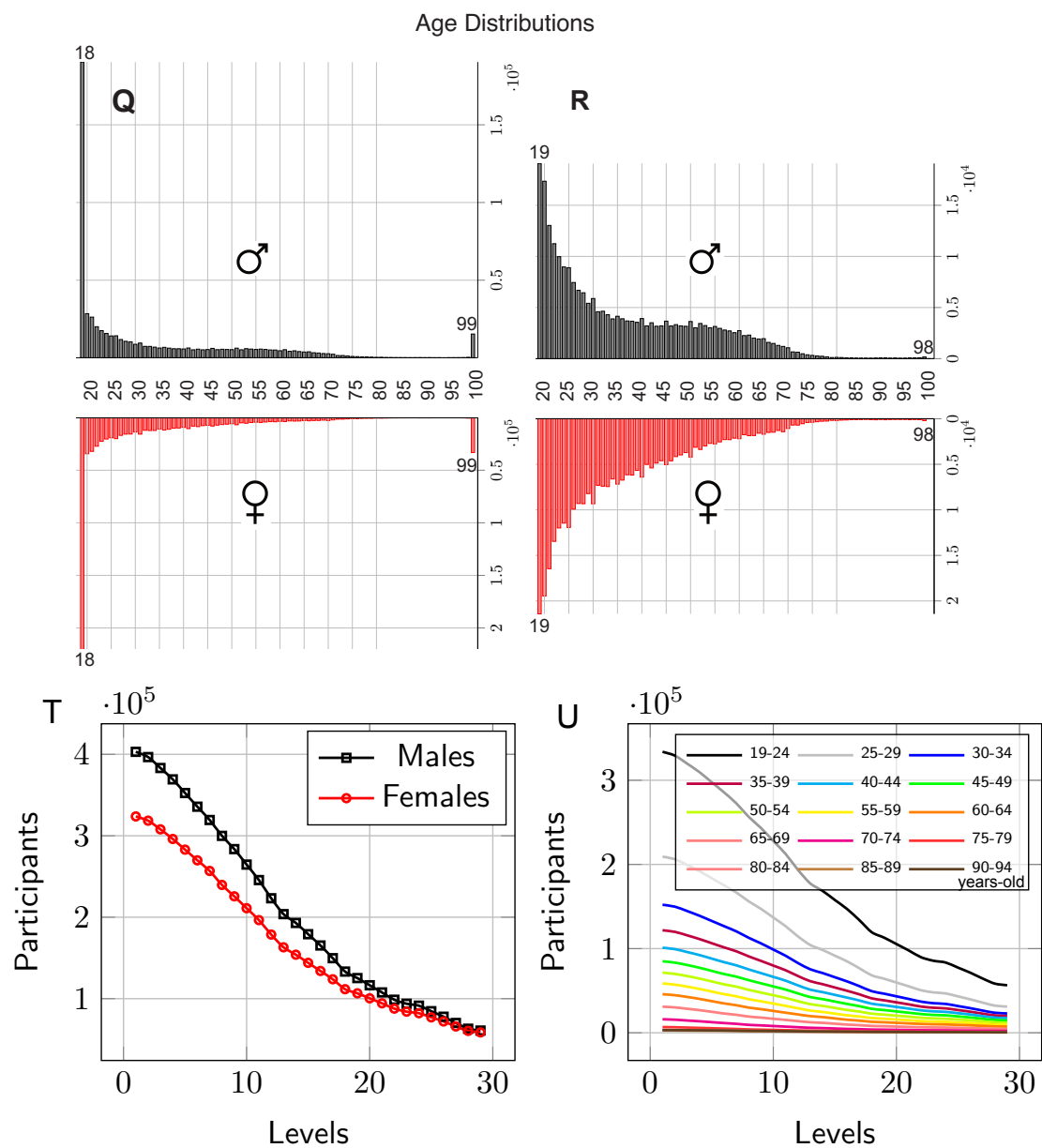
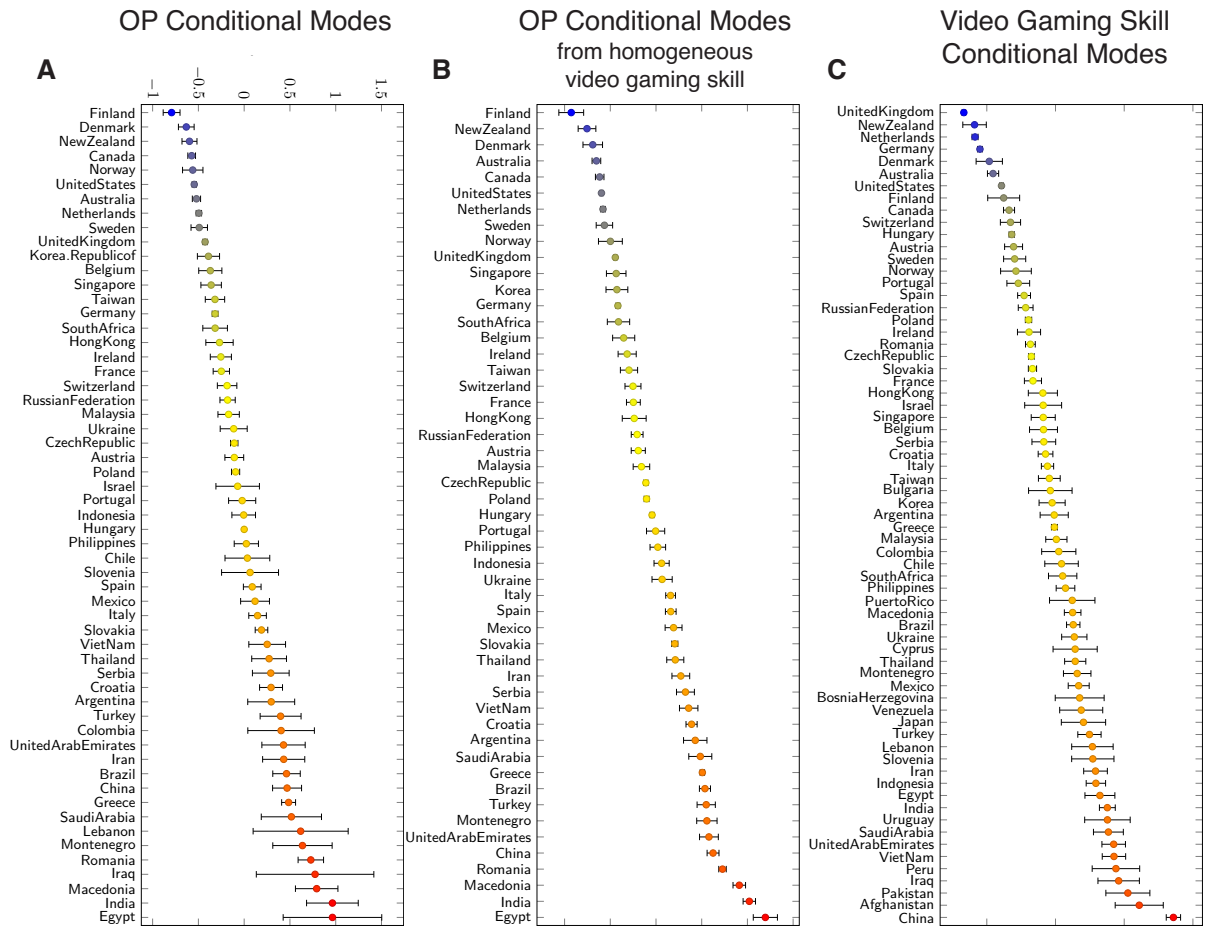


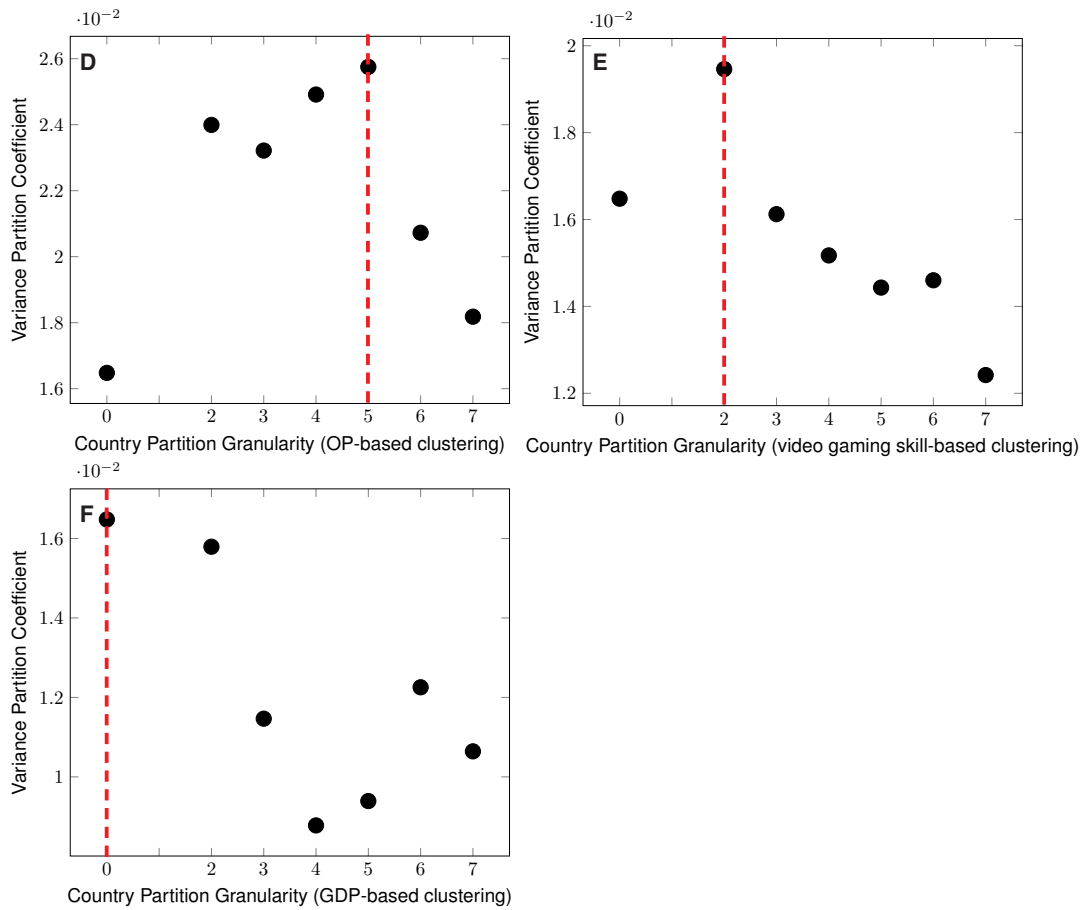
**Figure S1.** (Part A) Screenshots from the game. Related to Figure 1. **A-D** Set-up, game narrative, main menu showing game progression. **E** Wayfinding task. The number of stars (top-centre) decreases with time. **F-K** Maps of wayfinding levels 1, 2, 3, 6, 7 and 8. Map of level 8 is partly blurred. Starting position is indicated by a pale blue arrow, ordered checkpoints by red flags. **L-N** Path integration task. **P** Invitation to enter participant's gender.



**Figure S1.** (Part B) Age distributions and number of participants across levels. Related to Figure 2 Q Age distribution of all participants with demographics (1,446,954 participants). Spikes at 18 and 99 years indicate that the lower and upper bounds of the age range have been misused. R Age distribution of included participants (558,143 participants). T-U Number of participants across levels by gender and age groups.

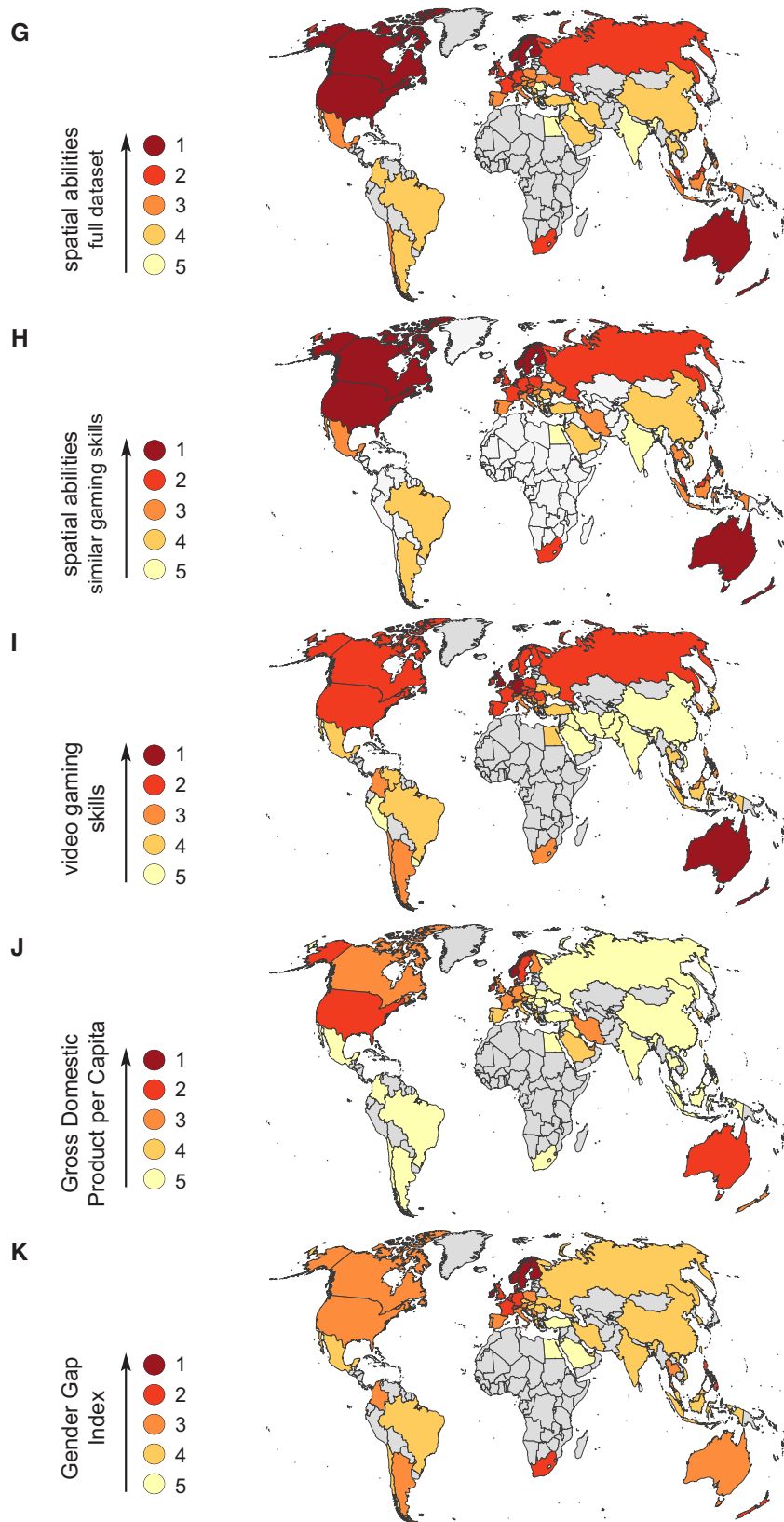


**Figure S2. (Part A) Country Rankings. Related to Figure 2B. A** Ranking based on Overall Performance corrected for video gaming skill (OPcorr), reproduced from Figure 2B for comparison. **B** Ranking based on OPcorr computed from a subset of participants with similar video gaming skills. We included participants within the [0.25 0.75] quantile interval of the distribution of performance at levels 1 and 2. **C** Ranking based on performance at level 1 and 2 reflecting video gaming skill. Rankings are based on the conditional modes from a multilevel model predicting OPcorr with fixed effect for age and gender and random effect for nationality.



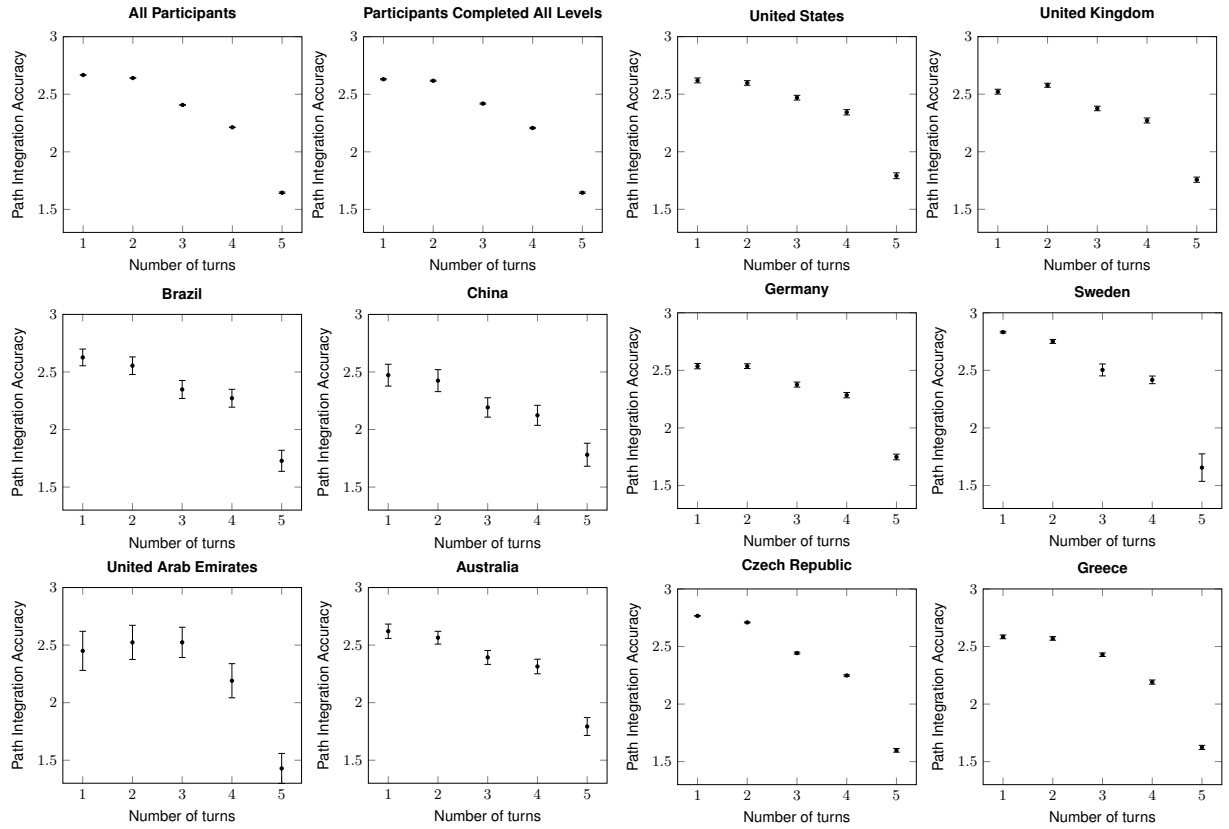
**Figure S2. (Part B) Percentage of variance explained vs. country clustering granularity. Related to Figure 2A.** We fit a multi-level model for Overall Performance corrected for video gaming skill (OPcorr), with fixed effects for age and gender and random effect for nationality. Nationality levels are either individual countries (granularity = 0) or cluster numbers for country partitions with 2, 3, 4, 5, 6 or 7 clusters. Variance Partition Coefficient (VPC) represents the percentage of variance explained by the random effect. **D** Countries are clustered based on OPcorr. VPC is maximal for 5 clusters. **E** Countries are clustered based on performance at level 1 and 2 reflecting video gaming skill. VPC is maximal for 2 clusters. **F** Countries are clustered based on GDP per Capita. VPC is maximal without country clustering. Red dashed line indicates country clustering granularity with highest VPC.





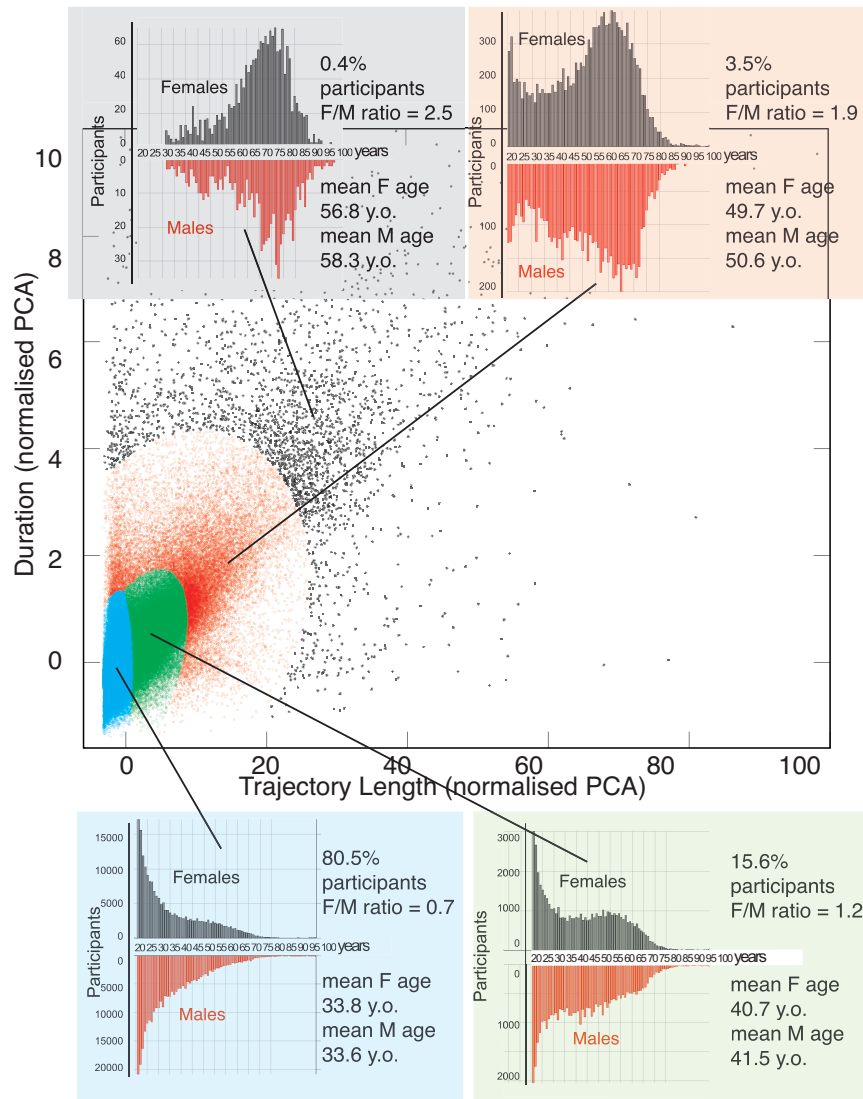
**Figure S2.** (Part C) **Geographical clustering.** Related to **Figure 2A.** **G** 5 clusters of spatial abilities, reproduced from Fig 1f. **H** 5 clusters of spatial abilities from a subset of participants with similar video gaming skills, as in Figure S2B. **I** 5 clusters of video gaming skills based on performance at levels 1 and 2, as in Figure S2C. **J** 5 clusters of Gross Domestic Product per Capita. **K** 5 clusters of Gender Gap Index. Cluster 1 contains the lowest index values.

Level 14: one turn    Level 34: two turns    Level 54: three turns    Level 44: four turns    Level 74: five turns

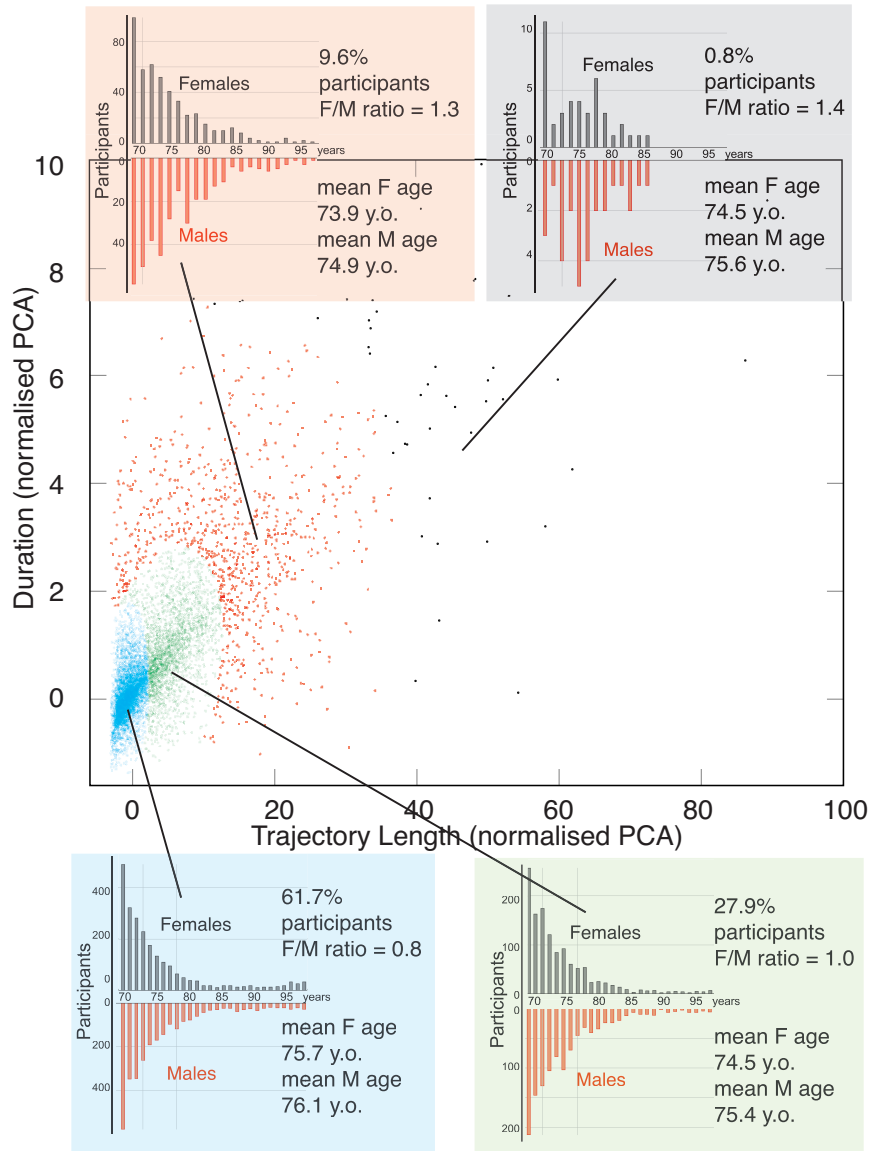


**Figure S2. (Part D) Path integration accuracy vs. path complexity. Related to Figure 2F.** Path integration accuracy is defined as the average number of stars obtained by participants (1, 2 or 3 stars). With increasing complexity: level 14 (1 turn), level 34 (2 turns), level 54 (3 turns), level 44 (4 turns) and level 74 (five turns). The first plot includes all participants:

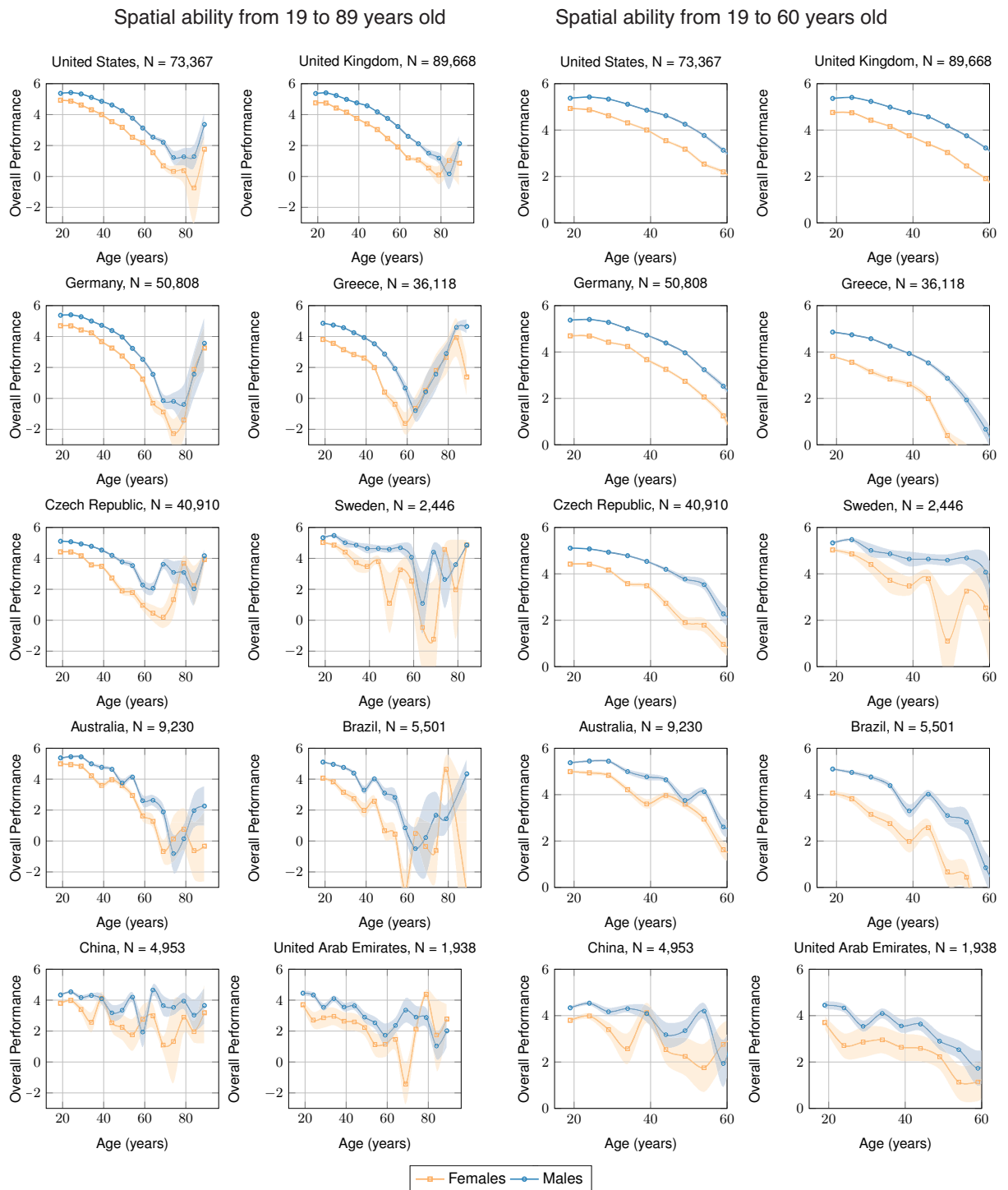
$N_{level14} = 942,628$ ,  $N_{level34} = 208,251$ ,  $N_{level44} = 116,044$ ,  $N_{level54} = 44,704$ ,  $N_{level74} = 19,038$ . The second plot includes participants that completed all five levels. Subsequent plots break down path integration accuracy by country. Error bars correspond to standard errors.



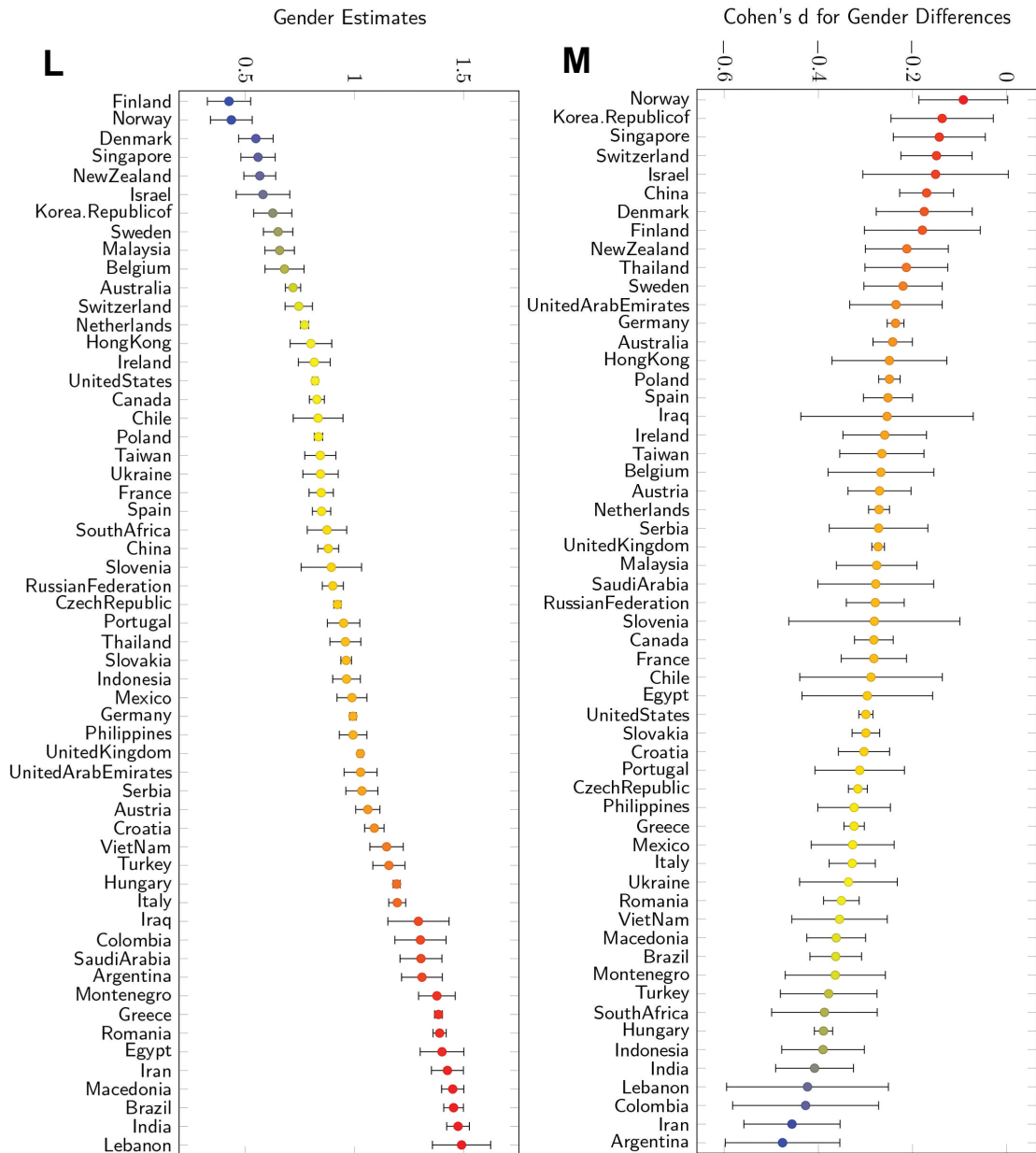
**Figure S2. (Part E) Clusters of efficiency. Related to Figure 2.** Gaussian Mixture Model (4 Gaussians) in the *Duration* × *Trajectory Length* space. For each cluster, the percentage of participants represented, their age and gender distributions are displayed. The lower *Duration* and *Trajectory Length*, the more efficient.



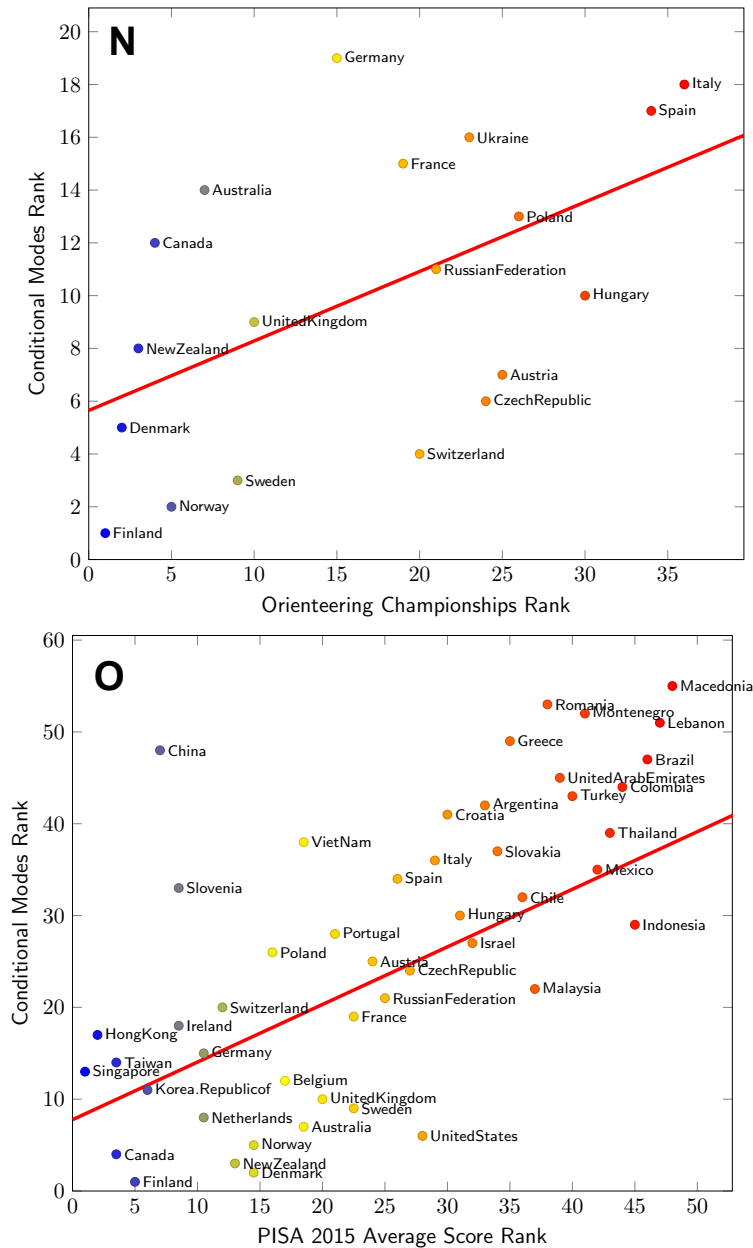
**Figure S2.** (Part F) Clusters of efficiency within participants above 70 years old ( $N = 8653$ ). Related to Figure 2. Gaussian Mixture Model (4 Gaussians) in the *Duration*  $\times$  *Trajectory Length* space. For each cluster, the percentage of participants represented, their age and gender distributions are displayed.



**Figure S2. (Part G) Evolution of spatial abilities across age and gender for 10 example countries. Related to Figure 2E.** Left: between 19 and 89 years old. Right: close up between 19 and 60 years old. Data points correspond to the average OPcorr within 5-year windows. Error bars correspond to standard error.



**Figure S2. (Part H) Gender differences across countries. Related to Figure 2D. L** Gender estimates from a multi-level model for Overall Performance corrected for video gaming skill ( $OP_{corr}$ ), with fixed effects for age and varying slope for gender, nested within nationality:  $OP_{corr} \sim age + (gender|nationality)$ . Positive values mean male advantage. **M** Cohen's d for gender difference in  $OP_{corr}$  (negative values mean male advantage), computed within each country. Gender estimates and Cohen's d are highly correlated (Pearson's correlation  $r = -0.77, p < 0.001$ ). Error bars represent standard errors.



**Figure S2. (Part I) N- Correlation between country performance (CM) and orienteering championships results normalized by population. Related to Figure 2.** Orienteering world championship country results significantly correlate with countries' CM (Pearson's correlation  $\rho = 0.55, p = 0.01$ ). Orienteering championships results are based on the cumulated points of the top 100 male and female orienteers (<http://ranking.orienteering.org>) normalized by country's population. The analysis was carried out on the subset of 19 nations represented in the top 100 orienteering ranking. A multiple linear regression was calculated to predict CM based on GDP per capita and orienteering championship results. Both orienteering championships ( $t(16) = -2.43, p = 0.02$ ) and GDP per capita ( $t(16) = -2.39, p = 0.03$ ) significantly predicted Countries' CM. **O- Correlation between country performance (CM) and average scores at the Programme for International Student Assessment (PISA 2015).** PISA scores are averaged over science, reading and mathematics results. PISA scores significantly correlate with countries' CM (Pearson's correlation  $\rho = 0.73, p < 0.001$ ). A multiple linear regression was calculated to predict CM based on GDP per capita and average PISA scores. Both average PISA scores ( $t(45) = -3.38, p = 0.001$ ) and GDP per capita ( $t(45) = -4.26, p < 0.001$ ) significantly predicted Countries' CM.