City Research Online

City St George's, University of London

Citation: Alessandretti, L., Lehmann, S. & Baronchelli, A. (2018). Understanding the interplay between social and spatial behaviour. EPJ Data Science, 7(1), 36. doi: 10.1140/epjds/s13688-018-0164-6

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/20650/

Link to published version: https://doi.org/10.1140/epjds/s13688-018-0164-6

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Alessandretti et al. *EPJ Data Science* (2018) 7:36
https://doi.org/10.1140/epjds/s13688-018-0164-6

EPJ Data Science
a SpringerOpen Journal

LETTER Open Access

Understanding the interplay between social and spatial behaviour

Laura Alessandretti[1*], Sune Lehmann[1,2] and Andrea Baronchelli[3]

*Correspondence: lauale@dtu.dk
[1]Technical University of Denmark, Lyngby, Denmark
Full list of author information is available at the end of the article

Abstract

According to personality psychology, personality traits determine many aspects of human behaviour. However, validating this insight in large groups has been challenging so far, due to the scarcity of multi-channel data. Here, we focus on the relationship between mobility and social behaviour by analysing trajectories and mobile phone interactions of ∼1000 individuals from two high-resolution longitudinal datasets. We identify a connection between the way in which individuals explore new resources and exploit known assets in the social and spatial spheres. We show that different individuals balance the exploration-exploitation trade-off in different ways and we explain part of the variability in the data by the big five personality traits. We point out that, in both realms, extraversion correlates with the attitude towards exploration and routine diversity, while neuroticism and openness account for the tendency to evolve routine over long time-scales. We find no evidence for the existence of classes of individuals across the spatio-social domains. Our results bridge the fields of human geography, sociology and personality psychology and can help improve current models of mobility and tie formation.

Keywords: Human mobility; Social behaviour; Computational social science; Digital data; Personality traits

1 Intro

Our social and spatial behaviour are shaped by both internal and external constraints. On one hand, external factors [1] such as time, cognition, age or the need for food constrain our possibilities. On the other hand, we are driven by internal needs, purposes and preferences. Specifically, within personality psychology, it has been conjectured that personality traits play a key role in shaping our choices across various situations [2, 3].

In the social realm, individuals cope with cognitive and temporal constraints by establishing and maintaining connections in a distinctive [4, 5] and persistent [4] manner. For example, the size of an individual's social circle is bounded under ∼150, the so-called Dunbar number [6], but varies among individuals around this limit [7]. These differences result from an interplay between physical and extrinsic factors such as gender [8], age [9] and socio-economic status [10] as well as from stable individual dispositions underlying personality [11].

Spatially, individuals are characterised by an *activity space* of repeatedly visited locations within which they move during their daily activities [12], but this geo-spatial signature varies in size [13] and spatial shape [14]. However, unlike the social case, the conjecture

© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

that individuals' spatial behaviour is persistent in time [15] had not been verified until recently.

Here, we capitalise on the recent discovery that the size of the activity space is conserved and correlates with the social circle size [16] to test the conjecture that the same personality dispositions in part determine social and spatial behaviour. We test this theory by analysing two long-term datasets consisting of ∼1000 individuals mobility trajectories and their phone interactions (for previous studies see Sect. 2 below).

First, we test the hypothesis that the strategies individuals adopt in order to choose where to go and with whom to interact are similar. Then, we identify and characterise the prevailing socio-spatial profiles appearing in the datasets. Finally, we show that socio-spatial profiles can be partially explained by the widely adopted big-five personality trait model, often used to describe aspects of the social and emotional life [7, 11, 17–22]. In Sect. 2, we review the relevant literature; in Sect. 3 we describe data collection and pre-processing, and we provide details of the methods implemented; in Sect. 4 we present our findings.

2 State of the art

Individual-level variability in social and spatial behaviour has mostly been investigated in isolation so far, with few notable efforts to reconcile the two. Here, we briefly review the empirical findings in the two domains.

2.1 The social domain

Individuals deal with limited time and cognitive capacity resulting in finite social networks [6, 23] by distributing time unevenly across their social circle [4, 24–28]. While this is a shared strategy, there is clear evidence for individual-level variation. First, social circles vary in terms of diversity: they differ in size [7]—within a maximum upper-bound of ∼150 individuals [6]—and in structure [4, 29]. Second, individuals display different attitudes towards exploration of social opportunities as they are more or less keen on creating new connections [30–33]. Finally, individuals manage social interactions over time in different ways. Some are characterised by high level of stability as they maintain a very stable social circle, while others renew their social ties at high pace [5].

These heterogeneities can be partially explained by factors including gender [8, 34], age [9, 35, 36], socio-economic status [10, 37] and physical attractiveness [38]. Moreover, as conjectured by personality psychologists [2, 39], differences in personalities partially explain the variability in social circle composition [7, 11, 17, 40–44], and the different attitudes towards forming [30, 45], developing [20, 46] and replacing [29] social connections. It is worth noticing that many of these findings are recent, resulting from the analysis of digital communication traces.

2.2 The spatial domain

Constraints including physical capabilities, the distribution of resources, and the need to coordinate with others limit our possibilities to move in space [1]. Individuals cope with these limitations by allocating their time within an activity space of repeatedly visited locations [47], whose size is conserved over several years according to a recent study based on high-resolution trajectories [16], and previous ones based on unevenly sampled and low spatial resolution data [48, 49]. The activity space varies across individuals in terms

of size [16] and shape [14]: it was shown that two distinct classes of individuals, *returners* and *explorers*, can be identified based on their propensity to visit new locations, similarly to the social domain [5]. Heterogeneities in spatial behaviour can be explained in terms of gender [50], age [51, 52], socio-economic [35, 53] and ethnic [54] differences. There has only been sporadic efforts to include personality measures in geographic research, despite the strong connections between the two [55]. Recent works [44, 56] suggest that spatial behaviour can be partially explained from personality traits. However, in [56], this understanding is based on biased data collected from location-based social networks. In [44], the connection between spatial behaviour and personality is not investigated extensively, as it is not the main focus of the study.

2.3 Social and spatial connection

Recently, connections between the social and spatial behaviour of pairs [57–62] and groups [63] of individuals have been demonstrated, and used to design predictive models of mobility [58, 64, 65] or social ties [59, 66–68]. Shifting the attention to the individual level, recent works based on online social network data [69, 70], mobile phone calls data [62] and evenly sampled high resolution mobility trajectories [16] have shown correlations between the activity space size and the ego network structure, calling for further research to more closely examine the connections between social and spatial behaviour at the individual level.

3 Methods

3.1 Data description and pre-processing

Our study is based on 850 high resolution trajectories and call records of participants in a 24 months longitudinal experiment, the Copenhagen Networks Study (CNS) [71]. Results on the connections between social and spatial behaviour were corroborated with data from another experiment with fixed rate temporal sampling, but lower spatial resolution and sample size: the Lausanne Mobile Data Challenge (MDC) [72, 73], lasted for 19 months (see Table 1).

3.1.1 CNS dataset

The Copenhagen Networks Study (CNS) experiment took place between September 2013 and September 2015 [71] and involved ∼1000 Technical University of Denmark students (∼22% female, ∼78% male) typically aged between 19 and 21 years old. Participants' position over time was estimated combining their smart-phones WiFi and GPS data using the method described in [16, 74]. The location estimation error is below 50 meters in 95% of the cases. Participants' calls and SMS activity was also collected as part of the experiment. Individuals' background information were obtained through a 310 questions survey including the Big Five Inventory [75], which measures how individuals score on five broad

Table 1 Characteristics of the mobility datasets considered. N is the number of individuals, δt the temporal resolution, T the duration of data collection, δx the spatial resolution, TC the median weekly time coverage, defined as the fraction of time an individual's location is known

	N	δt	T	δx	TC
CNS	850	16 s	24 months	10 m	0.84
MDC	185	60 s	19 months	100–200 m	0.73

domains of human personality traits: openness, conscientiousness, extraversion, agreeableness, neuroticism. The personality questionnaire used in the study is a version of the Big Five Inventory [75], translated from English into Danish. It contains 44 individual items and each trait is computed as the average of 7–10 items. Data collection was approved by the Danish Data Protection Agency. All participants provided individual informed consent. Mobility patterns of participants in the CNS experiment display statistical properties consistent with previous literature [13], as shown in [16].

3.1.2 MDC dataset

Data was collected by the Lausanne Data Collection Campaign between October 2009 and March 2011. The campaign involved an heterogeneous sample of ∼185 volunteers with mixed backgrounds from the Lake Geneva region (Switzerland), who were allocated smart-phones [73]. In this work we used GSM data, that has the highest temporal sampling. Following Nokia's privacy policy, individuals participating in the study provided informed consent [73]. The Lausanne Mobile Data Challenge experiment involves 62% male and 38% female participants, where the age range 22–33 year-old accounts for roughly 2/3 of the population [76].

3.2 Metrics

In this section, we define the concepts and metrics used to quantify the social and spatial behaviour of an individual i.

Exploration behaviour is characterised by the following quantities:

Number of new locations/week: $n_{\text{loc}}(i,t)$ is the number of locations discovered by i in the week preceding t.

Number of new ties/week: $n_{\text{tie}}(i,t)$ is the number of individuals who had contact with i (by SMS or call) for the first time in the week preceding t.

Note that locations/ties are considered 'new' only if discovered after 20 weeks from the beginning of data collection.

Exploitation behaviour can be quantified by considering:

Activity space: The set $\text{AS}(i,t) = \{\ell_1, \ell_2, \ldots, \ell_j, \ldots, \ell_C\}$ of locations ℓ_j that individual i visited at least twice and where she spent a time τ_j larger than 200 min during a time-window of $T = 20$ weeks preceding time t (see Additional file 1 for the analysis with $T = 30$ weeks). Among the locations in the activity space, i visited ℓ_j with probability $p(\ell_j) = \tau_j / \sum \tau_j$. (It is worth noting that this time-based definition of activity space includes all significant locations independently of their spatial position and it is only loosely connected with space-oriented definitions widespread in the geography literature such as the "standard deviational ellipse" and the "road network buffer" [77].)

Social circle: The set $\text{SC}(i,t) = \{u_1, u_2, \ldots, u_j, \ldots, u_k\}$ of individuals u_j with whom individual i had a number of contacts $n_j > 5$ by SMS or call during a time-window of $T = 20$ consecutive weeks preceding time t (see Additional file 1 for the analysis with $T = 30$ weeks). The probability that i has contact with a given member u_j of her social circle is $p(u_j) = n_j / \sum n_j$.

For these two sets $\text{AS}(i,t)$ and $\text{SC}(i,t)$, we consider their sizes $C(i,t)$ and $k(i,t)$, quantifying the number of favoured locations and social ties, respectively; their entropies $H_{\text{AS}}(i,t)$ and $H_{\text{SC}}(i,t)$, measuring how time is allocated among locations and ties; their stabilities $J_{\text{AS}}(i,t)$ and $J_{\text{SC}}(i,t)$, quantifying the fraction of conserved locations and ties, respectively, across

Table 2 Definition of the metrics characterising the activity space and the social circle. (1) The size of a set is the number of elements in the set. (2) We compute the entropy of a set considering the probability $p(j)$ associated to each element j of the set. (3) We measure the stability J_{AS} by computing the Jaccard similarity between the activity space at t and at $t - T$, with $T = 20$ weeks. J_{SC} is computed in the same way for the social circle. (4) We compute the rank turnover of a set by measuring for each of its elements j the absolute change in rank between two consecutive time windows of length $T = 20$ weeks. The rank is attributed based on the probability $p(j)$. The average absolute change in rank across all elements corresponds to the rank turnover

	Activity space	Social circle
(1) Size	$C(i,t) = \lvert AS_i(t)\rvert$	$k(i,t) = \lvert SC(i,t)\rvert$
(2) Entropy	$H_{\mathrm{AS}}(i,t) = -\sum_{j=1}^{C(i,t)} p(j)\log p(j)$	$H_{\mathrm{SC}}(i,t) = -\sum_{j=1}^{k(i,t)} p(j)\log p(j)$
(3) Stability	$J_{\mathrm{SC}}(i,t) = \frac{\lvert SC(i,t)\cap SC(i,t-T)\rvert}{\lvert SC(i,t)\cup SC(i,t-T)\rvert}$ *	$J_{\mathrm{AS}}(i,t) = \frac{\lvert AS(i,t)\cap AS(i,t-T)\rvert}{\lvert AS(i,t)\cup AS(i,t-T)\rvert}$ *
(4) Rank turnover	$R_{\mathrm{AS}}(i,t) = \sum_{j=1}^{N} \frac{\lvert r(j,t) - r(j,t-T)\rvert}{N}$ †	$R_{\mathrm{SC}}(i,t) = \sum_{j=1}^{N} \frac{\lvert r(j,t) - r(j,t-T)\rvert}{N}$ †

*Here $T = 20$ weeks, see Additional file 1 for the analysis with $T = 30$ weeks.

† $r(\ell_k,t)$ and $r(u_k,t)$ denote the rank of a location ℓ_k and individual u_k at t, respectively.

consecutive non-overlapping windows of $T = 20$ weeks (see Additional file 1 for $T = 30$); their rank turnovers $R_{\mathrm{AS}}(i,t)$ and $R_{\mathrm{SC}}(i,t)$ measuring the average absolute change in rank of an element in the set between consecutive windows. The mathematical definition of these quantities is provided in Table 2

3.2.1 Other metrics

In order to compare the difference in entropy between two different sets, we compute their Jensen–Shannon divergence (JSD). The JSD between two sets P_1 and P_2 is computed as $\mathrm{JSD}(P_1,P_2) = H(\frac{1}{2}(P_1+P_2)) - \frac{1}{2}[H(P_1)+H(P_2)]$ (see also [4]).

4 Results

Both in their spatial and social behaviour, individuals are constantly balancing a trade-off between the exploitation of familiar options (such as returning to a favourite restaurant or spending time with an old friend) and the exploration of new opportunities (such as visiting a new bar or going on a first date) [78]. We adopt this exploration-exploitation perspective to analyse the relationship between social and spatial strategies in our dataset [16].

We quantify the propensity for exploration and exploitation within each individual, i, using the metrics reported in Table 3, Fig. 1 and described in Sect. 3. We focus on two aspects of exploitation, (i) *diversity*, characterising how individuals allocate time among their set of familiar locations and friends, and (ii) *evolution*, characterising the tendency to change exploited locations and friends over time.

Exploration and exploitation are persistent in time. First, we verify that individual behaviour is persistent in time. For all the aforementioned measures, we compare the individual self-variation across time $d_{\mathrm{self}}(i)$ with a reference difference $d_{\mathrm{ref}}(i,j)$ between individuals i and j. In the case of the activity space size, for example, self-variation is measured as $d_{\mathrm{self}} = \langle\lvert C(i,t) - C(i,t-T)\rvert\rangle$, where $\langle\cdot\rangle$ is the average across time and $T = 20$ weeks (see Additional file 1 for $T = 30$); the reference difference is computed as $d_{\mathrm{ref}}(i,j) = \lvert\langle C(i,t)\rangle - \langle C(j,t)\rangle\rvert$. If $d_{\mathrm{self}}(i) < d_{\mathrm{ref}}(i,j)$ for most j, we can conclude that for individual i, fluctuations of the activity space size are negligible compared to the difference with other individuals. The same procedure is followed for all metrics with an adjustment in the case

Table 3 Metrics characterising social and spatial behaviour. The metrics are defined in Sect. 3

	Exploration	Exploitation: Diversity	Exploitation: Evolution
Spatial	New locations/week, n_{loc}	Activity space size, C Activity space entropy, H_{AS}	Activity space stability, J_{AS} Activity space rank turnover, R_{AS}
Social	New ties/week, n_{tie}	Social circle size, k Social circle entropy, H_{SC}	Social circle stability, J_{SC} Social circle rank turnover, R_{SC}

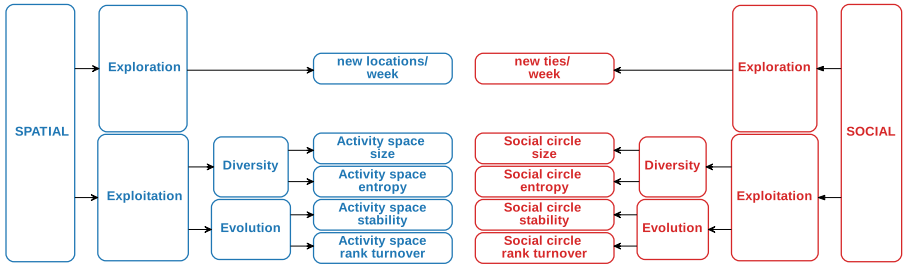

Figure 1 Schematic description of our framework

Table 4 CNS dataset: Persistence of social and spatial behaviour. For each of the social and spatial metrics, $\overline{d_{\text{self}}}$ is the average self-distance and $\overline{d_{\text{ref}}}$ is the reference distance between an individual and all others, averaged across individuals. The third column reports the fraction of cases where $\overline{d_{\text{self}}(i)} < \overline{d_{\text{ref}}(i,j)}$, averaged across the population

	$\overline{d_{\text{self}}}$	$\overline{d_{\text{ref}}}$	$\overline{d_{\text{self}}(i) < d_{\text{ref}}(i,j)}$
Social circle size, k	0.04 ± 0.09	12 ± 5	99%
Activity space size, C	0.04 ± 0.07	7 ± 3	99%
New ties/week, n_{tie}	0.05 ± 0.10	0.9 ± 0.5	96%
New locations/week, n_{loc}	0.10 ± 0.17	1 ± 1	95%
Social circle entropy, H_{SC}	0.002 ± 0.007	0.7 ± 0.2	99%
Activity space entropy, H_{AS}	0.002 ± 0.005	0.4 ± 0.1	99%
Social circle stability, J_{SC}	$(9 \pm 22) \cdot 10^{-4}$	0.13 ± 0.05	100%
Activity space stability, J_{AS}	$(9 \pm 26) \cdot 10^{-4}$	0.10 ± 0.04	99%
Social circle rank turnover, R_{SC}	0.05 ± 0.39	2 ± 1	99%
Activity space rank turnover, R_{AS}	0.04 ± 0.10	2 ± 1	99%

of entropies: The persistence of the entropy H_{AS} is verified by comparing the Jansen–Shannon divergences $d_{\text{self}} = \text{JSD}(\text{AS}(i,t), \text{AS}(i,t-T))$ and $d_{\text{ref}} = \text{JSD}(\text{AS}(i,t), \text{AS}(j,t))$. The same method was used for H_{SC} (see Sect. 3 and [4]).

Results from the CNS dataset reported in Table 4 show that for all metrics $d_{\text{self}}(i) < d_{\text{ref}}(i,j)$ holds in more than 99% of cases on average (MDC: 97%, see Additional file 1, Table S1). Moreover, the average self-variation across the population $\overline{d_{\text{self}}}$ is consistent with $\overline{d_{\text{self}}} = 0$ within errors, and $\overline{d_{\text{self}}}$ significantly smaller than the average reference difference $\overline{d_{\text{ref}}}$ (see Tables 4 and S1 in Additional file 1).

These results extend previous findings [4, 16] and suggest that each individual is characterised by a distinctive socio-spatial behaviour captured by the ensemble of these metrics averaged across time. In fact, these averages are heterogeneously distributed across the samples considered (see Fig. 2).

Exploration and exploitation are correlated in the social and spatial domain. A natural way to test the interdependency between social and spatial behaviours is measuring the correlation between a given social metric and a corresponding spatial one. We find positive

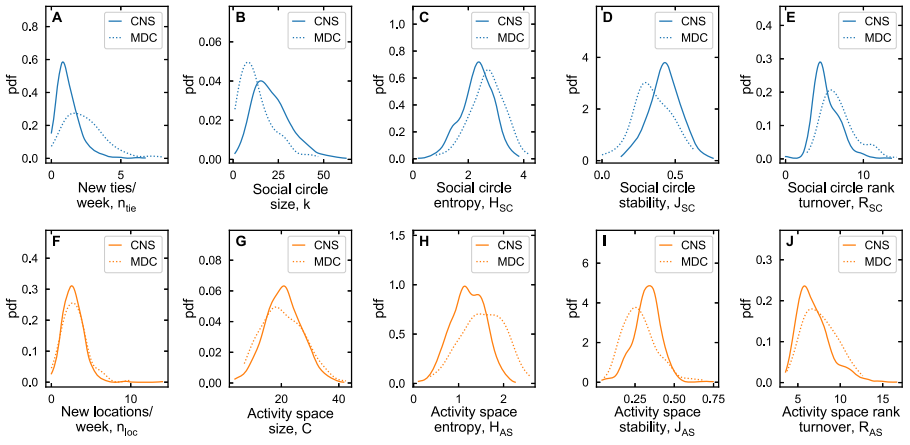

Figure 2 Distribution of social (above line) and spatial (bottom line) metrics for the CNS and MDC datasets

and significant correlations for all metrics and datasets (see Figs. 3 and S1 in Additional file 1).

We find that individuals with high propensity to explore new locations are also more keen on exploring social opportunities (see Fig. 3(A)). Those with diverse mobility routine are also likely to have a correspondingly large social circle (see Fig. 3(B)), and those that often replace social ties, have also an unstable set of favourite locations (see Fig. 3(C) and (D)).

We verify that the observed correlations are not spurious by performing multiple regression analyses that control for other possible sources of variation: gender, age, and time coverage (the average time an individual position is known). We implement five multiple linear regression models M1, M2, M3, M4 and M5. Each regression model predicts a given spatial metric (the activity space size C, the activity space entropy H_{AS}, the number of new locations/week n_{loc}, the activity space stability J_{AS} and the rank turnover R_{AS}) using the corresponding social metric and the control variables (age, gender and time coverage) as regressors. The relative importance of each regressor is assessed using the Lindeman, Merenda and Gold (LMG) [79] method.

Results obtained via weighted least square regression (see Table 5 for the CNS dataset and Table S2 in Additional file 1 for the MDC dataset) reveal that the social metrics are significant predictors for spatial metrics (p-value < 0.01 in all cases except for M4 in the MDC dataset), and they typically have more importance than factors such as gender, time, coverage and age group (see Fig. 4).

Among the control variables, gender is a significant predictor of spatial behaviour in the CNS dataset: Females display higher level of routine diversity and propensity towards exploration, in accordance with [80]. Time coverage, measuring the fraction of time an individual position is known, plays a significant role in explaining spatial entropy and activity space stability, since individuals who spend long time in the same place (or leave their phone in the same place) are more easily geo-localised. Age differences are not present within the sample of students participating in the CNS study, and they are not estimated to be relevant with respect to spatial behaviour in the MDC study.

We do not identify distinct classes of individuals. A natural question is whether or not, in the samples considered, there is evidence for distinct classes of individuals based on their

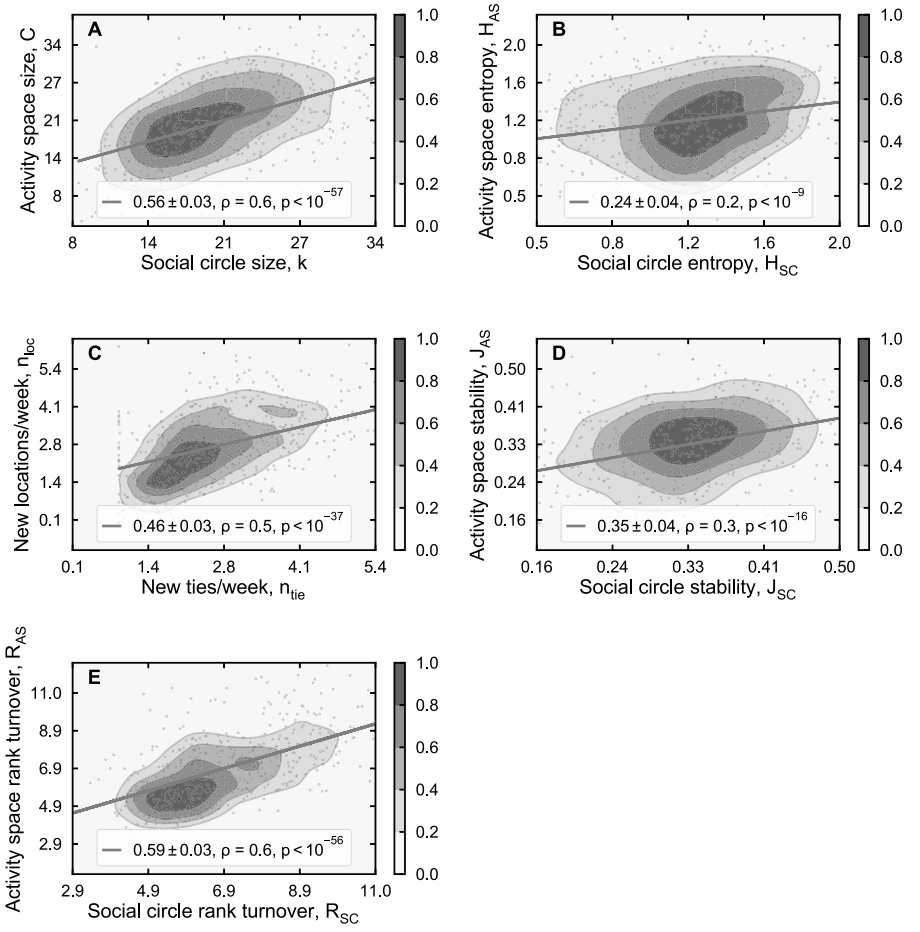

Figure 3 CNS dataset: correlation between the four dimensions of social and spatial behaviour. (**A**) Activity space vs social circle size. (**B**) Activity space vs social circle composition measured as their entropy. (**C**) Average number of new locations vs new ties per week. (**D**) Stability of the activity space vs the stability of the social circle measured as the Jaccard similarity between their composition in consecutive time-windows. (**E**) Rank turnover of the activity space vs the rank turnover of the social circle. Coloured filled areas correspond to cumulative probabilities estimated via Gaussian Kernel Density estimations. Grey lines correspond to linear fit with angular coefficient *b* reported in the legend. The Pearson correlation coefficient, with corresponding p-value, is reported in the legend

socio-spatial behaviour [5, 14]. We approach this problem by reducing the set of metrics to a smaller number of uncorrelated variables by applying Principal Component Analysis [81, 82]. The principal components represent the data through linear combinations of the original variables: In Table 6 we report the percentage of variance in the data explained by all components; in Table 7 we report the coefficients w describing how the original variables are linearly combined to obtain the first two principal components.

In both datasets, we find that the first principal component (PC 0) explains ∼40% of the differences between individuals (see Table 6). For the CNS dataset, the variables contributing the most to PC 0 (e.g. such that $w^2 > 0.1$) are, in order, the activity space size C, the social circle size k, the number of new locations/week n_{loc}, the activity space entropy H_{AS} and the number of new ties/week n_{tie}. n_{loc} and n_{tie} characterise the attitude towards exploration. The other metrics (C, k and H_{AS}) are related to routine diversity, or the tendency to dispose of a large set of familiar locations and friends. Since the sign of w is the

Table 5 Linear regression models for the CNS dataset. For each model, we show the coefficients (coeff) calculated by the regression model, the probability (p val) that the variable is not relevant, and the relative importance (LMG) of each regressor computed using the Lindeman, Merenda and Gold method [79]. Gender is a binary variable taking value 1 for females and 2 for males. For this dataset, age is not relevant as all participants have similar age. For each model, we report the R^2 goodness of fit, the F-test statistics with the corresponding p-value p_F

	coeff	p val	LMG
Model M1: Activity space size, C			
Social circle size, k	4 ± 0	$< 10^{-50}$	0.94
Gender	-0.4 ± 0.2	0.05	0.05
Time coverage	0.4 ± 0.2	0.06	0.01
[$R^2 = 0.32$, $F = 100.44$, $p_F = 0.0$]			
Model M2: Activity space entropy, H_{AS}			
Social circle entropy, H_{SC}	0.07 ± 0.01	$< 10^{-6}$	0.42
Gender	-0.06 ± 0.01	$< 10^{-4}$	0.22
Time coverage	-0.07 ± 0.01	$< 10^{-5}$	0.36
[$R^2 = 0.11$, $F = 27.30$, $p_F = 0.0$]			
Model M3: New locations/week, n_{loc}			
New ties/week, n_{tie}	0.60 ± 0.05	$< 10^{-32}$	0.9
Gender	-0.16 ± 0.05	$< 10^{-3}$	0.08
Time coverage	0.001 ± 0.047	1.0	0.01
[$R^2 = 0.22$, $F = 61.99$, $p_F = 0.0$]			
Model M4: Activity space stability, J_{AS}			
Social circle stability, J_{SC}	0.024 ± 0.004	$< 10^{-10}$	0.6
Gender	0.007 ± 0.003	0.05	0.04
Time coverage	0.017 ± 0.004	$< 10^{-5}$	0.36
[$R^2 = 0.16$, $F = 33.36$, $p_F = 0.0$]			
Model M5: Activity space rank turnover, R_{AS}			
Social circle rank turnover, R_{SC}	1 ± 0	$< 10^{-56}$	0.98
Gender	0.12 ± 0.07	0.06	0.01
Time coverage	-0.12 ± 0.07	0.07	0.01
[$R^2 = 0.36$, $F = 108.31$, $p_F = 0.0$]			

same for all the metrics above, we can conclude that individuals with higher propensity towards exploration tend to have a more diverse social and spatial routine, and vice-versa. Similar conclusions could be drawn by looking at results obtained for the MDC dataset.

The second principal component (PC 1) accounts for ∼15% of the total variation (see Table 6). It is dominated by the social circle stability J_{SC} (CNS: $w^2 = 0.21$, MDC: $w^2 = 0.52$) and the activity space stability J_{AS} (CNS: $w^2 = 0.24$, MDC: $w^2 = 0.26$) for both datasets (see Table 7). The sign of the coefficients w for J_{SC} and J_{AS} are the same, further confirming that these two metrics are correlated (see also Fig. 3). We can conclude that the second principal component accounts for the effects of routine evolution, or the tendency to change familiar locations and friends over long time scales. We consider the first two principal components, PC 0 and PC 1, to reduce the effects of noise and we test the hypothesis that there exists different classes of individuals applying the gap statistic method [83]. We apply it by looking at the gap between the within-cluster dispersion expected under a uniform distribution of the data and the dispersion obtained after applying K-means. For all possible choices of $K > 1$, we find that the gap is not large enough to support the existence of more than one class of individuals.

The big-five personality traits partly explain spatial and social behaviour. We verify if the differences between individuals can be explained by the Big five personality traits model [75], typically used to describe social and emotional life (see Table 8). We build two multiple linear regression models that use the Big five personality traits as regressors and one of

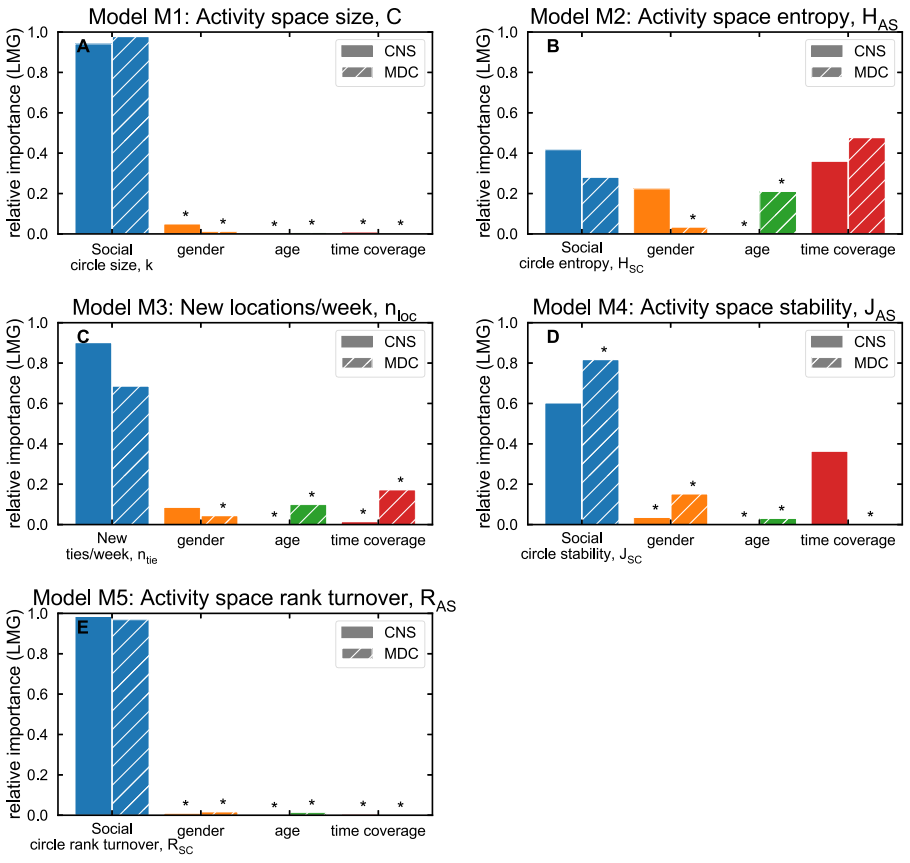

Figure 4 Relative importance of regressors. LMG of each regressor computed using the Lindeman, Merenda and Gold method [79] for models M1 (**A**), M2 (**B**), M3 (**C**), M4 (**D**) and M5 (**E**). Plain bars show results for the CNS dataset, dashed bars for the MDC dataset. Variables that are not significant in the regression model are marked with *

Table 6 Variance explained by principal components. The fraction of variance explained by each principal component for the CNS and MDC dataset

	PC 0	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9
CNS	0.39	0.17	0.12	0.08	0.07	0.06	0.04	0.03	0.03	0.01
MDC	0.43	0.14	0.13	0.08	0.07	0.06	0.04	0.03	0.02	0.01

the principal components describing socio-spatial behaviour as target. Results, shown in Table 9, show that three personality traits, neuroticism, openness and extraversion, are relevant predictors for socio-spatial behaviour. In particular, extraversion is the most important predictor of the first principal component: it positively correlates with the tendency to diversify routine and to explore opportunities. Neuroticism and openness explain instead the second principal component, since it correlates with the tendency to change routine over time (see also Fig. 5).

Finally, we perform all analyses considering only spatial metrics. Results are in line with those obtained considering all metrics: The first two principal components account for a large fraction of the variability in the data (see Table 10); The first component is dominated by the activity space size C, the number of new locations/week n_{loc} and the activity space entropy H_{AS}, while the second is mostly controlled by the activity space stability J_{AS}

Table 7 Contribution of the original variables to the first principal component. The scalar projection w of the first (PC 0) and second (PC 1) principal components, along the axis defined by each of the original variables. Each principal component has unit norm, hence the sum of w^2 is 1. Results are shown for the CNS and MDC datasets

	CNS		MDC	
	w (PC 0)	w (PC 1)	w (PC 0)	w (PC 1)
Social circle size, k	0.41	0.16	0.37	–0.15
Activity space size, C	0.42	–0.24	0.42	–0.08
New ties/week, n_{tie}	0.33	0.28	0.27	0.33
New locations/week, n_{loc}	0.38	–0.05	0.37	0.19
Social circle entropy, H_{SC}	0.31	0.30	0.34	0.09
Activity space entropy, H_{AS}	0.38	–0.16	0.30	–0.07
Social circle stability, J_{SC}	–0.16	–0.46	0.07	–0.72
Activity space stability, J_{AS}	–0.10	–0.49	–0.12	–0.51
Social circle rank turnover, R_{SC}	–0.20	0.28	–0.33	0.10
Activity space rank turnover, R_{AS}	–0.30	0.44	–0.38	0.17

Table 8 The Big-five traits and examples of adjectives describing them [84]

Trait	Related Adjectives
Extraversion	Active, Assertive, Energetic, Enthusiastic, Outgoing, Talkative
Agreeableness	Appreciative, Forgiving, Generous, Kind, Sympathetic
Conscientiousness	Efficient, Organised, Planful, Reliable, Responsible, Thorough
Neuroticism	Anxious, Self-pitying, Tense, Touchy, Unstable, Worrying
Openness to experience	Artistic, Curious, Imaginative, Insightful, Original, Wide Interests

Table 9 Extraversion, openness, and neuroticism explain socio-spatial behaviour. The result of a multiple linear regression explaining principal components of socio-spatial data (see Table 7). The value of each coefficient (coeff) is reported together with the probability (p val) that the coefficient is not relevant for the model. The relative importance of each coefficient (LMG) is computed using the LMG method [79]

	PC 0 $R^2 = 0.17, F = 21.40, p_F = 0.0$			PC 1 $R^2 = 0.03, F = 3.64, p_F = 0.0$		
	coeff	p val	LMG	coeff	p val	LMG
Extraversion	0.85 ± 0.09	$< 10^{-19}$	0.85	0.12 ± 0.06	0.05	0.14
Openness	-0.17 ± 0.08	0.03	0.02	0.13 ± 0.06	0.02	0.33
Neuroticism	0.25 ± 0.09	0.004	0.04	0.15 ± 0.06	0.02	0.3
Agreeableness	0.11 ± 0.08	0.2	0.04	-0.07 ± 0.06	0.2	0.12
Conscientiousness	0.06 ± 0.08	0.4	0.04	-0.07 ± 0.06	0.2	0.11

(Table 11). For the CNS dataset, extraversion is the most important predictor of the first principal component, while openness, extraversion and neuroticism account for the second component (see Table 12 and Fig. 6). The result presented above hold when choosing a time-window with length $T = 30$ weeks (see Additional file 1, Sect. S2).

5 Discussion

Using high resolution data from two large scale studies, we have investigated the connection between social and spatial behaviour for the first time. We have shown that, in both domains, individuals balance the trade-off between exploring new opportunities and exploiting known options in a distinctive and persistent manner. We have found that, to a significant extent, individuals adopt a similar strategy in the social and spatial sphere. These strategies are heterogeneous across the two samples considered, and there is no evidence suggesting that there exist distinct classes of individuals. Finally, we have shown that the

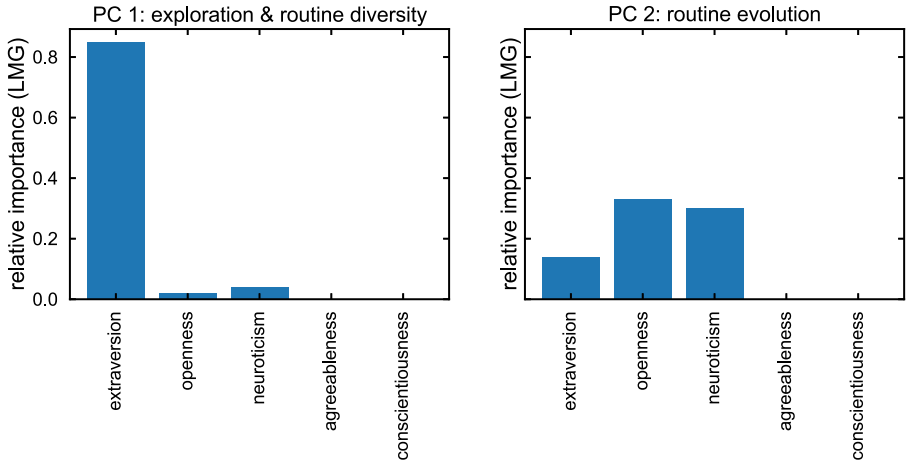

Figure 5 Relative importance of personality traits for socio-spatial behaviour. LMG of each regressor computed using the Lindeman, Merenda and Gold method [79] for the multiple regression model of the principal components (see also Table 9)

Table 10 Variance explained by principal components (only spatial data). The fraction of variance explained by each principal component for the CNS and MDC dataset

	PC 0	PC 1	PC 2	PC 3	PC 4
CNS	0.53	0.21	0.13	0.10	0.04
MDC	0.56	0.19	0.13	0.07	0.04

Table 11 Principal components (only spatial data). The weight of each metric in the first two principal components, for both datasets

	CNS		MDC	
	PC 0	PC 1	PC 0	PC 1
Activity space size, C	–0.58	0.02	0.55	0.12
New locations/week, $n_{\mathbf{loc}}$	–0.48	–0.19	0.51	–0.09
Activity space entropy, $H_{\mathbf{AS}}$	–0.50	–0.08	0.43	0.20
Activity space stability, $J_{\mathbf{AS}}$	–0.02	0.94	–0.19	0.95
Activity space rank turnover, $R_{\mathbf{AS}}$	0.43	–0.25	–0.47	–0.16

big five personality traits explain related aspects of both social and spatial behaviour. In particular, we have found that extraverted individuals are more explorative and have diverse routines in both the social and the spatial sphere while neuroticism and openness associate with high level of routine instability in the social and spatial domain.

Our findings confirm the usefulness of mobile phone data to study the connections between behaviour and personality [29, 40, 44, 85–87]. The results are in line with previous findings on the relation between personality and social behaviour: extraversion correlates with ego-network size [18, 41, 43] and diverse composition [88], openness to experience to social network turnover [29] and neuroticism does not correlate with social network size [11]. Furthermore, our findings establish a relation between personality and spatial behaviour, validating the theories suggesting that spatial choices are partially dictated by personality dispositions [15] and that a single set of personality traits underlies many aspects of a person's behaviour [2, 3].

Table 12 Extraversion, openness, and neuroticism explain spatial behaviour. The result of a multiple linear regression explaining principal components of spatial data (see Table 7). The value of each coefficient (coeff) is reported together with the probability (p val) that the coefficient is not relevant for the model. The relative importance of each coefficient (LMG) is computed using the LMG method [79]

	PC 0 $R^2 = 0.10, F = 12.83, p_F = 0.0$			PC 1 $R^2 = 0.03, F = 3.50, p_F = 0.0$		
	coeff	p val	LMG	coeff	p val	LMG
Extraversion	-0.50 ± 0.07	$< 10^{-10}$	0.77	-0.11 ± 0.05	0.02	0.3
Openness	0.19 ± 0.07	0.004	0.08	-0.11 ± 0.04	0.009	0.45
Neuroticism	-0.07 ± 0.07	0.4	0.03	-0.10 ± 0.05	0.03	0.21
Agreeableness	-0.10 ± 0.07	0.2	0.07	0.01 ± 0.05	0.8	0.01
Conscientiousness	-0.05 ± 0.07	0.5	0.05	0.03 ± 0.04	0.4	0.03

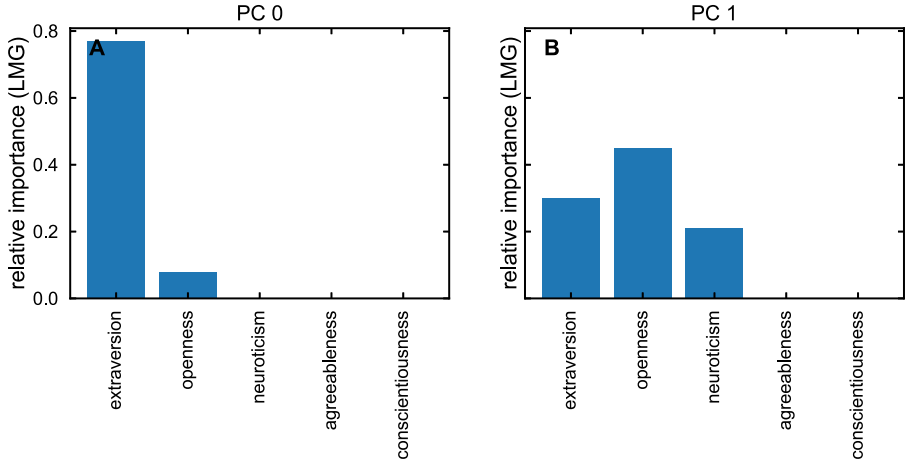

Figure 6 Relative importance of personality traits for spatial behaviour. LMG of each regressor computed using the Lindeman, Merenda and Gold method [79] for the multiple regression model of the principal components (see also Table 12)

Our findings on the connection between spatial behaviour and personality are consistent with the existing literature on personality. The correlation between exploration and extraversion could be explained by the fact that extraverted individuals are more likely to be risk-takers in various domains of life [89]. Extraverted individuals are also generally more likely to engage in social activities [90], which could partially explain why they allocate time among a larger set of locations. Furthermore, the key finding that individuals who score high in neuroticism and openness display a tendency to change familiar locations, and friends, over time fits well within the existing picture. In the case of neuroticism, it is well known that this trait is closely related with 'stability' [91], such that the trait of neuroticism is sometimes referred to as (low) 'emotional stability' [92]. Also, at the core of neuroticism is the tendency to experience negative emotions [93] including dissatisfaction [94], which in turn can lead into desire for change [95, 96]. Finally, it is known that people scoring high in neuroticism have a larger number of weak ties [42] and perceive that they tend to have less social support [97, 98], in line with our observation that they dispose of an unstable ego-network. Openness to experience has been shown to correlate with 'disloyal' behaviour also in other contexts such as politics [99] and shopping [100]. Our results, in agreement with previous studies on social [29, 40, 44, 85–87] and online

[101–104] behaviour, show that personality traits explain only partially how individuals behave in specific situations [105].

As a final point, we emphasize that the individual characterisation of spatial behaviour and connections with personality are fundamental to develop conceptual [55] and predictive [106] models of travel behaviour accounting for individual-level differences.

Additional material

Additional file 1: Supplementary information. (PDF 455 kB)

Acknowledgements
Portions of the research in this paper used the MDC Database made available by Idiap Research Institute, Switzerland and owned by Nokia.

Funding
This work was supported by the Danish Council for Independent Research ("Microdynamics of influence in social systems", grant id. 4184-00556, SL is PI).

List of abbreviations
CNS, Copenhagen Networks Study; MDC, (Lausanne) Mobile Data Challenge; GSM, Global System for Mobile Communications; JSD, Jensen–Shannon divergence; LMG, Lindeman, Merenda and Gold.

Availability of data and materials
CNS dataset: Data from the Copenhagen Networks study are not publicly available due to privacy considerations including European Union regulations and Danish Data Protection Agency rules. Due to the data security of participants, data cannot be shared freely, but are available to researchers who meet the criteria for access to confidential data, sign a confidentiality agreement, and agree to work under supervision in Copenhagen. Please direct your queries to Sune Lehmann, the Principal Investigator of the study, at sljo@dtu.dk. *MDC dataset*: The Lausanne Mobile Data Challenge data are available from Idiap Research Institute but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from Idiap Research Institute to eligible institutions upon reasonable request (https://www.idiap.ch/dataset/mdc/download).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
LA, SL and AB designed the study; LA performed the analysis; LA, SL and AB analysed the results; LA, SL and AB wrote the paper. All authors read and approved the final manuscript.

Author details
[1]Technical University of Denmark, Lyngby, Denmark. [2]Niels Bohr Institute, University of Copenhagen, København Ø, Denmark. [3]Department of Mathematics, City, University of London, London, UK.

Publisher's Note

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

Received: 8 January 2018 Accepted: 5 September 2018 Published online: 18 September 2018

References

1. Hägerstraand T (1970) What about people in regional science? Pap Reg Sci 24(1):7–24
2. Mischel W (1973) Toward a cognitive social learning reconceptualization of personality. Psychol Rev 80(4):252–283
3. Allport GW (1937) Personality: a psychological interpretation. Henry Holt And Company, Inc., New York
4. Saramäki J, Leicht EA, López E, Roberts SGB, Reed-Tsochas F, Dunbar RIM (2014) Persistence of social signatures in human communication. Proc Natl Acad Sci USA 111(3):942–947
5. Miritello G, Lara R, Cebrian M, Moro E (2013) Limited communication capacity unveils strategies for human interaction. Sci Rep 3:1950
6. Dunbar RIM (1993) Coevolution of neocortical size, group size and language in humans. Behav Brain Sci 16(4):681–694
7. Roberts SGB, Dunbar RIM, Pollet TV, Kuppens T (2009) Exploring variation in active network size: constraints and ego characteristics. Soc Netw 31(2):138–146
8. Igarashi T, Takai J, Yoshida T (2005) Gender differences in social network development via mobile phone text messages: a longitudinal study. J Soc Pers Relatsh 22(5):691–713
9. Wrzus C, Hänel M, Wagner J, Neyer FJ (2013) Social network changes and life events across the life span: a meta-analysis. Psychol Bull 139(1):53–80

10. McPherson M, Smith-Lovin L, Brashears ME (2006) Social isolation in America: changes in core discussion networks over two decades. Am Sociol Rev 71(3):353–375
11. Roberts SGB, Wilson R, Fedurek P, Dunbar RIM (2008) Individual differences and personal social network size and structure. Pers Individ Differ 44(4):954–964
12. Perchoux C, Chaix B, Cummins S, Kestens Y (2013) Conceptualization and measurement of environmental exposure in epidemiology: accounting for activity space related to daily mobility. Health Place 21:86–93
13. Gonzalez MC, Hidalgo CA, Barabasi A-L (2008) Understanding individual human mobility patterns. Nature 453(7196):779–782
14. Pappalardo L, Simini F, Rinzivillo S, Pedreschi D, Giannotti F, Barabási A-L (2015) Returners and explorers dichotomy in human mobility. Nat Commun 6:8166
15. Aitken SC (1991) Person-environment theories in contemporary perceptual and behavioural geography I: personality, attitudinal and spatial choice theories. Prog Hum Geogr 15(2):179–193
16. Alessandretti L, Sapiezynski P, Lehmann S, Baronchelli A (2016) Evidence for a conserved quantity in human mobility. arXiv:1609.03526
17. Lu Y-E, Roberts S, Lio P, Dunbar R, Crowcroft J (2009) Size matters: variation in personal network size, personality and effect on information transmission. In: CSE'09: international conference on computational science and engineering, vol 4, pp 188–193
18. Casciaro T (1998) Seeing things clearly: social structure, personality, and accuracy in social network perception. Soc Netw 20(4):331–351
19. Hojat M (1982) Loneliness as a function of selected personality variables. J Clin Psychol 38(1):137–141
20. Branje SJT, van Lieshout CFM, van Aken MAG (2004) Relations between big five personality characteristics and perceived support in adolescents' families. J Pers Soc Psychol 86(4):615–628
21. Amato PR (1990) Personality and social network involvement as predictors of helping behavior in everyday life. Soc Psychol Q 53:31–43
22. Hamburger YA, Ben-Artzi E (2000) The relationship between extraversion and neuroticism and the different uses of the Internet. Comput Hum Behav 16(4):441–449
23. Gonçalves B, Perra N, Vespignani A (2011) Modeling users' activity on Twitter networks: validation of Dunbar's number. PLoS ONE 6(8):e22656
24. Sutcliffe A, Dunbar R, Binder J, Arrow H (2012) Relationships and the social brain: integrating psychological and evolutionary perspectives. Br J Psychol 103(2):149–168
25. Arnaboldi V, Conti M, Passarella A, Pezzoni F (2012) Analysis of ego network structure in online social networks. In: 2012 international conference on privacy, security, risk and trust (PASSAT) and 2012 international conference on social computing (SocialCom), pp 31–40
26. Miritello G, Moro E, Lara R, Martínez-López R, Belchamber J, Roberts SG, Dunbar RIM (2013) Time as a limited resource: communication strategy in mobile phone networks. Soc Netw 35(1):89–95
27. Zhou W-X, Sornette D, Hill RA, Dunbar RIM (2005) Discrete hierarchical organization of social group sizes. Proc R Soc Lond B, Biol Sci 272(1561):439–444
28. Marlow C, Byron L, Lento T, Rosenn I (2009) Maintained relationships on Facebook
29. Centellegher S, López E, Saramäki J, Lepri B (2017) Personality traits and ego-network dynamics. PLoS ONE 12(3):e0173110
30. Wehrli S et al (2008) Personality on social network sites: an application of the five factor model. Working paper no. 7, ETH Sociology, Zurich
31. Kumar R, Novak J, Tomkins A (2010) Structure and evolution of online social networks. In: Link mining: models, algorithms, and applications. Springer, New York, pp 337–357
32. Newman ME (2001) Clustering and preferential attachment in growing networks. Phys Rev E 64(2):025102
33. Mislove A, Koppula HS, Gummadi KP, Druschel P, Bhattacharjee B (2008) Growth of the Flickr social network. In: Proceedings of the first workshop on online social networks. ACM, New York, pp 25–30
34. Dunbar RIM, Spoors M (1995) Social networks, support cliques, and kinship. Hum Nat 6(3):273–290
35. Carrasco J, Miller E, Wellman B (2008) How far and with whom do people socialize?: empirical evidence about distance between social network members. Transp Res Rec 2076:114–122
36. van den Berg P, Arentze T, Timmermans H (2009) Size and composition of ego-centered social networks and their effect on geographic distance and contact frequency. Transp Res Rec 2135:1–9
37. Campbell KE, Marsden PV, Hurlbert JS (1986) Social resources and socioeconomic status. Soc Netw 8(1):97–117
38. Reis HT, Nezlek J, Wheeler L (1980) Physical attractiveness in social interaction. J Pers Soc Psychol 38(4):604–617
39. Jaccard JJ (1974) Predicting social behavior from personality traits. J Res Pers 7(4):358–367
40. Staiano J, Lepri B, Aharony N, Pianesi F, Sebe N, Pentland A (2012) Friends don't lie: inferring personality traits from social network structure. In: Proceedings of the 2012 ACM conference on ubiquitous computing. ACM, New York, pp 321–330
41. Asendorpf JB, Van Aken MA (2003) Personality–relationship transaction in adolescence: core versus surface personality characteristics. J Pers 71(4):629–666
42. Kalish Y, Robins G (2006) Psychological predispositions and network structure: the relationship between individual predispositions, structural holes and network closure. Soc Netw 28(1):56–84
43. Pollet TV, Roberts SG, Dunbar RIM (2011) Extraverts have larger social network layers. J Individ Differ 32:161–169
44. de Montjoye Y-A, Quoidbach J, Robic F, Pentland A (2013) Predicting personality using novel mobile phone-based metrics. In: SBP. Springer, Berlin, pp 48–55
45. Selfhout M, Burk W, Branje S, Denissen J, Van Aken M, Meeus W (2010) Emerging late adolescent friendship networks and big five personality traits: a social network approach. J Pers 78(2):509–538
46. Asendorpf J, Denissen JJ (2006) Predictive validity of personality types versus personality dimensions from early childhood to adulthood: implications for the distinction between core and surface traits. Merrill-Palmer Q 52(3):486–513
47. Golledge RG (1997) Spatial behavior: a geographic perspective. Guilford Press, New York
48. Järv O, Ahas R, Witlox F (2014) Understanding monthly variability in human activity spaces: a twelve-month study using mobile phone call detail records. Transp Res, Part C, Emerg Technol 38:122–135

49. Schönfelder S, Axhausen KW (2010) Urban rhythms and travel behaviour: spatial and temporal phenomena of daily travel. Ashgate Publishing, Ltd., Farnham
50. Kwan M-P (2000) Gender differences in space-time constraints. Area 32(2):145–156
51. Vazquez-Prokopec GM, Bisanzio D, Stoddard ST, Paz-Soldan V, Morrison AC, Elder JP, Ramirez-Paredes J, Halsey ES, Kochel TJ, Scott TW et al (2013) Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment. PLoS ONE 8(4):e58802
52. Kang C, Gao S, Lin X, Xiao Y, Yuan Y, Liu Y, Ma X (2010) Analyzing and geo-visualizing individual human mobility patterns using mobile call records. In: 2010 18th international conference on geoinformatics, pp 1–7
53. Zenk SN, Schulz AJ, Matthews SA, Odoms-Young A, Wilbur J, Wegrzyn L, Gibbs K, Braunschweig C, Stokes C (2011) Activity space environment and dietary and physical activity behaviors: a pilot study. Health Place 17(5):1150–1161
54. Kwan M-P, Lee J (2004) Geovisualization of human activity patterns using 3D GIS: a time-geographic approach. In: Spatially integrated social science. Oxford University Press, Oxford
55. Van Acker V, Van Wee B, Witlox F (2010) When transport geography meets social psychology: toward a conceptual model of travel behaviour. Transp Rev 30(2):219–240
56. Chorley MJ, Whitaker RM, Allen SM (2015) Personality and location-based social networks. Comput Hum Behav 46:45–56
57. Backstrom L, Sun E, Marlow C (2010) Find me if you can: improving geographical prediction with social and spatial proximity. In: Proceedings of the 19th international conference on world wide web. ACM, New York, pp 61–70
58. McGee J, Caverlee J, Cheng Z (2013) Location prediction in social media based on tie strength. In: Proceedings of the 22nd ACM international conference on information & knowledge management. ACM, New York, pp 459–468
59. Sadilek A, Kautz H, Bigham JP (2012) Finding your friends and following them to where you are. In: Proceedings of the fifth ACM international conference on web search and data mining. ACM, New York, pp 723–732
60. Crandall DJ, Backstrom L, Cosley D, Suri S, Huttenlocher D, Kleinberg J (2010) Inferring social ties from geographic coincidences. Proc Natl Acad Sci USA 107(52):22436–22441
61. Grabowicz PA, Ramasco JJ, Gonçalves B, Eguíluz VM (2014) Entangling mobility and interactions in social media. PLoS ONE 9(3):e92196
62. Toole JL, Herrera-Yaqüe C, Schneider CM, González MC (2015) Coupling human mobility and social ties. J R Soc Interface 12(105):20141128
63. Onnela J-P, Arbesman S, González MC, Barabási A-L, Christakis NC (2011) Geographic constraints on social network groups. PLoS ONE 6(4):e16939
64. Jurgens D (2013) That's what friends are for: inferring location in online social media platforms based on social relationships. In: ICWSM-13, pp 273–282
65. Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: user movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, New York, pp 1082–1090
66. Wang D, Pedreschi D, Song C, Giannotti F, Barabasi A-L (2011) Human mobility, social ties, and link prediction. In: Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, New York, pp 1100–1108
67. Scellato S, Noulas A, Mascolo C (2011) Exploiting place features in link prediction on location-based social networks. In: Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, New York, pp 1046–1054
68. Pham H, Shahabi C, Liu Y (2013) EBM: an entropy-based model to infer social strength from spatiotemporal data. In: Proceedings of the 2013 ACM SIGMOD international conference on management of data. ACM, New York, pp 265–276
69. Cheng Z, Caverlee J, Lee K, Sui DZ (2011) Exploring millions of footprints in location sharing services. ICWSM 2011:81–88
70. Cranshaw J, Toch E, Hong J, Kittur A, Sadeh N (2010) Bridging the gap between physical location and online social networks. In: Proceedings of the 12th ACM international conference on ubiquitous computing. ACM, New York, pp 119–128
71. Stopczynski A, Sekara V, Sapiezynski P, Cuttone A, Madsen MM, Larsen JE, Lehmann S (2014) Measuring large-scale social networks with high resolution. PLoS ONE 9(4):e95978
72. Kiukkonen N, Blom J, Dousse O, Gatica-Perez D, Laurila J (2010) Towards rich mobile phone datasets: Lausanne data collection campaign. In: Proc. ICPS
73. Laurila JK, Gatica-Perez D, Aad I, Bornet O, Do T-M-T, Dousse O, Eberle J, Miettinen M et al (2012) The mobile data challenge: big data for mobile computing research. In: Pervasive computing, number EPFL-CONF-192489
74. Sapiezynski P, Gatej R, Mislove A, Lehmann S (2015) Opportunities and challenges in crowdsourced wardriving. In: Proceedings of the 2015 ACM conference on Internet measurement conference. ACM, New York, pp 267–273
75. John OP, Srivastava S (1999) The big five trait taxonomy: history, measurement, and theoretical perspectives. In: Handbook of personality: theory and research, vol 2, pp 102–138
76. Laurila JK, Gatica-Perez D, Aad I, Blom J, Bornet O, Do TMT, Dousse O, Eberle J, Miettinen M (2013) From big smartphone data to worldwide research: the mobile data challenge. Pervasive Mob Comput 9(6):752–771
77. Sherman JE, Spencer J, Preisser JS, Gesler WM, Arcury TA (2005) A suite of methods for representing activity space in a healthcare accessibility study. Int J Health Geogr 4(1):24
78. Hills TT, Todd PM, Lazer D, Redish AD, Couzin ID, Cognitive Search Research Group (2015) Exploration versus exploitation in space, mind, and society. Trends Cogn Sci 19(1):46–54
79. Grömping U et al (2006) Relative importance for linear regression in R: the package relaimpo. J Stat Softw 17:1
80. Mollgaard A, Lehmann S, Mathiesen J (2017) Correlations between human mobility and social interaction reveal general activity patterns. PLoS ONE 12(12):e0188973
81. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemom Intell Lab Syst 2(1–3):37–52
82. Mika S, Schölkopf B, Smola AJ, Müller K-R, Scholz M, Rätsch G (1999) Kernel PCA and de-noising in feature spaces. In: Advances in neural information processing systems, pp 536–542
83. Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a data set via the gap statistic. J R Stat Soc, Ser B, Stat Methodol 63(2):411–423

84. McCrae RR, John OP (1992) An introduction to the five-factor model and its applications. J Pers 60(2):175–215
85. Lambiotte R, Kosinski M (2014) Tracking the digital footprints of personality. Proc IEEE 102(12):1934–1939
86. Bogomolov A, Lepri B, Ferron M, Pianesi F, Pentland AS (2014) Daily stress recognition from mobile phone data, weather conditions and individual traits. In: Proceedings of the 22nd ACM international conference on multimedia. ACM, New York, pp 477–486
87. Chittaranjan G, Blom J, Gatica-Perez D (2013) Mining large-scale smartphone data for personality studies. Pers Ubiquitous Comput 17(3):433–450
88. Friggeri A, Lambiotte R, Kosinski M, Fleury E (2012) Psychological aspects of social communities. In: 2012 international conference on privacy, security, risk and trust (PASSAT) and 2012 international conference on social computing (SocialCom), pp 195–202
89. Nicholson N, Soane E, Fenton-O'Creevy M, Willman P (2005) Personality and domain-specific risk taking. J Risk Res 8(2):157–176
90. Lucas RE, Diener E, Grob A, Suh EM, Shao L (2000) Cross-cultural evidence for the fundamental features of extraversion. J Pers Soc Psychol 79(3):452–468
91. Robinson MD, Tamir M (2005) Neuroticism as mental noise: a relation between neuroticism and reaction time standard deviations. J Pers Soc Psychol 89(1):107–114
92. Judge TA, Bono JE (2001) Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: a meta-analysis. J Appl Psychol 86(1):80–92
93. Costa PT, McCrea RR (1992) Revised neo personality inventory (NEO PI-R) and neo five-factor inventory (NEO-FFI). Psychological Assessment Resources
94. Costa PT, McCrae RR (1980) Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. J Pers Soc Psychol 38(4):668–678
95. Zhou J, George JM (2001) When job dissatisfaction leads to creativity: encouraging the expression of voice. Acad Manag J 44(4):682–696
96. Rusbult CE, Zembrodt IM (1983) Responses to dissatisfaction in romantic involvements: a multidimensional scaling analysis. J Exp Soc Psychol 19(3):274–293
97. Daniel RW, Booth B, Reed D, Laughlin PR (1997) Personality, social networks, and perceived social support among alcoholics: a structural equation analysis. J Pers 65(3):649–692
98. Stokes JP (1985) The relation of social network and individual difference variables to loneliness. J Pers Soc Psychol 48(4):981–990
99. Bakker BN, Klemmensen R, Nørgaard AS, Schumacher G (2016) Stay loyal or exit the party? How openness to experience and extroversion explain vote switching. Polit Psychol 37(3):419–429
100. Matzler K, Bidmon S, Grabner-Kräuter S (2006) Individual determinants of brand affect: the role of the personality traits of extraversion and openness to experience. J Prod Brand Manag 15(7):427–434
101. Bachrach Y, Kosinski M, Graepel T, Kohli P, Stillwell D (2012) Personality and patterns of Facebook usage. In: Proceedings of the 4th annual ACM web science conference. ACM, New York, pp 24–32
102. Markovikj D, Gievska S, Kosinski M, Stillwell D (2013) Mining Facebook data for predictive personality modeling. In: Proceedings of the 7th international AAAI conference on weblogs and social media (ICWSM 2013), pp 23–26
103. Quercia D, Kosinski M, Stillwell D, Crowcroft J (2011) Our Twitter profiles, our selves: predicting personality with Twitter. In: 2011 IEEE third international conference on privacy, security, risk and trust (PASSAT) and 2011 IEEE third international conference on social computing (SocialCom), pp 180–185
104. Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci USA 110(15):5802–5805
105. Fleeson W, Noftle E (2008) The end of the person–situation debate: an emerging synthesis in the answer to the consistency question. Soc Pers Psychol Compass 2(4):1667–1684
106. Jiang S, Yang Y, Gupta S, Veneziano D, Athavale S, González MC (2016) The TimeGeo modeling framework for urban motility without travel surveys. Proc Natl Acad Sci USA 113:E5370–E5378

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com