



OPEN

DATA DESCRIPTOR

Explorative pedestrian mobility geolocated data from a citizen science experiment in a neighbourhood

Ferran Larroya^{1,2}✉, Roger Paez³, Manuela Valtchanova³ & Josep Perelló^{1,2}✉

Pedestrian geolocated data are key to a better understanding of micro-mobility within a neighbourhood. These data can bring new insights into walkability and livability in the context of urban sustainability. However, pedestrian open data are scarce and often lack a context for their transformation into actionable knowledge in a neighbourhood. Citizen science and public involvement practices are powerful instruments for obtaining these data and take a community-centred placemaking approach. The study shares some 3 000 geolocated records corresponding to 19 unique trajectories made and recorded by groups of participants from three distinct communities (72 participants and 19 groups) in a relatively small neighbourhood. The groups explored the neighbourhood through a number of actions and chose different places to stop and perform various social and festive activities. The study shares not only raw data but also processed records with specific filtering and processing to facilitate and accelerate data usage. Citizen science practices and the data-collection protocols involved are reported in order to offer a complete perspective of the research undertaken jointly with an assessment of how community-centred placemaking and operative mapping are incorporated into local urban transformation actions.

Background & Summary

Mobility and public spaces are both identified as key issues in the attempts to achieve more sustainable, walkable and livable cities¹. Understanding how cities affect the behaviour of their inhabitants and how urban public spaces can be reshaped by this behaviour are important challenges to be considered in the development and regeneration of public spaces². The analysis of available digital traces³ can contribute to this effort as part of the interdisciplinary approaches currently applied within the field of urban science^{4–9}.

A growing body of research is adopting a complex systems science approach to deepen the understanding of social behavior and inform urban planning strategies^{4–8}. This framework underscores the importance of accounting for urban inhomogeneities—such as unequal access to resources, variations in population density, diverse configurations of public space, and the heterogeneous social and cultural backgrounds of residents, taking a data-driven perspective⁹. Within this context, human mobility data has become a valuable resource, offering fine-grained insights into how people navigate cities, engage with public infrastructure, and form social connections across neighborhoods. Integrating such data into computational models enables the identification of movement patterns¹⁰ and the detection of spatial inequalities¹¹ that might otherwise remain hidden in aggregated datasets. By adopting this systems-based perspective, urban planners and policymakers are better equipped to design cities that are not only more resilient and adaptable, but also more inclusive and reflective of the diverse lived experiences within urban populations^{1,2}.

When studying urban mobility, digital traces have shown their potential for investigating long distance movements (e.g., in public transportation¹²) but there is still a lack of high resolution micro-mobility data¹³ to

¹OpenSystems, Departament de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franqués, 1, 08028 Barcelona, Catalonia, Spain. ²Universitat de Barcelona Institute of Complex Systems, Universitat de Barcelona, Barcelona, Catalonia, Spain. ³ELISAVA Barcelona School of Design and Engineering, Universitat de Vic - Universitat Central de Catalunya, La Rambla, 30-32, 08002, Barcelona, Catalonia, Spain. ✉e-mail: ferran.larroya@ub.edu; josep.perello@ub.edu

study and characterize the local micro-behaviours in people's neighbourhoods^{14,15}. Many research questions may be related to micro-mobility. Micro-mobility patterns of pedestrians and cyclists can be effectively modeled using gravity models¹⁶ or random walk approaches that incorporate attractive hotspots^{17,18}. These models are useful for identifying regularities in movement behavior, such as the distances people typically travel¹⁶ or the duration of their stops at specific locations¹⁸. Such analyses support a range of urban planning and public health applications. For instance, they can be used to evaluate geographic disparities in access to parks¹⁹, assess pedestrian exposure to air pollution at the city scale²⁰, or examine behavioral changes in response to COVID-19 policies^{14,21}. Focusing more specifically on the data described in this manuscript, these methods also bring insights into how people use public spaces for walking and which locations they choose for specific types of activities. The inferred walkability taking GPS (Global Positioning System) geolocated data in its finest resolution overlaps strongly with the analysis of livability in general, and with lively and sociable urban environments in particular²². In this context, the street-space approaches of urban placemaking²³ and operative mapping²⁴ can promote and encourage a more open and collaborative use of high-resolution GPS data, create reflection and discussion among citizens and city makers, and motivate public space transformations with new evidence^{25,26}.

Participatory research approaches such as citizen science²⁷ represent a valid option for providing high resolution datasets in small-scale urban contexts^{28,29} and contribute to a community-centred placemaking in more livable cities³⁰. Citizen science practices involve the active participation of the general public in scientific research tasks, which may include the formulation of scientific questions, the elaboration of the experimental protocol, and the collection and subsequent interpretation of data²⁷. Citizen science data are already contributing to the monitoring of sustainable goals in a bottom-up approach³¹ but social injustices emanating from the data processes in citizen science need to be considered in order to avoid exploitative and exclusionary implications³². To effectively contribute to community urban planning³³, citizen science thus should incorporate diversity of participants, consider social and environmental justice, and co-produce new knowledge responding to people's concerns in an actionable manner^{34,35}.

The present study focuses on the community action to explore temporary uses of public spaces in the “Primer de Maig” neighbourhood of the city of Granollers (Catalonia, Spain). The city belongs to the Barcelona metropolitan region and has 61 983 inhabitants, covers 1 487 ha and 4 160 inhabitants per km². The “Primer de Maig” neighbourhood is one of the 35 censal districts of the city, with a surface area of about 3 ha and 45 200 inhabitants per km² (1 356 residents). Figure 1 shows an aerial map of the neighbourhood. It is the district with the lowest net annual income per person (9 329 EUR) and per dwelling (25 026 EUR) (<https://www.ine.es/jaxiT3/Tabla.htm?t=30896>, in Spanish). This small neighbourhood was one of the first areas of peripheral homes in the city of Granollers. The buildings provided a rapid solution to the housing problem that arose in the middle of the 20th century due to the large-scale migration waves (<https://www.granollers.cat/can-jonch/grup-primer-maig>, in Catalan). Until that time, this area was mostly devoted to industry and agriculture. The neighbourhood is made up of 43 buildings, with a total of more than 550 homes. Its urban nature is based on the construction of typologies of repetitive housing blocks with a strong presence of blind party walls and surrounded by medium-small empty spaces. Public space is generous in terms of quantity but poor in terms of quality. These characteristics mark the idiosyncrasy of the neighbourhood and its possibilities of transformation.

Design for City Making Research Lab (Elisava Research, UVic-UCC) invited OpenSystems group (Complexity Lab Barcelona, UBICS) to join the experiment in Granollers (see Fig. 2 for details). The broad aim was to enrich and widen urban planners' visions by providing concrete evidence through GPS data to support the redesign of public space and the regeneration of “Primer de Maig” neighbourhood³⁶. The collection of GPS explorative trajectories was based on citizen science practices and on previous experiences in running public and participatory experiments^{37–39}. In 2012, a hundred citizens were engaged in an experiment in a public park (Parc de la Ciutadella, 4 ha, Barcelona) during a Science Festival¹⁸. Another experiment (2017) involved 10 public schools from the Barcelona metropolitan area and more than 400 participants, including students (3rd and 4th grade of high school) and teachers⁴⁰. In both cases, we collected data on purpose-based pedestrian mobility and built our own mobile App to obtain the GPS data from participants' mobile phones. However, periodic updates and maintenance of the App represented important logistic challenges.

For the experiment in the “Primer de Maig” neighbourhood, we used apps already existing on the market to track routes in different outdoor sports activities (Wikiloc App, <https://www.wikiloc.com>). The App being used was tested to confirm that it recorded high quality GPS data at a high resolution of about 4–5 seconds maximum in movement (when the App does not detect any movement, the data collection is paused). The Wikiloc App allows the registration of users with a Google account, and then gives access to the recorded trajectories and downloads the data. As is known, sharing GPS data entails some additional privacy risks and efforts must be made and to avoid possible identification of participants based on the possibility of locating their homes. In these cases a geo-masking technique (*k*-anonymity) has to be applied^{40,41}. The experiment in the “Primer de Maig” neighbourhood avoided this risk with a privacy-by-design approach as the participants explored the neighbourhood in groups of at least three, and started and finished the experiment at pre-established public locations that served as meeting points. By using rented devices with the research staff's Google accounts activated, we also prevented participants from providing any private personal information. The protocol guaranteed the total anonymity of the participants, who also read and signed an informed consent. For all these reasons, we are able to share both raw and processed GPS trajectories for all trajectories.

The participants represented different socio-demographic profiles which were all considered strategic agents in the neighbourhood. The groups were invited to explore the neighbourhood and find places for different activities in the form of actions. As part of the “Fem la Nostra Festa” (*Let's Celebrate*, possible translation from Catalan) event, the actions were planned to obtain socio-spatial activation of unconsolidated public spaces. They aimed to connect spatiality (e.g., physical, material and infrastructural conditions of the neighbourhood) and sociality (e.g., relational, cultural and performative components)⁴². The festive format of the activity aimed

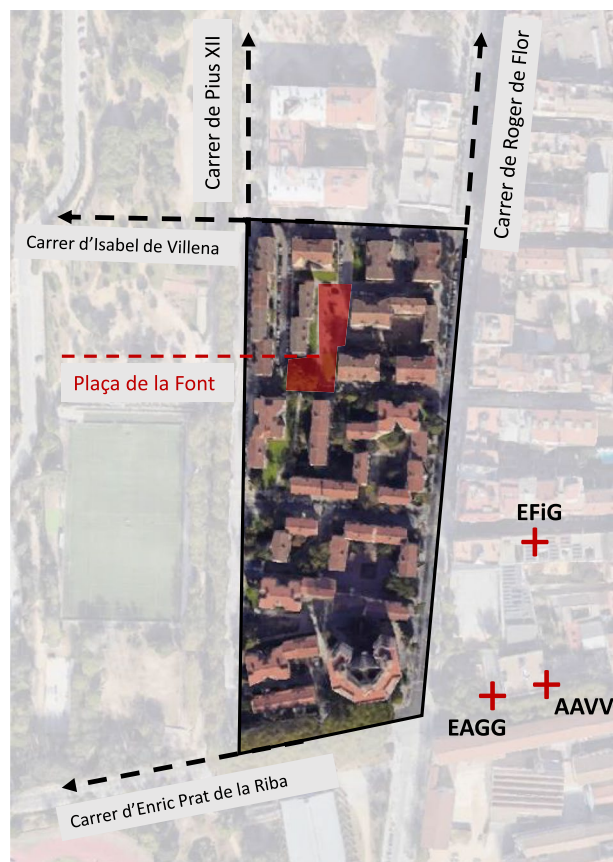


Fig. 1 Map of “Barri Primer de Maig”. The neighbourhood “Primer de Maig” in the city of Granollers (Catalonia, Spain) is limited by the streets: “Carrer d’Enric Prat de la Ribera”, “Carrer de Roger de Flor”, “Carrer de Pius XII” and “Carrer d’Isabel de Villena”. The red crosses represent the location of the three sites related to the communities involved: the school “Escola Ferrer i Guàrdia (EFiG)”, the neighbourhood association site “Associació de Veïns Sota el Camí Ral (AAVV)”, and the space for older people “Espai Actiu de la Gent Gran (EAGG)”. The meeting point was located at the “Plaça de la Font” and this square is also displayed in red. The neighbourhood has an area of approximately 3 ha (300 × 100 metres) and the distance between the three communities’ locations and the neighbourhood is less than 70 metres. The satellite image was retrieved from the Spanish National Geographic Institute (<https://www.ign.es/web/ign/portal>) with data from the European Copernicus satellites.

to reinforce an active attitude to connecting urban space and participants’ desires. For this reason, the citizen science experiment was specifically designed to capture these aspects through the predefined selection of six festive actions⁴³ (see Table 1), which we determined to be sufficient given the time and space constraints. Indeed, incorporating playful and festive actions can inspire communities to collectively reimagine and reinvent their public spaces within what is often referred as community placemaking^{33,44–46}. The GPS data collection was complemented with additional activities to share and discuss the aggregate GPS data from all groups. “Fem la Nostra Festa” triggered proactive and creative urban attitudes to discovering opportunities for socialization, to identifying the coexistence between different appropriations of public space, and to promoting sustained social development based on complexity and diversity.

Figure 2 shows the schematic description of the process. Here we share and report unique micro-mobility data from 72 participants (19 groups) with different socio-demographic traits. Participants acted as explorers of the neighbourhood, carrying around an exploration kit with different objects to perform specific festive actions in places of their choice (see Table 1). We share the original datasets and processed records with specific filtering and processing procedures for scientific research in order to facilitate data re-usability⁴⁷.

One of the main challenges in citizen science initiatives is participant engagement, which can limit the usefulness of the data due to size constraints³⁹. In this experiment, we collected approximately 3 000 GPS locations. However, a citizen science experiment can also address in a unique manner very specific aspects which cannot be studied by other means³⁹. Pedestrian mobility data in a very high resolution level (few meters) captured through a citizen science experiment can enable a cartographic analysis on walkability in a neighborhood^{29,48} or to very concrete purpose-based mobility patterns⁴⁰. These datasets can be obtained only with the conscious and active involvement of participants in a citizen science project. This approach can also help covering other active mobility types such as cycling^{49,50}.

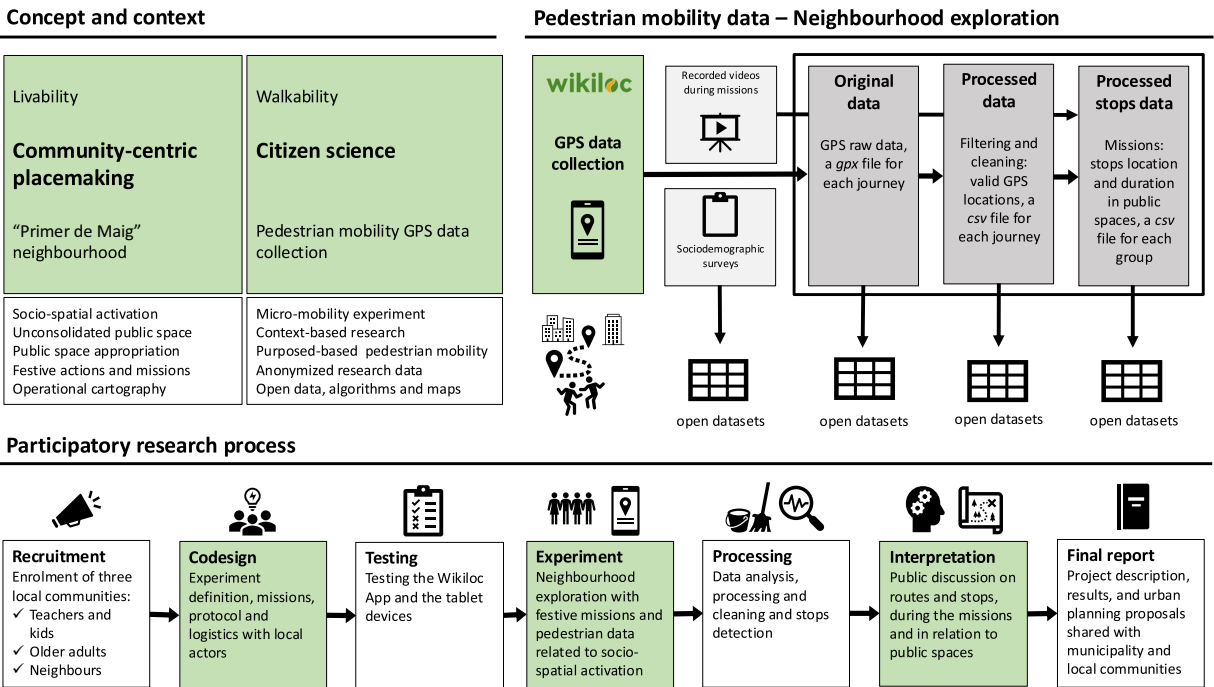


Fig. 2 Schematic description of the citizen science experiment. The data files available are reported jointly with the participatory process underpinning the whole research. The direct participation of citizens is highlighted in green and the four open datasets being created in grey.

Task	Code	Instructions
Dance	B	Explore and choose a place to dance using the speaker via Bluetooth
Play	J	Explore and choose a place to play with the ball
Chat	X	Explore and choose a place to sit and share personal neighbourhood stories
Lunch	M	Explore and choose a place to lay out the carpet and sit down to eat/snack
Celebration (toast)	C	Explore and choose a place to celebrate a special occasion with champagne/soft drink
Celebration (confetti)	CC	Explore and choose a place to celebrate a special occasion shooting the confetti gun

Table 1. List of festive actions to complete during the experiment. The groups had to choose a specific location in the neighbourhood to perform each of the six festive actions from the list. A kit comprising items for the festivities was provided. The tablets that each groups was carrying recorded GPS data, and filmed a video of the groups completing each action. This latter activity was a part of the data validation process.

Among the 19 groups involved, there were about 80 stops where well-identified actions were performed. This dataset is sufficiently large for the purpose of the experiment, that was to complement the cartographic studies of urban experts and to help identify, through the participants’ experiences, the most suitable places for redesigning the public space of the neighbourhood²⁶. Since we share the complete trajectories of the 19 participating groups and complement them with the sociodemographic information of the participants and with the specific type of actions, we believe that this dataset also represents a powerful resource in the field of exploratory mobility⁵¹. Expanding the dataset with urban-related characteristics (e.g. width of the streets, amount of green) or eventually climatic aspects⁵² could further enhance its applicability in urban science studies.

From an academic scientific perspective, data can indeed enrich and contribute to discussions about the most suitable models for exploratory pedestrian mobility¹⁰ (e.g., through stochastic or random walk models). Mobility can also be characterized by velocities (see Methods) or other statistical variables like turning angle or tortuosity¹⁰. Finally, from the perspective of data science, clustering techniques can be applied, for example to find high stop density regions, or to study the direction of movement flows in the busiest areas. This would help architects and urban planners to plan actions and interventions in public space in order to enhance livability.

Methods

The Universitat de Barcelona Ethics Committee (IRB00003099) approved this mobility experiment. All participants read and signed the informed consent form and parental/legal guardian consent if appropriate. No privacy issues were observed to be in conflict with the public release of the underlying processed data.

Co-designing the experiment with local communities. Between October 2021 and January 2022, the Design for City Making Research Lab and the OpenSystems group joined forces to design an urban mobility experiment in which the analysis of the data obtained could help to understand how the public space of the neighbourhood “Primer de Maig” in Granollers could be redesigned. The researchers decided to develop a set of public experiments^{37–39} in groups of people from local communities, where the participants had to explore the public space while choosing specific sites to perform actions of a festive nature, and carrying a festive exploration kit with various objects to perform the actions, see Table 1). The trajectories (GPS records) followed by the participants were then registered through the Wikiloc App using tablets. The experiments were planned as part of the neighbourhood festivities (April 29 and 30 and May 1, 2022).

An open call for participation was issued to the different associations or public centres in the neighbourhood. Three communities agreed to join the project: the school “Escola Ferrer i Guàrdia (EFiG)”, the neighbourhood association “Associació de Veïns Sota el Camí Ral (AAVV)”, and the older adults group “Espai Actiu de la Gent Gran (EAGG)”. Three 2-hour presentation sessions were organized with the three participating communities, hosted by the researchers. The aim of these sessions was to present the project to the interested members of each community, along with all the details corresponding to the objectives and the collection of data on pedestrian mobility. All doubts and questions were discussed. The meetings represented a chance to reflect on the urban structure of the neighbourhood. The attendees discussed the form of the pedestrian mobility experiments, and were actively engaged in the creation of the experimental protocol for the experiments.

The first meeting was held at the “Associació de Veïns Sota el Camí Ral (AAVV)” on February 16. The second meeting took place at the “Espai Actiu de la Gent Gran (EAGG)” on March 2. The last meeting was held on March 16 at the “Escola Ferrer i Guàrdia (EFiG)”. People interested in doing the experiment had about one month and a half to form the groups and confirm their participation in one of the three local communities involved. In this way, we were able to know the number of participating groups in advance and thus had enough time to prepare the material for all the experiments (e.g., the tablet devices and the festive exploration kits). The research team did not collect any personal data and only knew the number of groups and participants. The call for participation resulted in 19 groups of 3–5 people and a total number of 72 participants. From the school (EFiG), there were 15 teachers who formed four different groups and a class group (6th grade of primary school, 11–12 years old) of 26 students, who formed seven groups. From the two associations (AAVV and EAGG), eight groups were formed with a total of 31 participants.

Data acquisition with the tablet device App. A set of actions were proposed to the participants, which they had to complete in about 1h 30m. The actions consisted in performing specific activities of a festive nature, and the groups participating in the experiment had to find their preferred place inside the neighbourhood. The exploration kit included the objects to perform the set of actions and the related instructions which read: “Find a place to dance with your group. Connect the loudspeaker from the kit via Bluetooth and play a song with the iPad” or “Choose a place to make a toast: open the drink and make a toast to a special occasion for you”. The list of the six different festive actions is provided in Table 1. The aim of the kit was to create an enjoyable recreational experience and to encourage the participants to adopt an exploratory attitude towards the neighbourhood.

The experiments with the participating groups from the school (EFiG) took place on Friday April 29, 2022, during school hours. The groups started and finished the data collection in the vicinity of the school. The four groups of teachers participated in the experiment during the lunch break (between 1:30 and 3:00 pm). The seven groups of students from the same class participated as part of their timetable (3:30 – 5:00 pm). The following day, Saturday April 30 (11:30 am – 1:00 pm), the eight remaining groups (AAVV and EAGG) participated. These groups started gathering data at the EAGG and finished at “Plaça de la Font”, a square in the middle of the neighbourhood (see Fig. 1). The weather conditions were favourable on both days: there were no exceptional climate events (e.g., rain or very low/high temperatures).

Each participating group received a tablet device (iPad, with the iOS operating system and 1GB of high-speed 4G connectivity). On each of the devices, the Wikiloc App had been previously installed and a user id (Google account) created by the research staff was already logged in. Thus, the participants did not have to provide any personal information to use the App and the researchers were able to download the GPS data collected in `gpx` format once the experiment ended. A single account was assigned to each tablet device. Since there were eight tablets, only eight groups could perform the experiment simultaneously. Few groups used the same tablet device (at different times) and hence the same account. For this reason, a code name was assigned to each participating group. The group used the code name to save the recorded trajectory and the file generated `gpx` already contained the group code in the file name. The protocol thus avoided the possibility that participants might introduce personal data (e.g., their own name) in the file name. Finally, every tablet device was tested days prior to the experiment by the researchers, recording a trajectory with the Wikiloc App and exporting the GPS data.

Scientific protocol for the experiment. At the meeting point (EFiG for teachers and students on April 29 and EAGG on April 30 for the rest of the groups), the tablets were distributed to the different groups. The participants received the following instructions accompanied by some screenshots (see Fig. 3):

Before the experiment. (1) Each group member should fill out the anonymous socio-demographic data form (see Table 2) and sign an informed consent document. (2) The group must choose one member to be in charge of starting the App at the beginning of the experiment and stopping it at the end, and another member responsible for carrying the tablet with him/her throughout the journey (preferably someone already familiar with the technology). (3) The research team provides a tablet to the group in which the Wikiloc App is open and ready to start recording the GPS locations (see Fig 3a) jointly with a list of actions to complete along with the exploration kit (see Table 1).

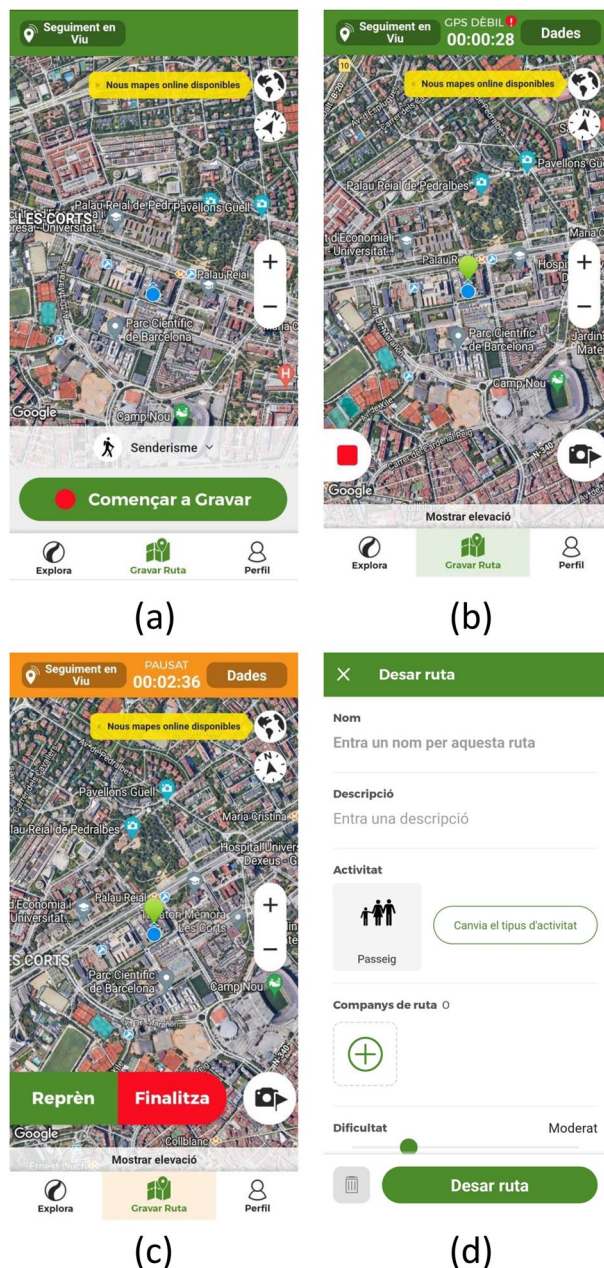


Fig. 3 Wikiloc App screenshots (in Catalan) during a test carried out in the University facilities (Faculty of Physics, Universitat de Barcelona). **(a)** Initial screen when opening the Wikiloc App. Participants had to click on “Començar a Gravar” (Start recording, in Catalan) to start recording the trajectory. **(b)** A real time recording of the journey; the recording time is displayed at the top of the screen, and the recording can be paused by pushing the red lower button. **(c)** The screen after pausing the recording of the journey; the participants had to click on “Finalitza” (End, in Catalan) to end the experiment and save the GPS data. **(d)** The screen after finalizing the recording. In this step, the participants returned the tablet to one of the researchers, who in turn saved the GPS data by clicking on “Desar ruta” (Save route, in Catalan), after entering the trajectory’s name “Nom” (Name, in Catalan) which was the code of the respective group.

During the experiment. (1) Click on “Començar a gravar” (Start recording, in Catalan, see Fig. 3a) when you start the journey. Avoid starting the experiment in indoor spaces as the GPS will lose precision. Walk between 50–100 meters before clicking on the App icon to start the experiment. This ensures a good connection to the GPS satellites. On the screen, you should see the recording time at the top and a red icon to pause the recording (see Fig. 3b). The App is collecting the GPS data from your journey. You do not need to have the Wikiloc visible on your screen; you can use other Apps during your journey. (2) During the experiment, the group must walk together and explore the neighbourhood while completing the list of festive actions with the exploration kit. The neighbourhood is limited by the streets: “Carrer d’Enric Prat de la Riba”, “Carrer de Roger de Flor”, “Carrer de Pius XII” and “Carrer d’Isabel de Villena” (see Fig. 1). The groups have 1 hour and 30 minutes to complete the

Grup:								
Nombre de participants en el grup:								
PARTICIPANTS	P1	P2	P3	P4	P5	P6	P7	P8
Rang d'edat								
<17								
18–35								
36–50								
51–65								
>65								
Identitat de gènere								
Dona								
Home								
No binari								
No vull respondre								
Coneixes el barri?								
Molt								
Poc								
Gens								
Freqüentes el barri?								
Cada dia								
Setmanalment								
Ocasionalment								
Gens								

Table 2. Socio-demographic form in Catalan. Each group had to fill out the socio-demographic form, firstly at the top adding the code of the group to which they belong ("Grup") and the number of participants in the group ("Nombre de participants en el grup"). Each participant (each column) in the group recorded their age range ("Rang d'edat"), the gender identity ("Identitat de gènere") male, female, non-binary or I don't want to answer. They also had to answer the question *Do you know the neighbourhood?* ("Coneixes el barri?") with a three-level answer: well, a little, or not at all. Finally, participants had to answer the question *Do you frequently go to the neighbourhood?* ("Freqüentes el barri?") with answers every day, weekly, occasionally or almost never.

list of festive actions in the list. (3) Use the tablet device to record a video of at least 30 seconds while performing the actions that comprise each action. This will allow the researchers to make a more accurate analysis of the locations chosen to perform the different actions.

After the experiment. (1) Upon completion of the actions, go to the meeting point - "Plaça de la Font" for the groups belonging to the associations AAVV and/or EAGG (April 30), and EFiG for the participants from the school (April 29), see Fig. 1. (2) Press the red icon to pause the recording (see Fig. 3b) and then "Finalitza" (End, in Catalan) to end the GPS recording (see Fig. 3c). (3) The App will then direct you to the save screen (see Fig. 3d). (4) Return the tablet to the researchers who are at the meeting point. They will check if you have correctly named and saved the recorded trajectory (using the group code name).

Group code names. As mentioned in the protocol, each group received a tablet device with the Wikiloc App with a session opened and ready to start recording the trajectory. A code was assigned to each group, which was also used to fill out the socio-demographic form and to write the session name with the GPS data collected in a single trajectory (and therefore the `gpx` also had the group code name as a file name).

The code names assigned to the four groups of teachers from the school (EFiG, Friday April 29) were: P1, P2, P3, and P4. The code names assigned to the students who participated in the experiment in the afternoon (EFiG, Friday April 29) were: E1, E2, E3, E4, E5, E6, and E7. On the following day, April 30, local participants from the residents' associations and older people (AAVV and EAGG) had the eight following codes: G1, G2, G3, G4, G5, G6, G7, and G8. For each group (code name), Table 3 shows the number of participants, the number of GPS data records and the number of stops, before and after data processing and filtering. Tables 4, 5 and 6 show the results of the socio-demographic form filled in by each participating group. The profile of the 72 participants was quite heterogeneous in terms of age due to the different profile of each community involved. There was a higher participation of women (44) than men (28) and in most cases the participants reported frequenting the neighbourhood and knowing it well.

Supervision, monitoring and support. Before the experiments, scientists from the OpenSystems group tested each tablet, installing the Wikiloc App and recording several trajectories. The `gpx` files with the GPS data were exported and analyzed to ensure that the mobility data were collected accurately and were of good quality. During the experiments, all the research staff also walked through the neighbourhood jointly with the groups to supervise and monitor the progress, and to ensure that the experiment adhered to the established protocol defined above, preventing, for example, participants from entering buildings and losing GPS connection.

	Code name	Participants	Original Data		Processed Data	
			GPS records	Stops	GPS records	Stops
School teachers	P1	3	306	16	291	16
	P2	3	251	19	233	17
	P3	5	—	—	—	—
	P4	4	175	21	150	17
School students	E1	3	219	25	161	18
	E2	3	159	21	133	17
	E3	4	275	21	206	16
	E4	4	170	27	123	20
	E5	4	210	32	182	29
	E6	4	300	27	257	21
	E7	4	321	45	228	32
Adults, families and older adults	G1	5	241	40	227	36
	G2	3	72	7	52	4
	G3	3	163	25	150	21
	G4	3	90	10	64	5
	G5	3	168	26	143	20
	G6	3	107	26	87	13
	G7	6	162	25	156	22
	G8	5	143	17	138	15
Total		72	3532	420	2981	339

Table 3. Number of participants, GPS data records and stops per each group. The columns represent the profile of the participants, the code names for each group, the number of participants and the number of GPS records and stops. We add the original data being collected and the data after filtering and cleaning processes. The last row includes the total number of participants, the GPS records and the number of stops identified. The participating groups are split in three main blocks corresponding to the three experimental sessions held: teachers (EFiG), students (EFiG), and residents jointly with the older adults (AAVV and EAGG). P3 failed in the data recording and saving process.

They also solved any doubts regarding the actions to be completed or any technical problems with the devices. After the experiment (on the second day, April 30), OpenSystems researchers exported all the `gpx` files with the recorded journeys and created some visualizations of the trajectories on maps to obtain a first impression of the sites the participants chose to perform the tasks and which streets were the most frequented (see Fig. 4). These preliminary results were shown and discussed with the participating groups during a public lunch that was held at the end of the experiments on the occasion of the neighbourhood festivities.

Actionable data to deliver policy recommendations. One of the main objectives was to work in the neighbourhood as a network. Due to its low cost, its great impact, and the fact that it encourages the residents' active participation, the Design for City Making Research Lab proposed to work on an intensive urbanization strategy (as opposed to the extensive strategies conventionally used in public space urbanization) based on the GPS data collected. The data were carefully analyzed and were combined with a public space analysis that identified the spaces of opportunity for higher impact. Within those spaces, several interventions were proposed in specific locations (see Filtering and processing GPS data Section). These interventions were conceived as small constructions including utilities and actionable elements aimed to rehabilitate public spaces. These support structures were meant to socially activate their immediate environment by prompting temporary uses of public space initiated autonomously by citizens. Through this format, the aim was to achieve a large social impact with a cost 10 times lower than that of conventional public space urbanization, while working in conjunction with the municipal departments of urban planning and community action. A report was delivered to the municipality and to the local communities during meetings³⁶.

Filtering and processing GPS data. Due to the urban context of the experiment, we were interested in keeping the trajectories of the participating groups within the boundaries of the neighbourhood (Fig. 1). For this reason, we performed a simple filtering process in which we removed the GPS records at the start and/or end of the trajectory that were outside the neighbourhood. In addition, we performed a more advanced processing in which we labeled each GPS record as “stop” or “moving”. We also computed the time difference between consecutive timestamps, the distance, and the instantaneous velocity between consecutive GPS records. We added this extra information in the processed datasets and saved them as new `csv` files. Finally, using the recorded videos while performing the missions at each stop and the processed GPS records, we also identified the coordinates where the actions took place and the duration and type of these actions, also saving them in new `csv` files. Figure 5 presents a schematic overview of the data processing workflow, which results in open sharing of four datasets: the original GPS trajectories, the processed GPS trajectories, the mission stops, and the sociodemographic profiles of the participants.

	Teachers (EFiG)	Students (EFiG)	Residents (AAVV/EAGG)	Total
Number of participants	15	26	31	72
Gender Identity				
Woman	12	10	22	44
Man	3	16	9	28
Age range				
<17	0	26	7	33
17-35	1	0	2	3
36-50	12	0	6	18
51-65	2	0	8	10
>65	0	0	8	8
Do you know the neighbourhood?				
Much	5	15	21	41
Little	6	10	7	23
None	4	1	3	8
Do you frequently visit the neighbourhood?				
Every day	6	14	22	42
Weekly	2	11	3	16
Rarely	1	0	6	7
Almost never	6	1	0	7

Table 4. Socio-demographic statistics of all the participants. Number of participants according to the socio-demographic form. The columns represent the participating groups in the three experimental sessions that were held (teachers from EFiG, students from EFiG, and people from the neighbourhood jointly with old adults from AAVV and EAGG). The last column aggregates all participants.

	Students EFiG (26)							Teachers EFiG (15)			
	E1	E2	E3	E4	E5	E6	E7	P1	P2	P3	P4
Number of participants	3	3	4	4	4	4	4	3	3	5	4
Gender Identity											
Woman	3	2	1	2	0	1	4	3	1	5	3
Man	0	1	3	2	4	3	0	0	2	0	1
Age range											
<17	3	3	4	4	4	4	4	0	0	0	0
17-35	0	0	0	0	0	0	0	0	0	1	0
36-50	0	0	0	0	0	0	0	2	3	4	3
51-65	0	0	0	0	0	0	0	1	0	0	1
>65	0	0	0	0	0	0	0	0	0	0	0
Do you know the neighbourhood?											
Much	1	3	1	2	4	3	1	2	1	2	0
Little	2	0	3	2	0	0	3	1	1	0	4
None	0	0	0	0	0	1	0	0	1	3	0
Do you frequently visit the neighbourhood?											
Every day	0	3	1	1	3	3	3	1	1	0	4
Weekly	3	0	3	3	1	0	1	0	2	0	0
Rarely	0	0	0	0	0	0	0	0	0	1	0
Almost never	0	0	0	0	0	1	0	2	0	4	0

Table 5. Socio-demographic statistics of the school groups (EFiG). Number of participants according to the socio-demographic form for the school EFiG. There were seven groups of students (26 participants) and four of teachers (15 participants).

GPS records beyond the neighbourhood boundaries. During the first seconds of movement (still outside the neighbourhood), GPS records may be noisy due to poor GPS connection (the participants were inside the building) or because the devices were still adjusting the satellites' signal accuracy. However, experiments started slightly outside the established boundaries of the neighbourhood, in the communities' locations (less than 70 meters, see Fig. 1), and the manual removal of the records outside the boundaries was enough to avoid noisy GPS data. In addition, although the final meeting point for the experiment was located within the neighbourhood boundaries for all participating groups ("Plaça de la Font", see Fig. 1), the EFiG school participants finished

AAVV and EAGG (31)								
	G1	G2	G3	G4	G5	G6	G7	G8
Number of participants	5	3	3	3	3	3	6	5
Gender Identity								
Woman	3	2	1	3	3	2	4	4
Man	2	1	2	0	0	1	2	1
Age range								
<17	0	0	0	0	0	0	4	3
17–35	0	0	0	0	0	0	1	1
36–50	0	0	2	0	1	1	1	1
51–65	3	2	1	1	1	0	0	0
>65	2	1	0	2	1	2	0	0
Do you know the neighbourhood?								
Much	0	3	3	2	2	0	6	5
Little	5	0	0	1	1	0	0	0
None	0	0	0	0	0	3	0	0
Do you frequently visit the neighbourhood?								
Every day	0	2	3	3	3	0	6	5
Weekly	0	1	0	0	0	2	0	0
Rarely	5	0	0	0	0	1	0	0
Almost never	0	0	0	0	0	0	0	0

Table 6. Socio-demographic statistics of participating groups from the community (AAVV/EAGG). Number of participants according to the socio-demographic form for each group of the two local communities (AAVV and EAGG). There were eight groups (31 participants).

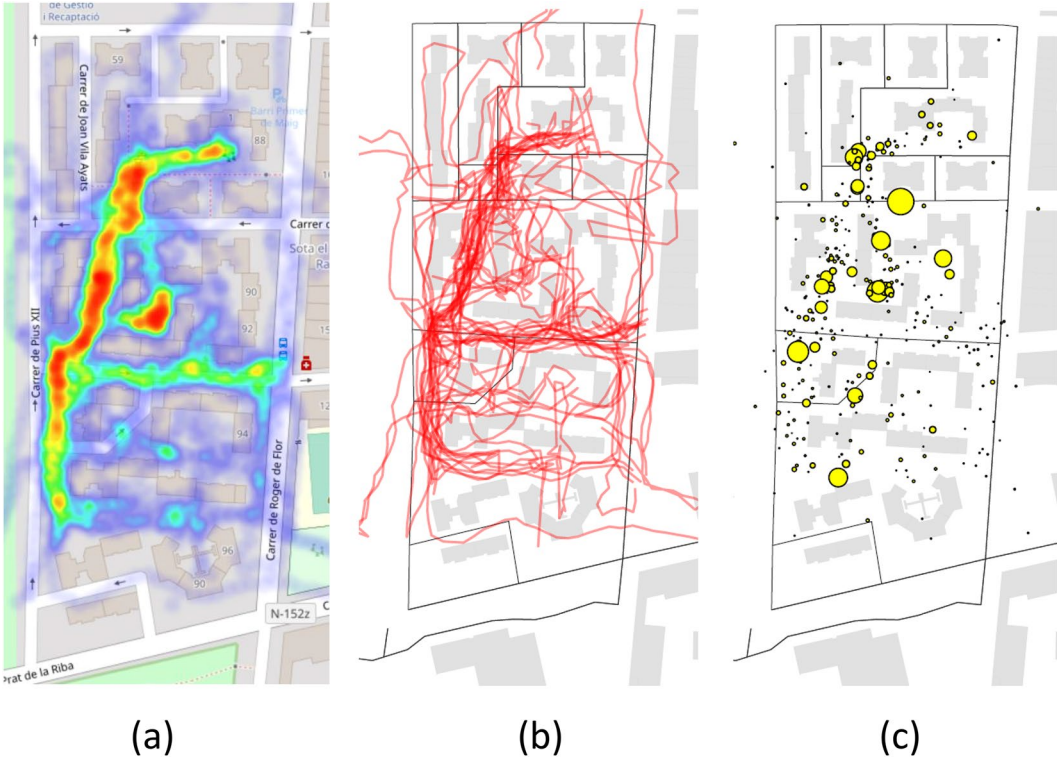


Fig. 4 GPS trajectories and stops of all participating groups. **(a)** Heat-map of the GPS locations (in motion). **(b)** Display of all trajectories in the neighbourhood. **(c)** All the GPS locations labelled as stopped (the size of the yellow circle depends on the duration of the stop: the larger the circle, the longer the stop).

their trajectories outside the boundaries, at the school itself. In this case, it was necessary to remove not just the first data records of the school groups' trajectories but the last part as well (outside the established limits). Figure 6a shows an example of a journey started and finished at the school. The GPS records outside the limits

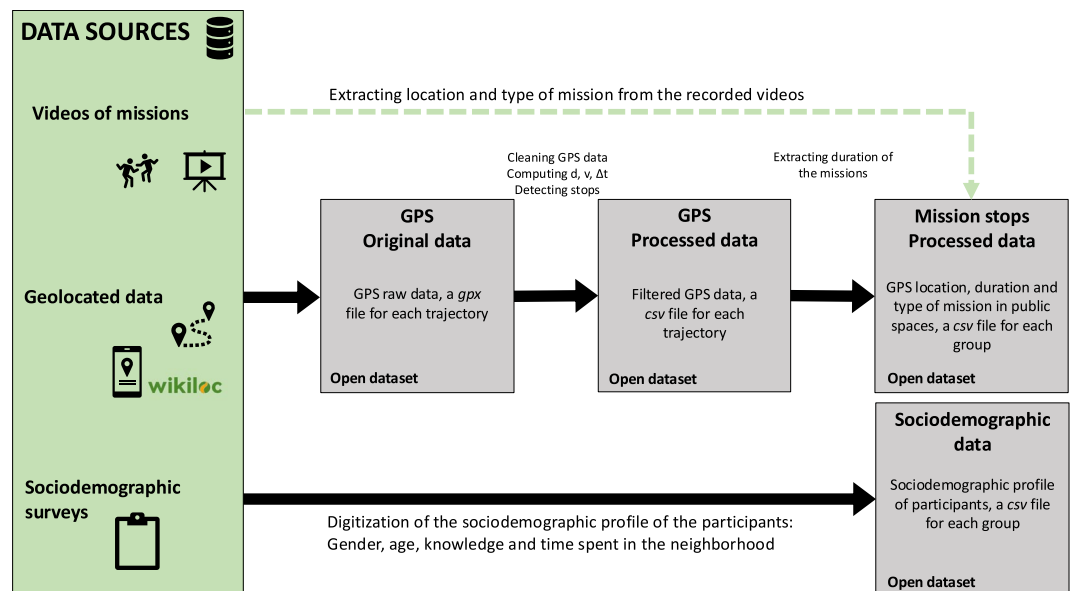


Fig. 5 Schematic diagram of data processing. We provide both raw and processed GPS trajectories. Based on the processed GPS data and the videos recorded during the missions, we construct and share an additional dataset that includes the location, duration, and type of mission at each stop. We also recorded, digitalized and share the sociodemographic profile of the participants.

of the district were carefully removed (see Fig. 6b). It should be noted that one group of teachers (group P1), went outside the established limits during the experiment, walking through the left side of the neighbourhood (see Figs. 1 and 4) where the football pitch and a green area are located. We decided to keep the full trajectory because they carried out some of the actions in that space, and also to avoid splitting the trajectory into two different transects.

In this cleaning process, a total of 551 GPS records were removed from the dataset, thus reducing the final number of GPS records from 3 532 to 2 981. Table 3 shows the number of GPS records and the number of stops before and after this filtering process.

Definition of a stop. Data processing also identified the stops made during the trip. We tagged each GPS record either as stopped (“stop”) or in motion (“moving”). The Wikiloc App has an automatic pause algorithm, when it detects that there is no movement. We performed reverse engineering to detect the stops:

1. We first obtain the difference between consecutive GPS timestamps in seconds, $\Delta_i(t) = t_{i+1} - t_i$. Note that $\Delta_i(t)$ is computed in a time-advanced way: the time difference of the record at the time stamp t_i is the time difference between the timestamps t_i and t_{i+1} . The duration is irregular and can vary from a few seconds (4, 5, 6...) to long durations, when participants perform the action related to each action.
2. Then we consider that the device at the timestamp t_i is stopped if $\Delta_i(t)$ is greater than or equal to a certain threshold Δ_c . After several in-the-field tests, we came up with a threshold of $\Delta_c = 10$ seconds as the optimal value to capture the stops performed during the trajectory (see Technical Validation section). We thus label each GPS record as “stop” or “moving” in the dataset.

Table 3 shows the number of participants, records (GPS locations) and the number of stops of each group (journey) before and after processing the collected data. The number of detected stops fell by 19.2% (from 420 to 339) when only the stops within the neighbourhood were considered. The number of stops includes not just stops made to complete the list of actions but also stops due to the actual mobility patterns (crossing the street, changing orientation, etc.).

Distance and instantaneous velocity. Data processing also includes the distance between geo-locations and the corresponding instantaneous velocity. Both are key variables for characterizing the mobility of each participating group. The Euclidean distance between consecutive GPS records reads

$$d(t) = |\vec{r}(t + \Delta(t)) - \vec{r}(t)|, \quad (1)$$

where $\vec{r}(t)$ is the GPS two-coordinate vector of a pedestrian position at time t (a given GPS timestamp) and we remove the label i henceforth for simplicity. We must take into account that the duration of consecutive records $\Delta(t)$ is irregular. We can also define instantaneous velocity as

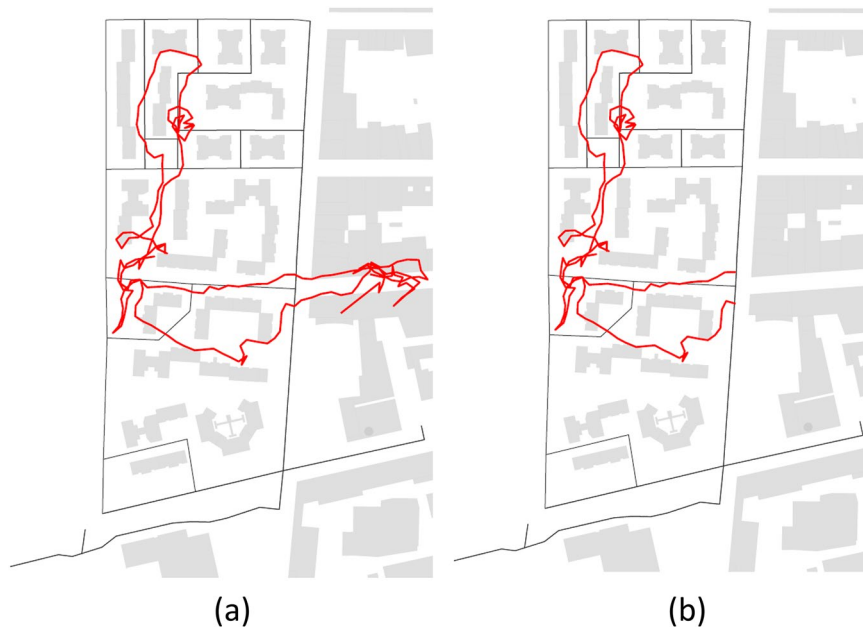


Fig. 6 Example of a trajectory clean-up. The trajectory corresponds to a group from the school EFiG. (a) Map representation of the original GPS location data from a single trajectory. (b) Map representation after the clean-up process, with the removal of the records outside the neighbourhood at the beginning and at the end of experiment.

$$v(t) = \frac{|\vec{r}(t + \Delta(t)) - \vec{r}(t)|}{\Delta(t)} = \frac{d(t)}{\Delta(t)}. \quad (2)$$

As for the case of $\Delta(t)$, the calculation of $d(t)$ and $v(t)$ is done in a time-advanced way.

Stops due to actions. We performed a further processing of the mobility data in which we detected the location of the stops to complete the actions and their duration. To do so, we used the videos recorded by the participating groups while they were performing the actions. Through the videos they recorded, we obtained and validated the location of the place they chose to perform each action. Then, by looking at the dataset and representing the stops and the trajectories in maps, we were able to identify which of the locations in the dataset labeled as “stop” correspond to the locations extracted from the videos (actions). We finally used the column of the time difference between consecutive records, $\Delta(t)$, to obtain the duration (in seconds) of the stop (action). If we detected two or more consecutive records labeled as stop, they were considered as a single stop. They were then merged as a single record with the location corresponding to the first detected stop and with a stop duration which is the aggregate stop duration.

Figure 7 shows an example of the procedure to determine the location of stops related to each action and its duration. In the first step, the locations of the stops (due to actions) extracted from the videos are represented in a map (purple stars in Fig. 7a) with the code of the action and the group. All the labeled stops in the GPS dataset are also represented (yellow dots). One can then easily relate the stops location from the video to the stops detected in the recorded trajectory and obtain the duration of these stops by analyzing the GPS dataset of the group. Occasionally, as stated above, consecutive stops may appear. In these cases, they are considered as a single stop, with their duration the sum of all durations. Then, for clarity, the stops due to the actions are represented (see Fig. 7b) indicating the duration and the code of the action and the group code name. A new `csv` file for each group is saved with the information of the location and the duration of the action stops. These new `csv` files do not contain the records of the whole trajectory but only the GPS locations of the stops corresponding to the festive actions (and their duration).

Data Records

The dataset⁴⁷ is available at CORA (*Repositori de Dades Obertes de Recerca*, <https://dataverse.csuc.cat>), with this section being the primary source of information on the availability and content of the data being described. The data repository⁴⁷ contains both raw and processed datasets, distributed in four different folders. The folder original data contains the raw GPS data collected by the participating groups (in `gpx` format). The folder processed data contains the same files as the original data folder after filtering and cleaning the GPS raw data (see section above) and in `csv` format. Finally, the folder called stops actions data contains information about the GPS location and the duration of the actions performed by the different participating groups and related to each action, in `csv` files. We also share the results of the socio-demographic form in the

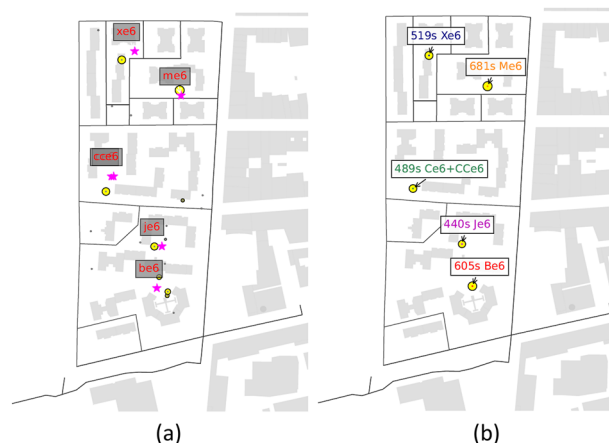


Fig. 7 Detecting the stops to complete the tasks and their duration. Example of a group of students (E6) from the school (EFiG). **(a)** All the stops detected from the GPS data (yellow dots) together with the locations of the stops to complete the actions extracted from the recorded videos (purple stars). With the map visualization and the information in the dataset, one can relate each stop location to each action and obtain the duration in seconds. **(b)** Only the stops due to the actions and their duration in seconds, as well as the code name of each action and group. Grey areas are building units.

folder called `sociodemographic_data`, which contains the `csv` files for each group and for the aggregate of all participants.

Raw GPS data. The `original_data` folder contains the raw data collected in the experiment. It consists of 18 different files in `gpx` format, corresponding to each one of the participating groups. One of the groups of teachers (P3), had problems with the tablet device which did not record the trajectory. For this reason there were 18 `gpx` files even though 19 groups participated. Although the trajectory was not recorded, thanks to the video recordings during the actions in the neighbourhood we have the exact locations where the P3 group stopped to perform the actions as well as an approximate duration of the stops. The `gpx` file name contains the anonymized nickname (code) of the participant group. The files with the code P (`p1.gpx`, `p2.gpx`, ...) represent the teachers of the school, and those with code E (`e1.gpx`, `e2.gpx`, ...) represent the students of school. Those with code G (`g1.gpx`, `g2.gpx`, ...) represent the people from the neighbourhood and old adults. Each of the `gpx` files has seven columns containing the information described in Table 7. Table 8 shows an example where the number of rows corresponds to the number of records, and so it is different for each participating group.

Filtered and processed GPS data. The `processed_data` folder contains the 18 `csv` files reporting the participants' trajectories after the filtering and cleaning process. In the processed data files, the columns `cmnt` and `desc` are removed as they do not give relevant information. These columns indicate the comments and the description of the trajectory respectively, and in our dataset they both have the same value as the `name` column, which indicates the name of the trajectory (group code, e.g. E1, see Tables 7 and 8). The column `ele` indicates the elevation (in meters) and is preserved although it was not used in our current related research. Four new columns are also included (see Table 9), representing the distance, velocity and time increment between consecutive GPS records and the label of each record indicating whether it is "stop" or "moving". The number of columns is then increased to 9. We note again that $\Delta(t)$, $d(t)$ and $v(t)$ are calculated in a time-advanced way (cf. Eqs. (1) and (2)): the time difference of the record at the time stamp t_i is the time difference between the timestamps at locations $t_i + \Delta_i(t)$ and t_i (the same applies for the distance and velocity). For this reason, the last record (last row) does not have these three values reported. Table 10 shows an example of the processed data. The processed `csv` files are saved with the same name as the original data (using the code of the group) but adding the suffix `processed` (e.g., `g4_processed.csv`).

Stop duration and actions. The `stops_actions_data` folder contains 19 `csvs` processed files, corresponding to each of the participating groups. The `csv` files are saved using the code name of the group and the suffix `stops` (e.g., `e6_stops.csv`). Each file provides detailed information about the stops the groups made to complete the festive actions: the geo-location of the stop, the duration and to which action it corresponds (see Table 11 as an example). It is worth mentioning that some groups carried out simultaneous actions at the same location. This is indicated in the action code (e.g., `Cp4 + CCp4`, see Table 11). To determine the location of the stops and the action that took place, we used the videos recorded by the participating groups as they carried out each of the actions around the neighbourhood (see Fig. 7). The `csv` files only include the stops due to the actions carried out by the participants, that were of major urbanistic relevance. The stops due to the micro-mobility of the participants (due to traffic lights, crossing the street, reorientation processes, etc.) are not included. However, they can be identified with the `stop` column of the processed `csv` files with the GPS data records. Tables 12, 13

Column	Description
lat	Latitude coordinate of each record in degrees
lon	Longitude coordinate of each record in degrees
name	Code that identifies the group
cmt	GPS comment of the waypoint
desc	A text description of the element
ele	Elevation in meters of the record
time	Timestamp of each record in YYYY:MM:DD HH:MM:SS format (UTC)

Table 7. Columns of the raw `gpx` files. Description of the seven columns of the `gpx` files from the original data, collected using the Wikiloc App.

	lat	lon	name	cmt	desc	ele	time
0	41.603439	2.283413	E1	E1	E1	138.518	2022-04-29T15:39:40Z
1	41.603425	2.283494	E1	E1	E1	139.298	2022-04-29T15:39:43Z
2	41.603432	2.283558	E1	E1	E1	139.917	2022-04-29T15:39:45Z
...	...	...	...	...	...	...	...
216	41.603345	2.283466	E1	E1	E1	142.929	2022-04-29T16:35:37Z
217	41.603397	2.283473	E1	E1	E1	143.060	2022-04-29T16:52:34Z
218	41.603241	2.283207	E1	E1	E1	143.393	2022-04-29T16:52:38Z

Table 8. Example of a raw `gpx` file display. This is a group of students from the school (EFiG), with 219 records (GPS locations).

Column	Description
Δt	Time difference between consecutive timestamps t , in seconds
d	Euclidean distance at a time t between consecutive GPS locations, in metres
v	Instantaneous velocity at time t (distance over time lapse), in metres/second
stops	Indicates if the participant is stopped ("stop") or in motion ("moving")

Table 9. New columns of the filtered and processed `csv` files. Description of the four new columns added to the processed dataset.

	lat	lon	name	ele	time	Δt	d	v	stops
0	41.602788	2.282938	G7	143.753	2022-04-30 11:28:20+00:00	6.0	6.117981	1.019664	moving
1	41.602843	2.282936	G7	143.716	2022-04-30 11:28:26+00:00	6.0	5.306684	0.884447	moving
2	41.602889	2.282919	G7	143.832	2022-04-30 11:28:32+00:00	7.0	5.524785	0.789255	moving
...	...	...	...	...	...	...	...	...	...
153	41.604316	2.282312	G7	154.909	2022-04-30 13:02:44+00:00	5.0	5.823689	1.164738	moving
154	41.604325	2.282243	G7	155.159	2022-04-30 13:02:49+00:00	4.0	5.269312	1.317328	moving
155	41.604364	2.282207	G7	155.178	2022-04-30 13:02:53+00:00	NaN	NaN	NaN	stop

Table 10. Example of a processed and cleaned `csv` file display. This is a participating group from the association of neighbours and older people (AAVV, EAGG), with 156 GPS records.

and 14 show the duration of each action carried out by each of the school groups (teachers and students, EFiG) and by each of the local association groups (AAVV and EAGG).

Socio-demographic data. The data repository⁴⁷ also contains the answers to the socio-demographic form that each participating group filled in. In the folder `sociodemographic` data there are 20 `csv` files, 19 of which correspond to each of the groups. The file name contains the group code (e.g., `g5_sociodem.csv`). The last `csv` file corresponds to the sociodemographic data of the aggregate of all participants (`all_sociodem.csv`). Table 15 shows an example of a `csv` file visualization of the socio-demographic form (group G7). Tables 4, 5 and 6 show the socio-demographic statistics of each participating group and for the aggregate of all participants.

	accio	longitud	latitud	durada
0	Cp4+CCp4	2.281650	41.603890	361.0
1	Jp4	2.282149	41.603109	197.0
2	Mp4	2.281773	41.603244	1811.0
3	Xp4	2.282434	41.603065	295.0
4	Bp4	2.282149	41.604199	258.0

Table 11. Example of a processed and cleaned `csv` file display, with the stops due to the actions. This group comprised teachers from the school (EFiG), with five stops to perform the actions. The column “accio” is the code of the action (see Table 1) followed by the group code (e.g. p4). The columns “longitud” and “latitud” represent respectively the longitude and latitude coordinates of the stop and the column “durada” is the duration of the stop in seconds.

Teachers		
Code Group	Task	Duration (s)
P1	B	557
	J	2151
	J	231
	X+M	525
P2	B	151
	J	113
	X	200
	M	317
	C+CC	1240
P3	B	166
	J + M	1060
	X	330
	C+CC	120
P4	B	258
	J	197
	X	295
	M	1811
	C + CC	361

Table 12. Duration of the actions (teachers from EFiG). For each group of the teachers from the school (EFiG), the table contains the list of festive actions they completed (with the code, see Table 1) and the duration of each action in seconds.

Technical Validation

In the Filtering and Processing GPS Data Section, we introduced a method to classify GPS records as “stop” or “moving”, based on a temporal threshold Δ_c . This parameter defines the minimum time interval between two consecutive recorded positions required to label a record as a stop. Through empirical validation, we selected $\Delta_c = 10$ seconds as the optimal threshold.

The selection was guided by a series of controlled in-the-field tests in which trajectories were recorded using the Wikiloc App, incorporating predefined stops at known locations. The recorded `gpx` data were subsequently analyzed using the stop classification proposed criterion, testing various threshold values ($\Delta_c = 6, 7, 8, 9, 10, \dots$ s). As one can expect, thresholds that are too small tend to overestimate the number of stops, introducing false extra stops, whereas excessively large thresholds fail to detect actual stopping events. Our approach aimed to identify the smallest value of Δ_c that reliably detects all known stops without introducing spurious ones.

An illustrative case is presented in Fig. 8, based on a six-minute trajectory through the “L’Esquerra de l’Eixample” neighborhood in Barcelona, during which two deliberate 30-second stops were made at traffic lights. With $\Delta_c = 6$ s (Fig. 8a), more than two stops are detected erroneously. Conversely, $\Delta_c = 50$ s (Fig. 8c) fails to detect any stop. The intermediate value of $\Delta_c = 10$ s (Fig. 8b) accurately identifies the two planned stops. Additional experiments across different urban settings and devices (using Wikiloc App) consistently supported $\Delta_c = 10$ s as a robust and generalizable threshold for stop detection in pedestrian trajectories.

Alternative approaches to stop detection are available in the literature, including those based on instantaneous velocity or distance thresholds, particularly valuable when the recording device does not automatically pause during stationary periods¹⁸. To validate the effectiveness of our proposed method, we analyzed the complementary cumulative distribution function (CCDF) of instantaneous velocities for GPS records classified as either “moving” or “stop”. As illustrated in Fig. 9a, the stops identified correspond to the lowest velocity values within the dataset. Specifically, we found that 95% of the velocities classified as “stop” are below 0.57 m/s, whereas 95% of those corresponding to “moving” exceed 0.65 m/s. This distinct separation between the two

Students		
Code Group	Task	Duration (s)
E1	B	151
	J	100
	J	249
	X	521
	M + C + CC	985
E2	B	670
	J	342
	X + M + C + CC	2345
E3	B	1130
	J	262
	X	39
	M + C	731
	CC	1240
E4	B + J	601
	B	87
	J + X	403
	M	963
	C + CC	1090
E5	B + J	486
	J	117
	X	817
	M + C + CC	858
E6	B	605
	J	440
	X	519
	M	681
	C + CC	489
E7	B + M	322
	B + J	927
	X	492
	C + CC	1065

Table 13. Duration of the actions (students from EFiG). For each group of the students from the school (EFiG), the table contains the list of festive actions they completed (with the code, see Table 1) and the duration of each action in seconds.

distributions supports the robustness of our stop-detection approach, indicating that the detected stops are consistently associated with the lowest velocity measurements. Figure 9b illustrates the temporal evolution of velocity for one of the participating groups throughout the experiment. The pattern reveals fluctuations around typical pedestrian speed levels, interspersed with prolonged periods without data points. These gaps correspond to stops made to complete the various proposed tasks. During these stationary periods, the mobility data collection App (Wikiloc) automatically pauses the recording when it detects no movement. The red markers indicate velocities classified as “stops”, further supporting that these events are associated with the lowest velocity values observed along the trajectory.

Analyzing instantaneous velocities is also an effective strategy for identifying outliers, which often arise due to GPS inaccuracies or signal instability. Figure 9c and d illustrate the temporal evolution of instantaneous velocity for two distinct trajectories (groups), both of which exhibit evident outliers with velocity values exceeding even 20 m/s, which are clearly inconsistent with typical pedestrian movement and indicative of measurement error. Although no additional filtering was applied in this study to remove such anomalies, several established techniques exist for trajectory data cleaning. A widely used approach consists of iteratively discarding velocity records that exceed a predefined threshold, followed by recalculation of velocities until all remaining data fall within acceptable bounds⁸. An alternative strategy for handling outliers involves removing the corresponding records and segmenting the trajectory into distinct transects. Regardless of the specific approach, such outliers should be carefully addressed in studies aimed at characterizing pedestrian movement, as their presence could distort statistical analyses and derived mobility patterns.

In our complete dataset, which includes 2963 instantaneous velocity records, only 69 values (2%) are equal to or exceed 3 m/s. 36 (1%) are greater or equal than 4 m/s. 18 values (0.98%) are equal to or exceed 5 m/s. 13 (0.44%) are greater than or equal to 8 m/s, 10 (0.34%) reach or exceed 10 m/s, and just 6 (0.20%) surpass 20 m/s. These values indicate that such high-velocity outliers are statistically negligible.

Local People		
Code Group	Action	Duration (s)
G1	B + J	909
	B	100
	J	479
	J	1882
	X	1896
	M + C	809
G2	B	427
	J	400
	X	360
	M	946
	C	677
	CC	82
G3	J + M	1845
	J	512
	X	96
	C + CC	2669
G4	X + M	1243
	C + CC	477
G5	B + J	690
	J	267
	X + M	1071
	C	1223
	CC	277
G6	B	828
	J + M	2677
	X + C	795
G7	B + J + X + M + C	3130
	CC	1955
G8	B	918
	J	713
	X	400
	M + C + CC	656

Table 14. Duration of the actions (local people from AAVV and EAGG). For each group of the participants from the associations (AAVV and EAGG), the table contains the list of festive actions they completed (with the code, see Table 1) and the duration of each action in seconds.

	group	community	question	answer	number_of_participants
0	g7	AAVV/EAGG	Gender Identity	Woman	4
1	g7	AAVV/EAGG	Gender Identity	Man	2
2	g7	AAVV/EAGG	Gender Identity	Non-binary	0
...	...	...	...	...	...
4	g7	AAVV/EAGG	Age range	<17	4
5	g7	AAVV/EAGG	Age range	18-35	1
...	...	...	...	...	...
14	g7	AAVV/EAGG	Do you frequently visit the neighbourhood?	Rarely	0
15	g7	AAVV/EAGG	Do you frequently visit the neighbourhood?	Almost never	0

Table 15. Example of a socio-demographic csv file display. Answers to the socio-demographic form of a group of 6 participants from the AAVV/EAGG communities. The csv file displays the code of the group, the community to which it belongs, the socio-demographic question, the possible answers and the number of participants who answered.

The use of videos recorded by the participants themselves during the execution of their tasks enabled us to validate our stop detection method. Since each video is geotagged, we were able to verify the location of each mission and compare it with the processed GPS data, from which the precise duration of each activity could be

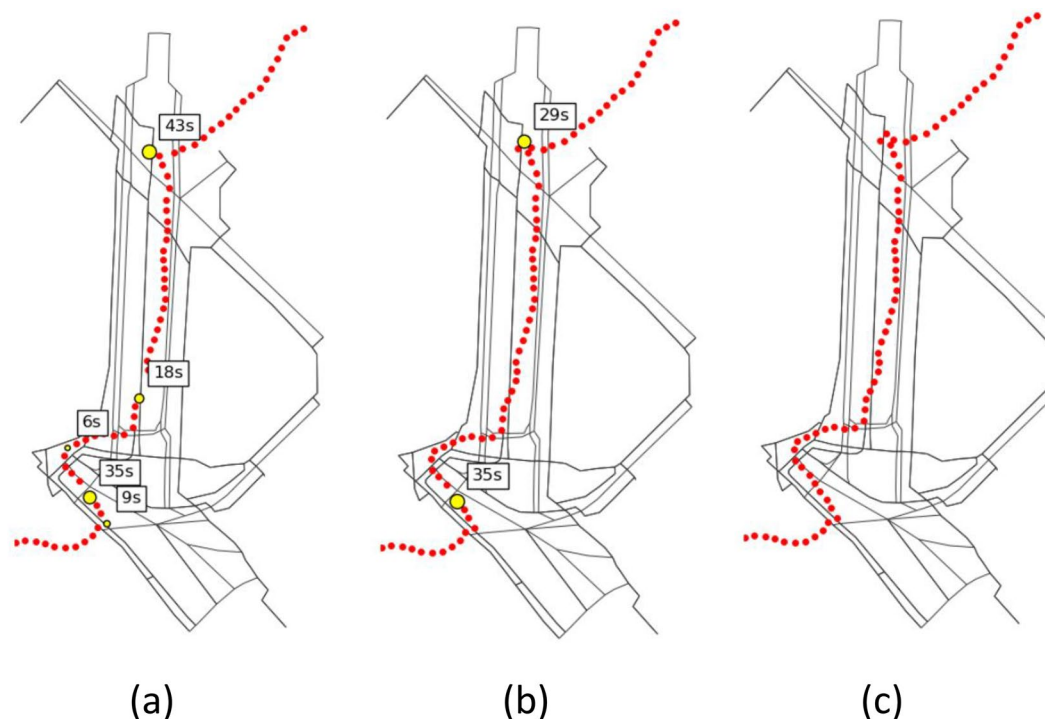


Fig. 8 Example of the method to capture the stops in a trajectory. Various thresholds Δ_c are applied to detect which is the minimum time difference to capture a stop. The red dots represent the trajectory and the yellow dots the stops captured by the method proposed. The duration in seconds is also included. In this particular journey, two stops were made due to traffic lights, of about 30 seconds. (a) Corresponds to a threshold of $\Delta_c = 6$ s, which overestimates the number of stops. (b) Corresponds to the optimal threshold, $\Delta_c = 10$ s, which is the lowest one that captures the two stops. (c) Corresponds to a threshold of $\Delta_c = 50$ s, which underestimates the number of stops, without detecting the two stops.

determined. Additionally, the videos allowed us to identify the type of mission performed (see Table 1), which may be useful for future studies. Due to privacy considerations and the need to preserve participant anonymity, the recorded videos are not publicly available.

Usage Notes

The data collected from the experiment represent a rich source of exploratory mobility information for the study of people's micro-motives around the neighbourhood "Primer de Maig" in the city of Granollers. The mobility data collected in the experiment can be accessed in the folder `original data` from the data repository. Each `gpx` file represents a participant group and their trajectory record. Each individual trajectory belongs to a different group with different socio-demographic characteristics. The Python notebook called `Data Processing and Cleaning.ipynb` from the GitHub repository contains a description of the treatment and data clean-up discussed in this Data Descriptor, as well as the scripts to reproduce all the trajectory representation on maps (Figs. 4, 6 and 8). Each processed individual trajectory is saved with the group name (e.g., `p2_processed.csv`) in a new `csv` file contained in the `processed data` folder. Therefore, clean pedestrian mobility data can also be accessed through this folder. In addition, the Python Notebook mentioned above contains the script for the calculation of the time increment, distance and velocity between consecutive records, as well as the calculation of the stops. These variables are added as new columns to the processed `csv` file of each participant group. The code repository also contains different Jupyter notebooks with the necessary functions and scripts to study and characterize the participants' movement through some statistical patterns which are not discussed in this Data Descriptor. All these files contain the detailed description of the study and the explanation of the code. Finally, through the `stops actions` data folder one can access the locations and the duration of the stops made by each of the participating groups to complete the list of actions, and through the `sociodemographic data` folder one can access the socio-demographic statistical data.

Code availability

The https://github.com/ferranlarroyaub/mobility_granollers repository contains the Python code and scripts for processing the input data and for replicating the statistical analysis and the figures in this document. The 3.8 Python version is used to build the code with the main libraries `networkx` and `osmnx` to plot the journeys on OpenStreet maps, `pandas` and `numpy` to process, clean and analyse the data in Data-Frame format and performing the basic statistical calculations, and `scipy` for more advanced calculations such as fitting models to

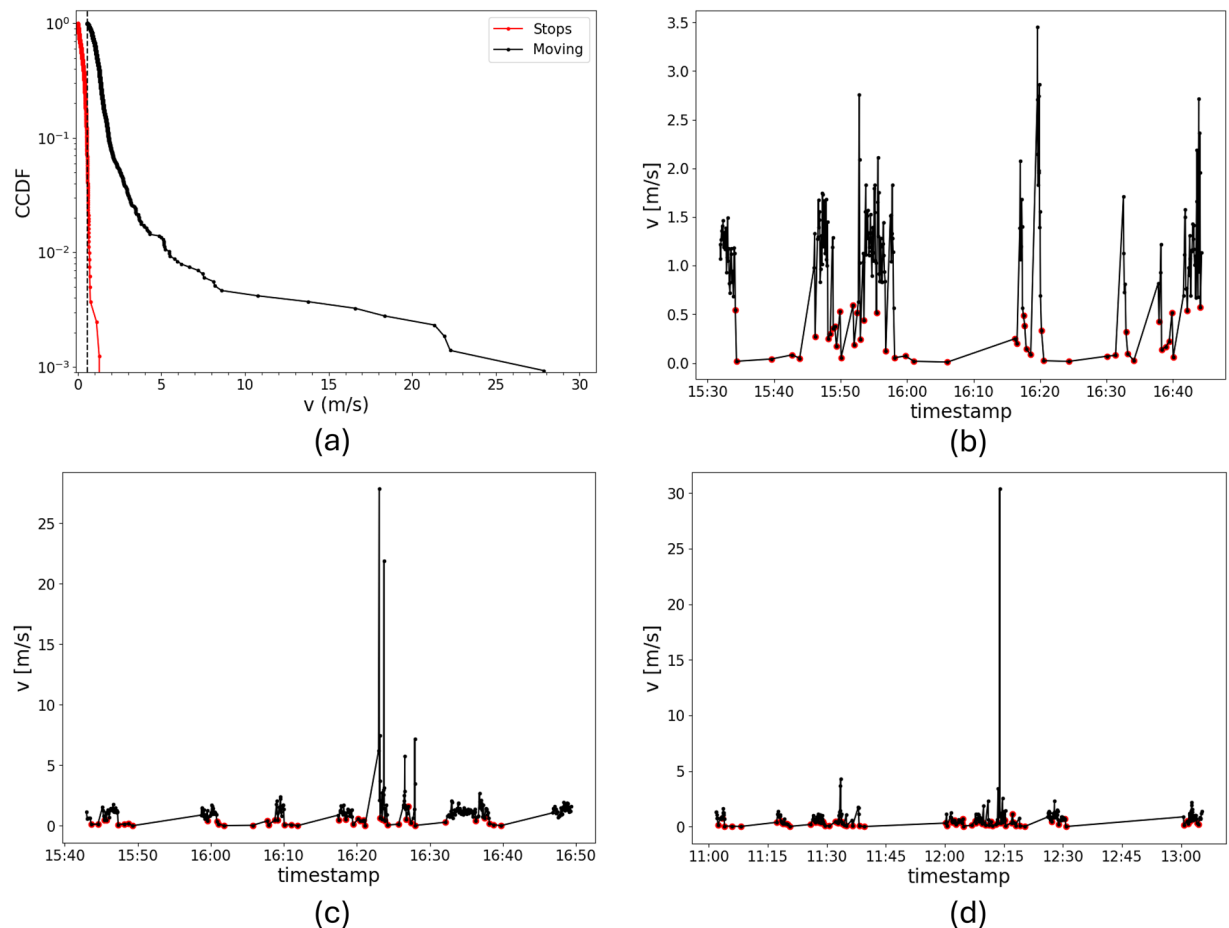


Fig. 9 Complementary Cumulative Distribution Function (CCDF) of instantaneous velocities and velocities time series. The instantaneous velocity between consecutive GPS records is computed using Eq. (2). (a) Corresponds to the CCDF of instantaneous velocity for those records labeled as “moving” (black line) and for those labeled as “stop” (red line). The vertical dashed line indicates the minimum value of the instantaneous velocity labeled as “moving” (0.55 m/s). (b) Temporal evolution of the instantaneous velocity for one of the participating groups throughout the experiment. Red dots correspond to those records classified as “stops”. (c) and (d) Two examples of the temporal evolution of instantaneous velocity for two different trajectories (groups). In both cases, the data include outliers—velocity records exceeding 20 m/s—which are clearly inconsistent with pedestrian movement and likely result from GPS errors or signal instability.

the empirical data and matplotlib for plotting purposes. The Python code is built in different Jupyter notebook files which contain the detailed description of the study and the code documentation.

Received: 24 October 2024; Accepted: 2 June 2025;

Published online: 19 June 2025

References

1. Ravazzoli, E. & Torricelli, G. P. Urban mobility and public space. A challenge for the sustainable liveable city of the future. *J. Public Space* **2**(2), 37–50 (2017).
2. Carmona, M. Principles for public space design, planning to do better. *Urban Des. Int.* **24**, 47–59 (2019).
3. Jiang, S. *et al.* The TimeGeo modeling framework for urban mobility without travel surveys. *Proc. Natl. Acad. Sci. USA* **113**, E5370–E5378 (2016).
4. Batty, M. *The new science of cities* (MIT Press, Cambridge, 2013).
5. Bettencourt, L. M. *Introduction to Urban Science: Evidence and Theory of Cities as Complex Systems* (MIT Press, Cambridge, 2021).
6. Bai, X. *et al.* Defining and advancing a systems approach for sustainable cities. *Curr. Opin. Environ. Sustain.* **23**, 69–78 (2016).
7. Barthélemy, M. *The Structure and Dynamics of Cities* <https://doi.org/10.1017/9781316271377> (Cambridge Univ. Press, 2016).
8. Liu, C. *et al.* Revealing spatiotemporal interaction patterns behind complex cities. *Chaos* **32**(8) (2022).
9. Bibri, S. E. & Krogstie, J. A Novel Model for Data-Driven Smart Sustainable Cities of the Future: A Strategic Roadmap to Transformational Change in the Era of Big Data. *Future Cities Environ.* **7**(1), 1–25 (2021).
10. Barbosa, H. *et al.* Human mobility: Models and applications. *Phys. Rep.* **734**, 1–74 (2018).
11. Moro, E., Calacci, D., Dong, X. & Pentland, A. Mobility patterns are associated with experienced income segregation in large US cities. *Nat. Commun.* **12**(1), 4633 (2021).
12. Xu, Y., Clemente, R. D. & González, M. C. Understanding vehicular routing behavior with location-based service data. *EPJ Data Sci.* **10**, 12 (2021).

13. Kraemer, M. U. G. *et al.* Mapping global variation in human mobility. *Nat. Hum. Behav.* **4**(8), 800–810 (2020).
14. Hunter, R. F. *et al.* Effect of COVID-19 response policies on walking behavior in US cities. *Nat. Commun.* **12**(1), 3652 (2021).
15. Yang, Y., Pentland, A. & Moro, E. Identifying latent activity behaviors and lifestyles using mobility data to describe urban dynamics. *EPJ Data Sci.* **12**, 15 (2023).
16. Li, R. *et al.* Gravity model in dockless bike-sharing systems within cities. *Phys. Rev. E* **103**(1), 012312 (2021).
17. Codling, E. A., Plank, M. J. & Benhamou, S. Random walk models in biology. *J. R. Soc. Interface* **5**(25), 813–834 (2008).
18. Gutiérrez-Roig, M. *et al.* Active and reactive behaviour in human mobility: the influence of attraction points on pedestrians. *R. Soc. Open Sci.* **3**(7), 160177 (2016).
19. Chen, P., Wang, W., Qian, C., Cao, M. & Yang, T. Gravity-based models for evaluating urban park accessibility: Why does localized selection of attractiveness factors and travel modes matter? *Environ. Plan. B: Urban Anal. City Sci.* **51**(4), 904–922 (2024).
20. Armengol, J. M. *et al.* City-scale assessment of pedestrian exposure to air pollution: A case study in Barcelona. *Urban Clim.* **58**, 102183 (2024).
21. Yabe, T., Bueno, B. G. B., Dong, X., Pentland, A. & Moro, E. Behavioral changes during the COVID-19 pandemic decreased income diversity of urban encounters. *Nat. Commun.* **14**(1), 2310 (2023).
22. Tobin, M. *et al.* Rethinking walkability and developing a conceptual definition of active living environments to guide research and practice. *BMC Public Health* **22**(1), 450 (2022).
23. Thomas, D. *Placemaking: An urban design methodology* <https://doi.org/10.4324/9781315648125> (Routledge, New York, 2016).
24. Paez, R. *Operative Mapping: Maps as Design Tools* (ACTAR publishers, Barcelona, 2019).
25. Valtchanova, M. & Paez, R. *Operative Mapping and Collaborative Actions as Design Tools for Critical Socio-spatial Urban Interventions. Proceedings of the International Conference: Repurposing Places for Social and Environmental Resilience* (eds. Karandinou, A. & Lam, F.) 133–136 (CounterArchitecture, University of East London, Arup, London, 2023).
26. Paez, R., Valtchanova, M., Larroya, F. & Perelló, J. *Maps as Design Tools: Space, Time, and Experience. In: The Routledge Handbook of Cartographic Humanities* (eds. Rossetto, T. & Lo Presti, L.) 172–181, <https://doi.org/10.4324/9781003327578-22> (Abingdon-on-Thames/New York: Routledge, 2024).
27. Vohland, K. *et al.* *The Science of Citizen Science* <https://doi.org/10.1007/978-3-030-58278-4> (Springer, Cham, 2021).
28. Irwin, A. No PhDs needed: how citizen science is transforming research. *Nature* **562**(7726), 480–482 (2018).
29. Kapenekakis, I. & Chorianopoulos, K. Citizen science for pedestrian cartography: collection and moderation of walkable routes in cities through mobile gamification. *Hum. Cent. Comput. Inf. Sci.* **7**(1), 10 (2017).
30. Palmer, L. New inroads on community-centric placemaking. *Nat. Cities* **1**, 2–4 (2024).
31. Ballerini, L. & Bergh, S. I. Using citizen science data to monitor the Sustainable Development Goals: a bottom-up analysis. *Sustain. Sci.* **16**(6), 1945–1962 (2021).
32. Christine, D. I. & Thinyane, M. Citizen science as a data-based practice: A consideration of data justice. *Patterns* **2**(4) (2021).
33. Sandercock, L. *Mapping possibility: Finding purpose and hope in community planning* (Routledge, London, 2023). <https://doi.org/10.4324/9781003325451>
34. Albert, A., Balazs, B., Butkeviciénė, E., Mayer, K. & Perelló, J. *Citizen Social Science: New and Established Approaches to Participation in Social Research. In: Vohland, K. et al. The Science of Citizen Science* Ch. 7, https://doi.org/10.1007/978-3-030-58278-4_7 (Springer, Cham, 2021).
35. Bonhoure, I., Cigarini, A., Vicens, J., Mitats, B. & Perelló, J. Reformulating computational social science with citizen social science: the case of a community-based mental health care research. *Humanit. Soc. Sci. Commun.* **10**(1), 1–14 (2023).
36. Paez, R., Valtchanova, M., Perelló, J., Larroya, F. & Sánchez, E. *Civic Placemaking 3: Disseny, espai públic i cohesió social* https://doi.org/10.46467/ElisavaResearch_CivicPlacemaking3 (2022).
37. Sagarra, O., Gutiérrez-Roig, M., Bonhoure, I. & Perelló, J. Citizen Science Practices for Computational Social Science Research: The Conceptualization of Pop-Up Experiments. *Front. Phys.* **3**, 93 (2016).
38. Perelló, J. New knowledge environments: On the possibility of a citizen social science. *Metode Sci. Stud. J.* **12**, 25–31 (2022).
39. Perelló, J., Larroya, F., Bonhoure, I. & Peter, F. Citizen science for social physics: digital tools and participation. *Eur. Phys. J. Plus* **139**(7), 572 (2024).
40. Larroya, F. *et al.* Home-to-school pedestrian mobility GPS data from a citizen science experiment in the Barcelona area. *Sci. Data* **10**(1), 428 (2023).
41. Wang, J. & Kwan, M. P. Daily activity locations k-anonymity for the evaluation of disclosure risk of individual GPS datasets. *Int. J. of Health Geogr.* **19**(1), 1–14 (2020).
42. Sennett, R. *Building and Dwelling: Ethics for the City* (Farrar, Straus & Giroux, New York, 2018).
43. Paez, R. Design as Playground: Exploring Spatial Design Through Playful Practices. *Space Cult.* **27**(2), 187–208 (2024).
44. Sepe, M. Placemaking, livability and public spaces. Achieving sustainability through happy places. *J. Public Space* **2**(4), 63–76 (2017).
45. Harley, J. B. *The new nature of maps: essays in the history of cartography* <https://doi.org/10.56021/9780801865664>
46. Pitidis, V., Coaffee, J. & Lima-Silva, F. Advancing equitable -resilience imaginaries' in the Global South through dialogical participatory mapping: Experiences from informal communities in Brazil. *Cities* **150**, 105015 (2024).
47. Larroya, F. & Perelló, J. Explorative pedestrian mobility GPS data from a citizen science experiment in a neighbourhood. *CORA. Repositori de Dades de Recerca* <https://doi.org/10.34810/data898> (2023).
48. Ertz, O. *et al.* Citizen participation & digital tools to improve pedestrian mobility in cities. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **46**, 29–34 (2021).
49. Ferster, C., Nelson, T., Laberee, K., Vanlaar, W. & Winters, M. Promoting crowdsourcing for urban research: Cycling safety citizen science in four cities. *Urban Sci.* **1**(2), 21 (2017).
50. Pappers, J. & De Wilde, L. Evaluating citizen science data: a citizen observatory to measure cyclists' waiting times. *Transp. Res. Interdiscip. Perspect.* **14**, 100624 (2022).
51. Pappalardo, L. *et al.* Returners and explorers dichotomy in human mobility. *Nat. Commun.* **6**(1), 8166 (2015).
52. Zhang, R. *et al.* Optimal allocation of local climate zones based on heat vulnerability perspective. *Sustain. Cities Soc.* **99**, 104981 (2023).

Acknowledgements

We thank the 72 volunteers from the school “Escola Ferrer i Guàrdia (EFiG)”, the associations “Associació de Veïns Sota el Camí Ral” and “Espai Actiu de la Gent Gran” for their participation. The study was partially supported by the Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación MCIN/AEI/10.13039/501100011033, grant number PID2019-106811GB-C33 [FL and JP]; by MCIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe”, grant number PID2022-140757NB-I00 [FL and JP]; and by Horizon Europe WIDERA program [SENSE, contract number 101058507, FL and JP]. We also acknowledge the support of Generalitat de Catalunya through Complexity Lab Barcelona, grant number 2021SGR00856 [FL and JP]; and by the ICI programme (Interculturalitat i Cohesió Social) funded by Fundació Bancària -la Caixa- [RP and MV].

Author contributions

Ferran Larroya - data acquisition - project conception - data validation - writing - proofreading. Josep Perelló - data acquisition - project conception - writing - proofreading. Roger Paez - data acquisition - project conception - writing - proofreading. Manuela Valtchanova - data acquisition - project conception - writing - proofreading.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Additional information

Correspondence and requests for materials should be addressed to F.L. or J.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025