

RESEARCH

# Supplementary material for Understanding the interplay between social and spatial behaviour

Laura Alessandretti, Sune Lehmann and Andrea Baronchelli

## 1 Results obtained with the MDC dataset

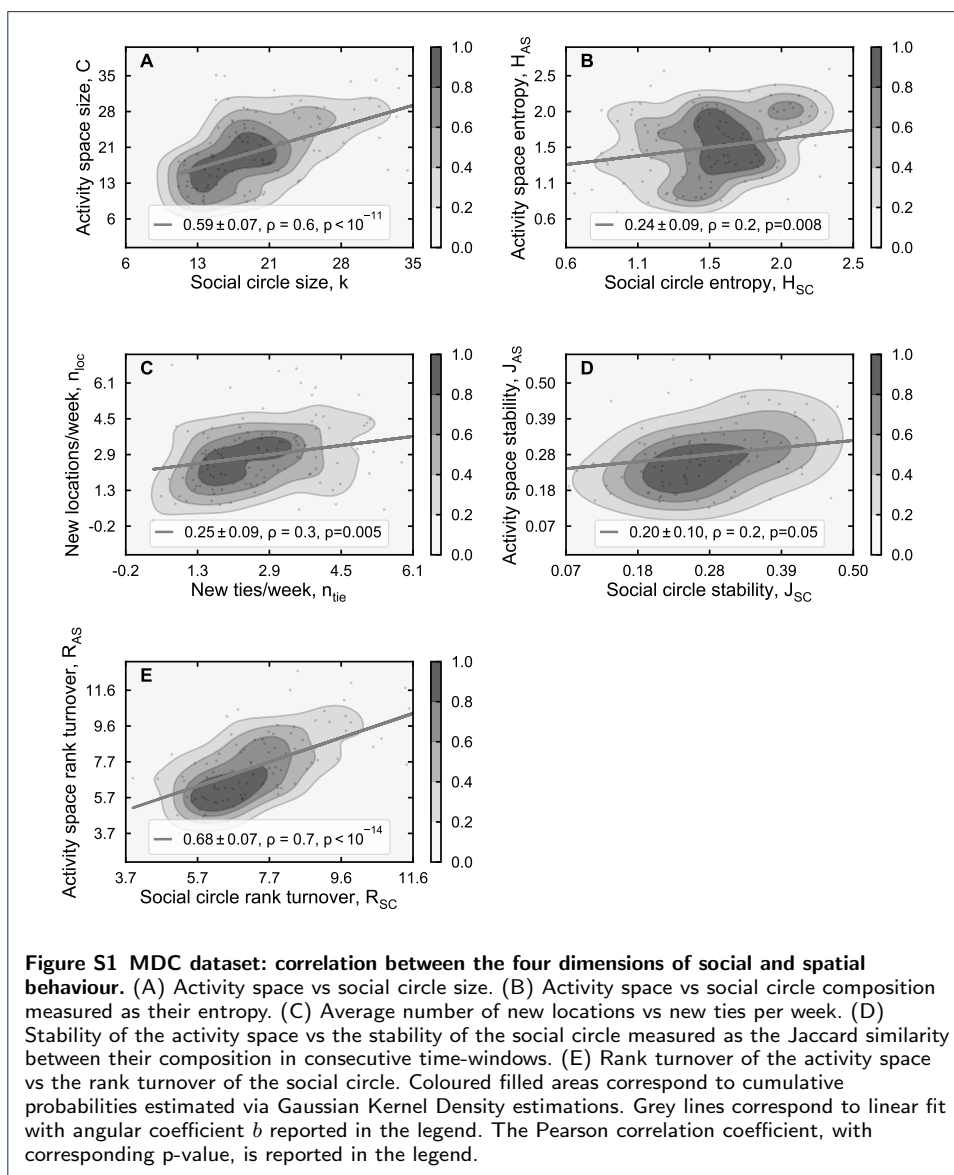
Tables S1, S2 and Fig. S1 report the results of the persistence analysis, the multiple regression analysis, and the correlation analysis for the MDC dataset.

|  | $\overline{d_{self}}$ | $\overline{d_{ref}}$ | $\overline{d_{self}(i) < d_{ref}(i, j)}$ |
|--|-----------------------|----------------------|--|
| Social circle size, $k$                | $0.05 \pm 0.13$       | $10 \pm 5$           | 97%                                      |
| Activity space size, $C$               | $0.07 \pm 0.12$       | $8 \pm 3$            | 97%                                      |
| New ties/week, $n_{tie}$               | $0.2 \pm 0.3$         | $2 \pm 1$            | 91%                                      |
| New locations/week, $n_{loc}$          | $0.2 \pm 0.6$         | $2 \pm 1$            | 90%                                      |
| Social circle entropy, $H_{SC}$        | $0.006 \pm 0.014$     | $0.7 \pm 0.3$        | 97%                                      |
| Activity space entropy, $H_{AS}$       | $0.004 \pm 0.008$     | $0.5 \pm 0.2$        | 97%                                      |
| Social circle stability, $J_{SC}$      | $0.002 \pm 0.005$     | $0.15 \pm 0.05$      | 99%                                      |
| Activity space stability, $J_{AS}$     | $0.002 \pm 0.004$     | $0.12 \pm 0.05$      | 99%                                      |
| Social circle rank turnover, $R_{SC}$  | $0.07 \pm 0.15$       | $2 \pm 1$            | 98%                                      |
| Activity space rank turnover, $R_{AS}$ | $0.2 \pm 0.6$         | $2 \pm 1$            | 97%                                      |

**Table S1 MDC dataset: Persistence of social and spatial behaviour.** For each of the social and spatial metrics,  $\overline{d_{self}}$  is the average self-distance and  $\overline{d_{ref}}$  is the reference distance between an individual and all others, averaged across individuals. The third column reports the fraction of cases where  $\overline{d_{self}(i) < d_{ref}(i, j)}$ , averaged across the population.

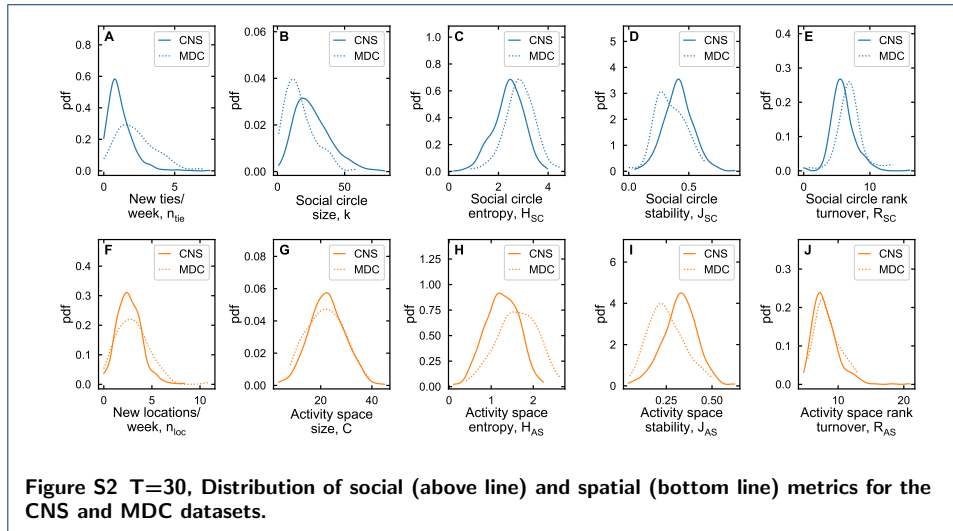
| <b>Model M1: Activity space size, <math>C</math></b>               | coeff                          | p val        | LMG  |
|--|--------------------------------|--------------|------|
| Social circle size, $k$  | $5 \pm 1$                      | $< 10^{-11}$ | 0.98 |
| gender   | $0.1 \pm 0.6$                  | 0.8          | 0.01 |
| age group  | $0.6 \pm 0.6$                  | 0.3          | 0.01 |
| time coverage  | $-0.4 \pm 0.6$                 | 0.4          | 0.0  |
| [ $R^2 = 0.40$ , $F = 16.80$ , $p_F = 0.0$ ]                       |                                |              |      |
| <b>Model M2: Activity space entropy, <math>H_{AS}</math></b>       |                                |              |      |
| Social circle entropy, $H_{SC}$                                    | $0.11 \pm 0.04$                | 0.009        | 0.28 |
| gender   | $0.04 \pm 0.04$                | 0.3          | 0.03 |
| age group  | $-0.08 \pm 0.04$               | 0.06         | 0.21 |
| time coverage  | $-0.14 \pm 0.04$               | 0.002        | 0.48 |
| [ $R^2 = 0.20$ , $F = 6.50$ , $p_F = 0.0$ ]                        |                                |              |      |
| <b>Model M3: New locations/week, <math>n_{loc}</math></b>          |                                |              |      |
| New ties/week, $n_{tie}$   | $0.5 \pm 0.1$                  | 0.002        | 0.69 |
| gender   | $0.01 \pm 0.15$                | 0.9          | 0.04 |
| age group  | $0.2 \pm 0.1$                  | 0.2          | 0.1  |
| time coverage  | $-0.3 \pm 0.1$                 | 0.06         | 0.17 |
| [ $R^2 = 0.13$ , $F = 3.78$ , $p_F = 0.0$ ]                        |                                |              |      |
| <b>Model M4: Activity space stability, <math>J_{AS}</math></b>     |                                |              |      |
| Social circle stability, $J_{SC}$                                  | $0.02 \pm 0.01$                | 0.1          | 0.82 |
| gender   | $-0.006 \pm 0.012$             | 0.6          | 0.15 |
| age group  | $-0.003 \pm 0.012$             | 0.8          | 0.03 |
| time coverage  | $(-10 \pm 1213) \cdot 10^{-5}$ | 1.0          | 0.0  |
| [ $R^2 = 0.04$ , $F = 0.80$ , $p_F = 0.5$ ]                        |                                |              |      |
| <b>Model M5: Activity space rank turnover, <math>R_{AS}</math></b> |                                |              |      |
| Social circle rank turnover, $R_{SC}$                              | $1 \pm 0$                      | $< 10^{-15}$ | 0.97 |
| gender   | $0.04 \pm 0.15$                | 0.8          | 0.02 |
| age group  | $-0.2 \pm 0.1$                 | 0.1          | 0.01 |
| time coverage  | $-0.06 \pm 0.15$               | 0.7          | 0.0  |
| [ $R^2 = 0.55$ , $F = 27.24$ , $p_F = 0.0$ ]                       |                                |              |      |

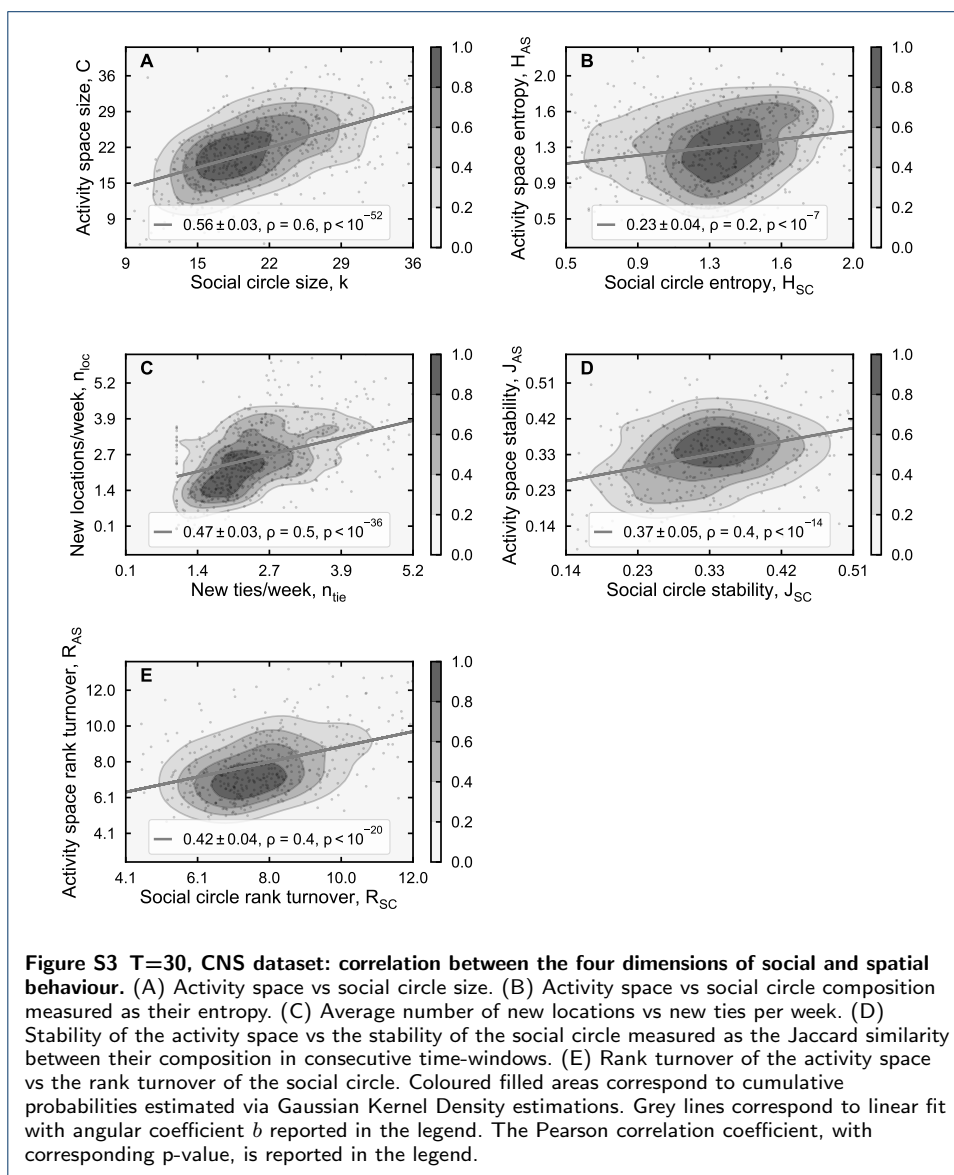
**Table S2 Linear regression models for the MDC dataset.** For each model, we report the  $R^2$  goodness of fit, the  $F$  - test statistics with the corresponding p-value  $p_F$ . We show the coefficients (coeff) calculated by the regression model, the probability (p val) that the variable is not relevant, and the relative importance (LMG) of each regressor computed using the Lindeman, Merenda and Gold method. Gender is a binary variable taking value 1 for females and 2 for males.

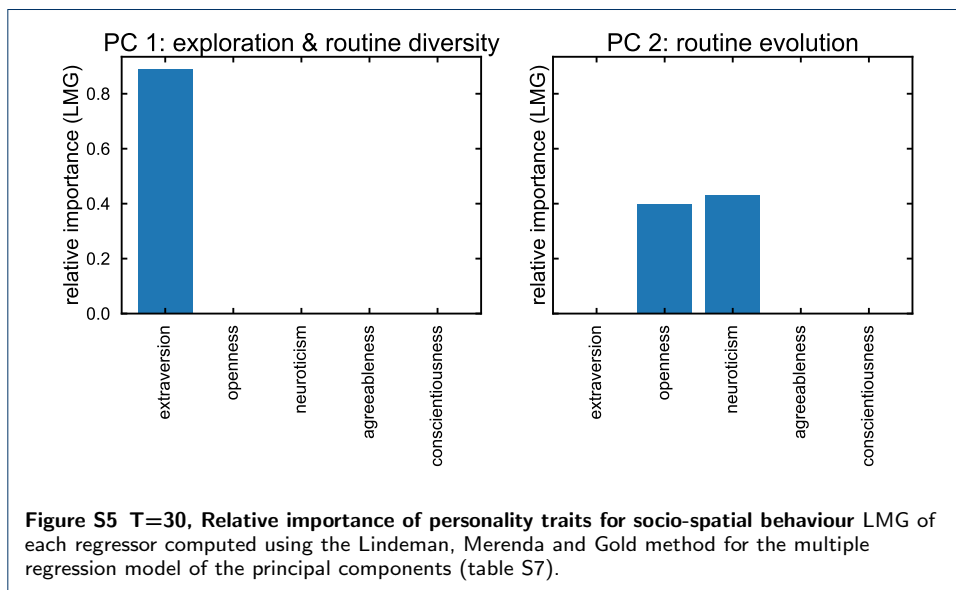
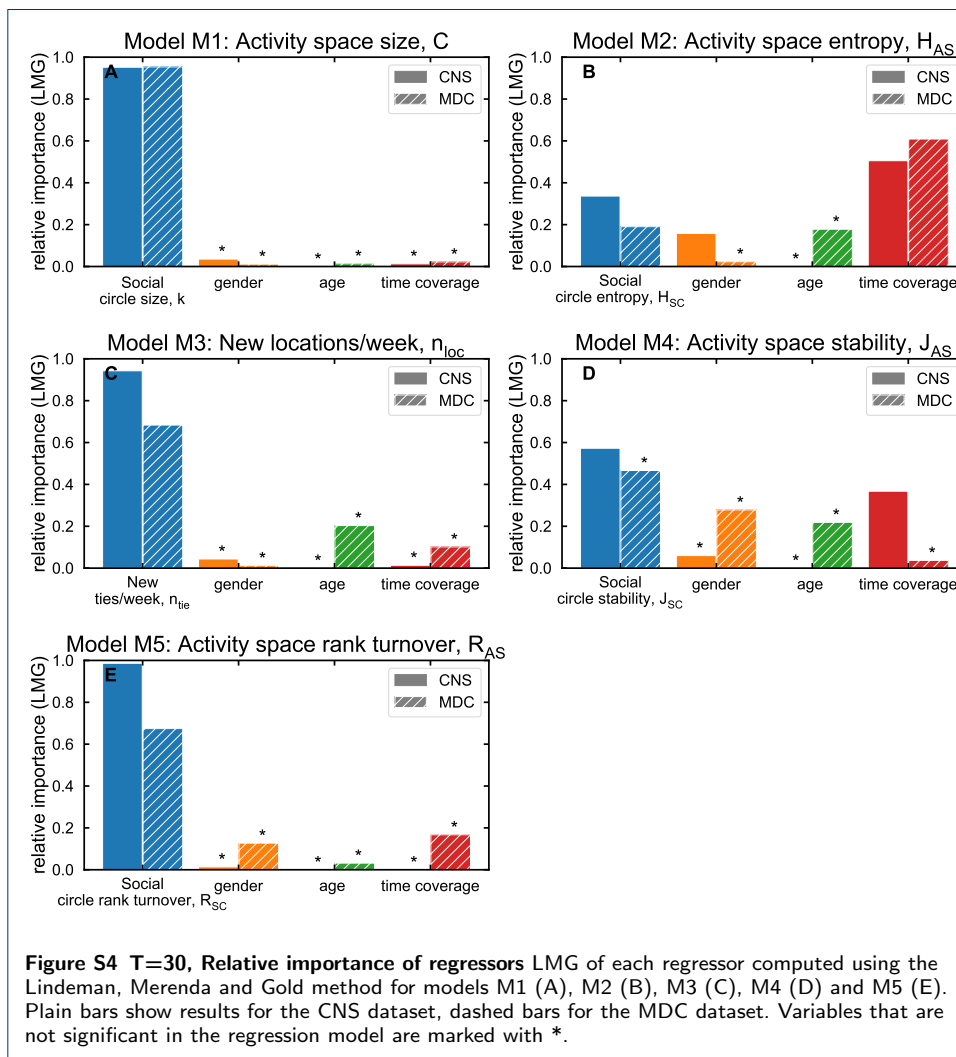


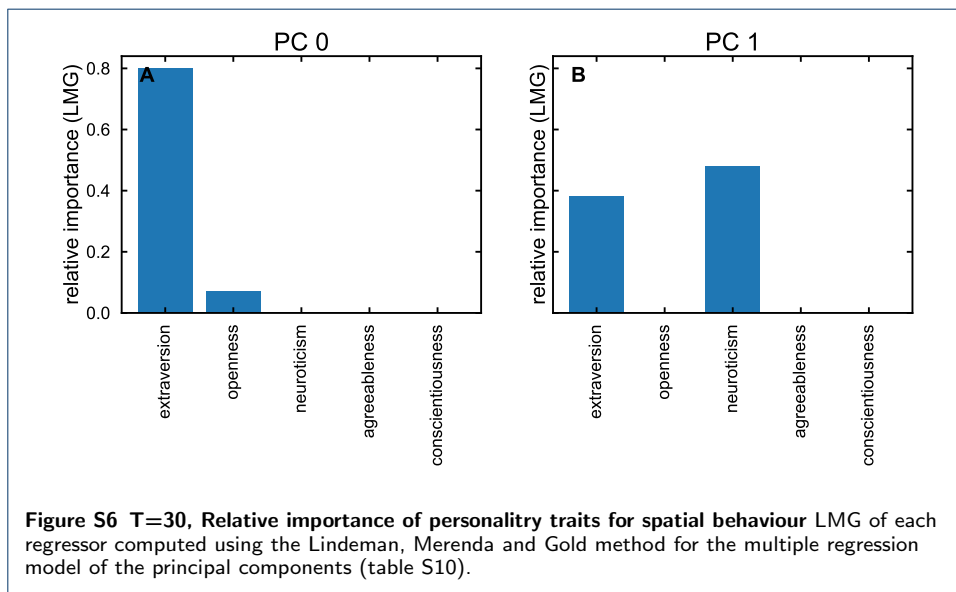
## 2 Results obtained with other windows

Figs. S3, S5, S5, S6 and Tables S3, S4, S5, S6, S7, S8, S9, S10 report the results obtained choosing a time-window with length  $T = 30$  weeks (see main manuscript, section ‘Methods’).











|  | $\overline{d_{self}}$      | $\overline{d_{ref}}$ | $\overline{d_{self}(i) < d_{ref}(i, j)}$ |
|--|----------------------------|----------------------|--|
| Social circle size, $k$                | $0.04 \pm 0.13$            | $15 \pm 6$           | 100%                                     |
| Activity space size, $C$               | $0.04 \pm 0.07$            | $8 \pm 3$            | 99%                                      |
| New ties/week, $n_{tie}$               | $0.06 \pm 0.12$            | $0.9 \pm 0.5$        | 96%                                      |
| New locations/week, $n_{loc}$          | $0.1 \pm 0.2$              | $1 \pm 1$            | 95%                                      |
| Social circle entropy, $H_{SC}$        | $0.002 \pm 0.005$          | $0.7 \pm 0.3$        | 99%                                      |
| Activity space entropy, $H_{AS}$       | $0.002 \pm 0.006$          | $0.4 \pm 0.1$        | 99%                                      |
| Social circle stability, $J_{SC}$      | $(6 \pm 15) \cdot 10^{-4}$ | $0.14 \pm 0.05$      | 100%                                     |
| Activity space stability, $J_{AS}$     | $(6 \pm 11) \cdot 10^{-4}$ | $0.10 \pm 0.04$      | 100%                                     |
| Social circle rank turnover, $R_{SC}$  | $0.04 \pm 0.11$            | $2 \pm 1$            | 99%                                      |
| Activity space rank turnover, $R_{AS}$ | $0.04 \pm 0.20$            | $2 \pm 1$            | 99%                                      |

**Table S3 T=30, CNS dataset: Persistence of social and spatial behaviour.** For each of the social and spatial metrics,  $\overline{d_{self}}$  is the average self-distance and  $\overline{d_{ref}}$  is the reference distance between an individual and all others, averaged across individuals. The third column reports the fraction of cases where  $\overline{d_{self}(i) < d_{ref}(i, j)}$ , averaged across the population.

| <b>Model M1: Activity space size, <math>C</math></b>               | coeff              | p val        | LMG  |
|--|--------------------|--------------|------|
| Social circle size, $k$  | $4 \pm 0$          | $< 10^{-46}$ | 0.95 |
| gender   | $-0.3 \pm 0.2$     | 0.2          | 0.04 |
| time coverage  | $0.5 \pm 0.2$      | 0.05         | 0.01 |
| [ $R^2 = 0.32$ , $F = 91.23$ , $p_F = 0.0$ ]                       |                    |              |      |
| <b>Model M2: Activity space entropy, <math>H_{AS}</math></b>       |                    |              |      |
| Social circle entropy, $H_{SC}$                                    | $0.07 \pm 0.02$    | $< 10^{-4}$  | 0.34 |
| gender   | $-0.05 \pm 0.02$   | $< 10^{-3}$  | 0.16 |
| time coverage  | $-0.09 \pm 0.02$   | $< 10^{-8}$  | 0.51 |
| [ $R^2 = 0.12$ , $F = 26.82$ , $p_F = 0.0$ ]                       |                    |              |      |
| <b>Model M3: New locations/week, <math>n_{loc}</math></b>          |                    |              |      |
| New ties/week, $n_{tie}$   | $0.58 \pm 0.05$    | $< 10^{-30}$ | 0.94 |
| gender   | $-0.09 \pm 0.05$   | 0.04         | 0.04 |
| time coverage  | $0.03 \pm 0.05$    | 0.5          | 0.01 |
| [ $R^2 = 0.22$ , $F = 55.56$ , $p_F = 0.0$ ]                       |                    |              |      |
| <b>Model M4: Activity space stability, <math>J_{AS}</math></b>     |                    |              |      |
| Social circle stability, $J_{SC}$                                  | $0.027 \pm 0.004$  | $< 10^{-9}$  | 0.57 |
| gender   | $0.009 \pm 0.004$  | 0.02         | 0.06 |
| time coverage  | $0.020 \pm 0.004$  | $< 10^{-5}$  | 0.37 |
| [ $R^2 = 0.18$ , $F = 30.32$ , $p_F = 0.0$ ]                       |                    |              |      |
| <b>Model M5: Activity space rank turnover, <math>R_{AS}</math></b> |                    |              |      |
| Social circle rank turnover, $R_{SC}$                              | $0.81 \pm 0.08$    | $< 10^{-19}$ | 0.99 |
| gender   | $0.09 \pm 0.08$    | 0.3          | 0.01 |
| time coverage  | $-0.001 \pm 0.084$ | 1.0          | 0.0  |
| [ $R^2 = 0.18$ , $F = 31.70$ , $p_F = 0.0$ ]                       |                    |              |      |

**Table S4**  $T=30$ , Linear regression models for the CNS dataset. For each model, we report the  $R^2$  goodness of fit, the  $F$ -test statistics with the corresponding p-value  $p_F$ . 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. 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.

|     | PC 0 | PC 1 | PC 2 | PC 3 | PC 4 | PC 5 | PC 6 | PC 7 | PC 8 | PC 9 |
|-----|------|------|------|------|------|------|------|------|------|------|
| CNS | 0.40 | 0.18 | 0.10 | 0.07 | 0.07 | 0.06 | 0.04 | 0.03 | 0.03 | 0.02 |
| MDC | 0.39 | 0.19 | 0.12 | 0.10 | 0.06 | 0.05 | 0.05 | 0.03 | 0.02 | 0.01 |

**Table S5 T=30, Variance explained by principal components.** The fraction of variance explained by each principal component for the CNS and MDC dataset.

|  | CNS   |       | MDC   |       |
|--|-------|-------|-------|-------|
|  | PC 0  | PC 1  | PC 0  | PC 1  |
| Social circle size, $k$                | 0.41  | 0.18  | -0.36 | 0.04  |
| Activity space size, $C$               | 0.42  | -0.23 | -0.40 | -0.05 |
| New ties/week, $n_{tie}$               | 0.33  | 0.27  | -0.24 | -0.35 |
| New locations/week, $n_{loc}$          | 0.39  | -0.11 | -0.37 | -0.22 |
| Social circle entropy, $H_{SC}$        | 0.29  | 0.33  | -0.36 | -0.23 |
| Activity space entropy, $H_{AS}$       | 0.38  | -0.10 | -0.35 | 0.13  |
| Social circle stability, $J_{SC}$      | -0.12 | -0.50 | -0.11 | 0.56  |
| Activity space stability, $J_{AS}$     | -0.06 | -0.50 | -0.03 | 0.62  |
| Social circle rank turnover, $R_{SC}$  | -0.17 | 0.26  | 0.28  | -0.19 |
| Activity space rank turnover, $R_{AS}$ | -0.35 | 0.37  | 0.42  | -0.18 |

**Table S6 T=30, Principal Components.** The weight of each metric in the first two principal components, for both datasets.

|                   | PC 0<br>$R^2 = 0.17, F = 17.08, p_F = 0.0$ |              |      | PC 1<br>$R^2 = 0.02, F = 2.05, p_F = 0.1$ |       |      |
|-------------------|--|--------------|------|---|-------|------|
|                   | coeff                                      | p val        | LMG  | coeff                                     | p val | LMG  |
| extraversion      | $0.9 \pm 0.1$                              | $< 10^{-15}$ | 0.89 | $0.06 \pm 0.07$                           | 0.4   | 0.04 |
| openness          | $-0.18 \pm 0.09$                           | 0.06         | 0.02 | $0.13 \pm 0.07$                           | 0.05  | 0.4  |
| neuroticism       | $0.1 \pm 0.1$                              | 0.1          | 0.03 | $0.15 \pm 0.07$                           | 0.04  | 0.43 |
| agreeableness     | $0.05 \pm 0.10$                            | 0.6          | 0.02 | $-0.04 \pm 0.07$                          | 0.5   | 0.09 |
| conscientiousness | $0.04 \pm 0.10$                            | 0.7          | 0.04 | $-0.03 \pm 0.07$                          | 0.7   | 0.04 |

**Table S7 T=30, Extraversion, openness, and neuroticism explain socio-spatial behaviour.** The result of a multiple linear regression explaining principal components of socio-spatial data (Table S6). 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.

|     | PC 0 | PC 1 | PC 2 | PC 3 | PC 4 |
|-----|------|------|------|------|------|
| CNS | 0.57 | 0.21 | 0.10 | 0.08 | 0.04 |
| MDC | 0.55 | 0.24 | 0.12 | 0.06 | 0.03 |

**Table S8 T=30, Variance explained by principal components (only spatial data).** The fraction of variance explained by each principal component for the CNS and MDC dataset.

|  | CNS   |       | MDC   |       |
|--|-------|-------|-------|-------|
|  | PC 0  | PC 1  | PC 0  | PC 1  |
| Activity space size, $C$               | -0.55 | 0.01  | -0.55 | -0.10 |
| New locations/week, $n_{loc}$          | -0.48 | -0.16 | -0.48 | -0.35 |
| Activity space entropy, $H_{AS}$       | -0.48 | -0.14 | -0.44 | 0.20  |
| Activity space stability, $J_{AS}$     | -0.06 | 0.96  | -0.02 | 0.88  |
| Activity space rank turnover, $R_{AS}$ | 0.48  | -0.16 | 0.51  | -0.22 |

**Table S9 T=30, Principal Components (only spatial data).** The weight of each metric in the first two principal components, for both datasets.

|                   | PC 0                               |             |      | PC 1                              |       |      |
|-------------------|------------------------------------|-------------|------|-----------------------------------|-------|------|
|                   | $R^2 = 0.11, F = 11.55, p_F = 0.0$ |             |      | $R^2 = 0.02, F = 2.02, p_F = 0.1$ |       |      |
|                   | coeff                              | p val       | LMG  | coeff                             | p val | LMG  |
| extraversion      | $-0.56 \pm 0.09$                   | $< 10^{-9}$ | 0.8  | $-0.12 \pm 0.05$                  | 0.03  | 0.38 |
| openness          | $0.20 \pm 0.08$                    | 0.01        | 0.07 | $-0.04 \pm 0.05$                  | 0.5   | 0.08 |
| neuroticism       | $0.03 \pm 0.08$                    | 0.7         | 0.07 | $-0.14 \pm 0.05$                  | 0.01  | 0.48 |
| agreeableness     | $-0.05 \pm 0.08$                   | 0.5         | 0.03 | $-0.002 \pm 0.052$                | 1.0   | 0.01 |
| conscientiousness | $-0.005 \pm 0.081$                 | 1.0         | 0.03 | $-0.03 \pm 0.05$                  | 0.6   | 0.04 |

**Table S10 T=30, Extraversion, openness, and neuroticism explain spatial behaviour.** The result of a multiple linear regression explaining principal components of spatial data (Table S6). 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.