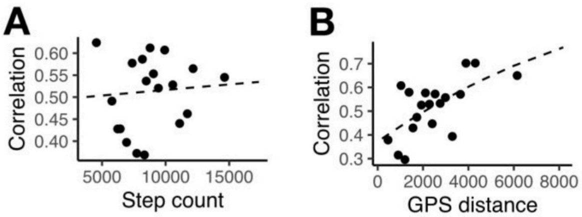

Figure E.2. (A) Individual-level Pearson's correlation coefficients between distances between GPS points and step counts (y-axis) shows no association with step count (x-axis). Dashed line shows beta regression fit, for participants that had at least 14 days of simultaneous step count and GPS records (n = 99). Each dot represents five participants, ranked and binned by step count (participants ranked 90–99 were binned into one group of 9). (B) Same as (F) but with correlation coefficients regressed on sum of distances between GPS records, and participants ranked and binned by sum of distance between GPS records. Correlation coefficient displays a strong association with distances between GPS records.

F. Estimating erroneous heart beats

As an accuracy analysis of the interbeat intervals obtained from the fitness wristbands, we select for each participant a random day with at least 12 h of fitness wristband data and a random night with at least 5 h data, and calculate the proportion of interbeat interval observations that are classified as errors for each following a procedure outlined by Lipponen and Tarvainen (2019). This procedure accounts for extra, missed, long, short, and ectopic beats by calculating the difference between successive interbeat intervals standardized over the quartile deviation of the 91 surrounding intervals, and the difference between an interbeat interval and the median of the 11 surrounding intervals standardized over the quartile deviation of the 91 surrounding intervals. A scaling factor is used for both standardization; following Lipponen and Tarvainen (2019) here we use 5.2. All observations that obtain values greater than one for either of the two measurements are considered errors.

References

[1] M. Kondo, J. Fluehr, T. McKeon, C. Branas, Urban green space and its impact on human health, Int. J. Environ. Res. Publ. Health 15 (2018) 445.
[2] M.J. Nieuwenhuijsen, Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities, Environ. Health 15 (2016) S38.
[3] N. Linder, M. Giusti, K. Samuelsson, S. Barthel, Pro-environmental habits: an underexplored research agenda in sustainability science, Ambio (2021), https://doi.org/10.1007/s13280-021-01619-6.
[4] L. Kong, Z. Liu, J. Wu, A systematic review of big data-based urban sustainability research: state-of-the-science and future directions, J. Clean. Prod. 273 (2020), 123142.
[5] V. Mayer-Schönberger, K. Cukier, Big Data: a Revolution that Will Transform How We Live, Work, and Think, Mariner Books, Houghton Mifflin Harcourt, 2014.
[6] L. Sinnenberg, et al., Twitter as a tool for health research: a systematic review, Am. J. Publ. Health 107 (2017) e1–e8.
[7] S. Aslam, Twitter by the Numbers: Stats, Demographics & Fun Facts. https://www.omnicoreagency.com/twitter-statistics/, 2022.
[8] I. Mansoor, Twitter Revenue and Usage Statistics. https://www.businessofapps.com/data/twitter-statistics/, 2022.
[9] B. Jacobs, Popma, Big Data and the need for privacy by design, Big Data Soc. 6 (2019), 2053951718824352.
[10] D. Lazer, R. Kennedy, G. King, A. Vespignani, The parable of Google flu: traps in big data analysis, Science 343 (2014) 1203–1205.
[11] S.J. Mooney, D.J. Westreich, A.M. El-Sayed, Epidemiology in the era of big data, Epidemiology 26 (2015) 390–394.
[12] X. Jin, Significance and challenges of big data research, Big Data Research 6 (2015).
[13] G. Boulton, The challenges of a big data earth, Big Earth Data 2 (2018) 1–7.
[14] A. Al-Badi, A. Tarhini, A.I. Khan, Exploring big data governance frameworks, Proc. Comput. Sci. 141 (2018) 271–277.
[15] L.P. de Vries, B.M.L. Baselmans, M. Bartels, Smartphone-based ecological momentary assessment of well-being: a systematic review and recommendations for future studies, J. Happiness Stud. 22 (2021) 2361–2408.
[16] T.R. Kirchner, S. Shiffman, Spatio-temporal determinants of mental health and well-being: advances in geographically-explicit ecological momentary assessment (GEMA), Soc. Psychiatr. Psychiatr. Epidemiol. 51 (2016) 1211–1223.
[17] D.M. Clark, J.D. Teasdale, Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences, J. Abnorm. Psychol. 91 (1982) 87–95.
[18] S. Shiffman, A.A. Stone, M.R. Hufford, Ecological momentary assessment, Annu. Rev. Clin. Psychol. 4 (2008) 1–32.
[19] A. Jones, et al., Compliance with ecological momentary assessment protocols in substance users: a meta-analysis, Addiction 114 (2019) 609–619.
[20] M. Boukhechba, et al., Contextual analysis to understand compliance with smartphone-based ecological momentary assessment, in: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare, ACM, 2018, pp. 232–238, https://doi.org/10.1145/3240925.3240967.
[21] K. Doherty, A. Balaskas, G. Doherty, The design of ecological momentary assessment technologies, Interact. Comput. 32 (2020) 257–278.
[22] D. Donaire-Gonzalez, et al., ExpoApp: an integrated system to assess multiple personal environmental exposures, Environ. Int. 126 (2019) 494–503.
[23] J. Sztajzel, Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system, Swiss Med. Wkly. (2004), https://doi.org/10.4414/smw.2004.10321.
[24] T.J. VanderWeele, On the promotion of human flourishing, Proc. Natl. Acad. Sci. U.S.A. 114 (2017) 8148–8156.
[25] M. Giusti, U. Svane, C.M. Raymond, T. Beery, A framework to assess where and how children connect to nature, Front. Psychol. 8 (2018).
[26] H.S. Park, S.W. Smith, Distinctiveness and influence of subjective norms, personal descriptive and injunctive norms, and societal descriptive and injunctive norms on behavioral intent: a case of two behaviors critical to organ donation, Hum. Commun. Res. 33 (2007) 194–218.
[27] T. Healy, The measurement of social capital at international level, Soc. Capital 24 (2002).
[28] E. Villalonga-Olives, The development of a bridging social capital questionnaire for use in population health research, Popul. Health 10 (2016).
[29] V. Jayawardene, S. Sadiq, M. Indulska, An Analysis of Data Quality Dimensions, vol. 40, 2013.
[30] N. Laranjeiro, S.N. Soydemir, J. Bernardino, A survey on data quality: classifying poor data, in: 2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC), 2015, pp. 179–188, https://doi.org/10.1109/PRDC.2015.41.
[31] K. Georgiou, et al., Can wearable devices accurately measure heart rate variability? A systematic review, Folia Med. 60 (2018).
[32] S. Gradl, M. Wirth, R. Richer, N. Rohleder, B.M. Eskofier, An overview of the feasibility of permanent, real-time, unobtrusive stress measurement with current wearables, in: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, ACM, 2019, pp. 360–365, https://doi.org/10.1145/3329189.3329233.
[33] J.A. Bunn, J.W. Navalta, C.J. Fountaine, Current State of Commercial Wearable Technology in Physical Activity Monitoring 2015-2017, 2018, p. 13.
[34] J.A. Lipponen, M.P. Tarvainen, A robust algorithm for heart rate variability time series artefact correction using novel beat classification, J. Med. Eng. Technol. 43 (2019) 173–181.
[35] K.M. Brown, The haptic pleasures of ground-feel: the role of textured terrain in motivating regular exercise, Health Place (2016), https://doi.org/10.1016/j.healthplace.2016.08.012.
[36] T. Strain, K. Wijndaele, S. Brage, Physical activity surveillance through smartphone apps and wearable trackers: examining the UK potential for nationally representative sampling, JMIR Mhealth Uhealth 7 (2019), e11898.
[37] S.N. Zenk, S.A. Matthews, A.N. Kraft, K.K. Jones, How many days of global positioning system (GPS) monitoring do you need to measure activity space environments in health research? Health Place 51 (2018) 52–60.
[38] J.R. Meliker, C.D. Sloan, Spatio-temporal epidemiology: principles and opportunities, Spat. Spatio-temporal Epidemiol. 2 (2011) 1–9.
[39] P. Juarez, et al., The public health exposome: a population-based, exposure science approach to health disparities research, IJERPH 11 (2014) 12866–12895.
[40] J.M. Rippe, Lifestyle medicine: the health promoting power of daily habits and practices, Am. J. Lifestyle Med. 12 (2018) 499–512.
[41] B. Bálint, P. Mooney, E. Nováková, L. Bastin, J. Jokar Arsanjani, Data quality in citizen science, in: K. Vohland, et al. (Eds.), The Science of Citizen Science, Springer International Publishing, 2021, pp. 139–157, https://doi.org/10.1007/978-3-030-58278-4_8.
[42] R. Lemmens, V. Antoniou, P. Hummer, C. Potsiou, Citizen science in the digital world of apps, in: K. Vohland, et al. (Eds.), The Science of Citizen Science, Springer International Publishing, 2021, pp. 461–474, https://doi.org/10.1007/978-3-030-58278-4_23.
[43] S. Leible, S. Schlager, M. Schubotz, B. Gipp, A review on blockchain technology and blockchain projects fostering open science, Front. Blockchain 2 (2019).
[44] F.L. Benítez-Martínez, M.V. Hurtado-Torres, E. Romero-Frías, A neural blockchain for a tokenizable e-Participation model, Neurocomputing 423 (2021) 703–712.
[45] S. Lynch, OpenLitterMap.com – open data on plastic pollution with blockchain rewards (littercoin), Open Geospatial Data, Soft. Stand. 3 (2018) 6.
[46] J.R. Cedeno Jimenez, A. Folini, M.A. Brovelli, Collaborative validation of user-contributed data using a geospatial blockchain approach: the simile case study, Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. (2022) 89–95. XLVIII-4/W1-2022.