

Self-evaluations of navigation ability: a big data approach

You (Lilian) Cheng*¹, Chuanxiuyue (Carol) He*², Hugo J Spiers³, Antoine Coutrot⁴, Michael Hornberger⁵, Mary Hegarty², Elizabeth R. Chrastil¹

*: co-first author: youc3@uci.edu | c_he@ucsb.edu
1. University of California, Irvine, US; 2. University of California, Santa Barbara, US;
3. University College London, UK; 4. Université de Lyon, France; 5. University of East Anglia, UK

Poster No. 2307

INTRODUCTION

Background:

- A previous study has reported effects of age and sex on self-reported navigation ability¹
- A recent meta-analysis of the effect of sex differences on navigation ability did not consider self-reported navigation ability²
- Sea Hero Quest is a phone-based navigation game with data from over 4 million users³
- Previous studies using the Sea Hero Quest dataset reported effects of demographic variables (e.g., age, sex, home environments) on objective navigation performance^{4,5}

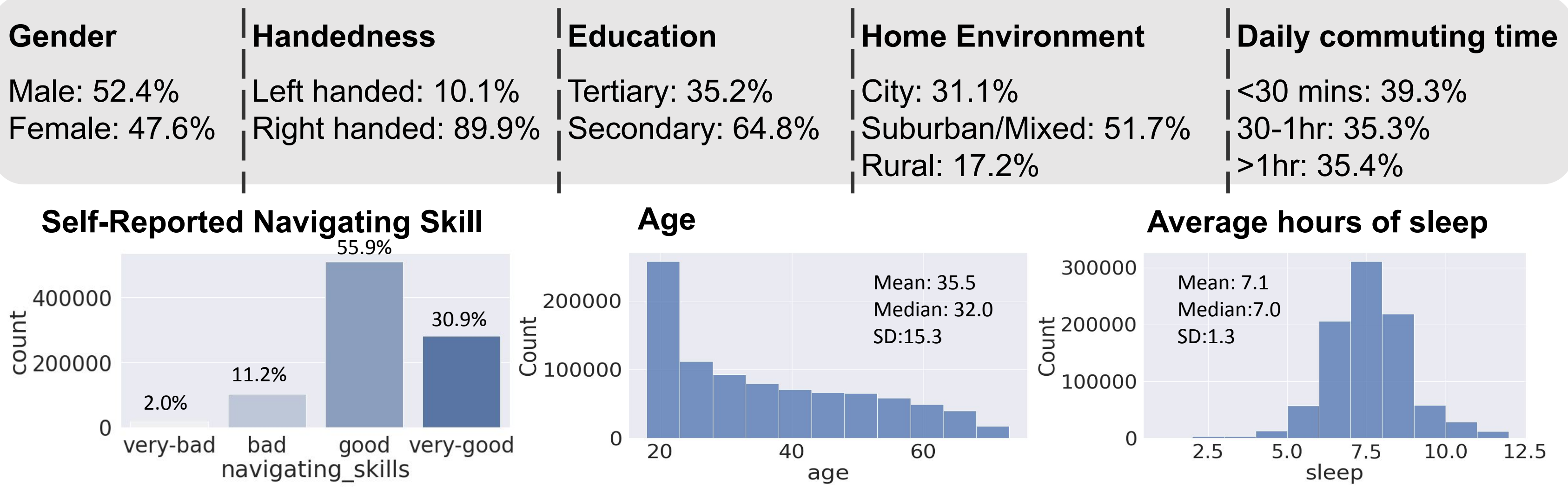
Questions:

- Primary
Are demographic variables related to **self-reported** navigating skills (SRNS)?
- Secondary
What is the structure of demographic characteristics in the Sea Hero Quest dataset? Are there any subpopulations? Does each subpopulation have different navigating skills?

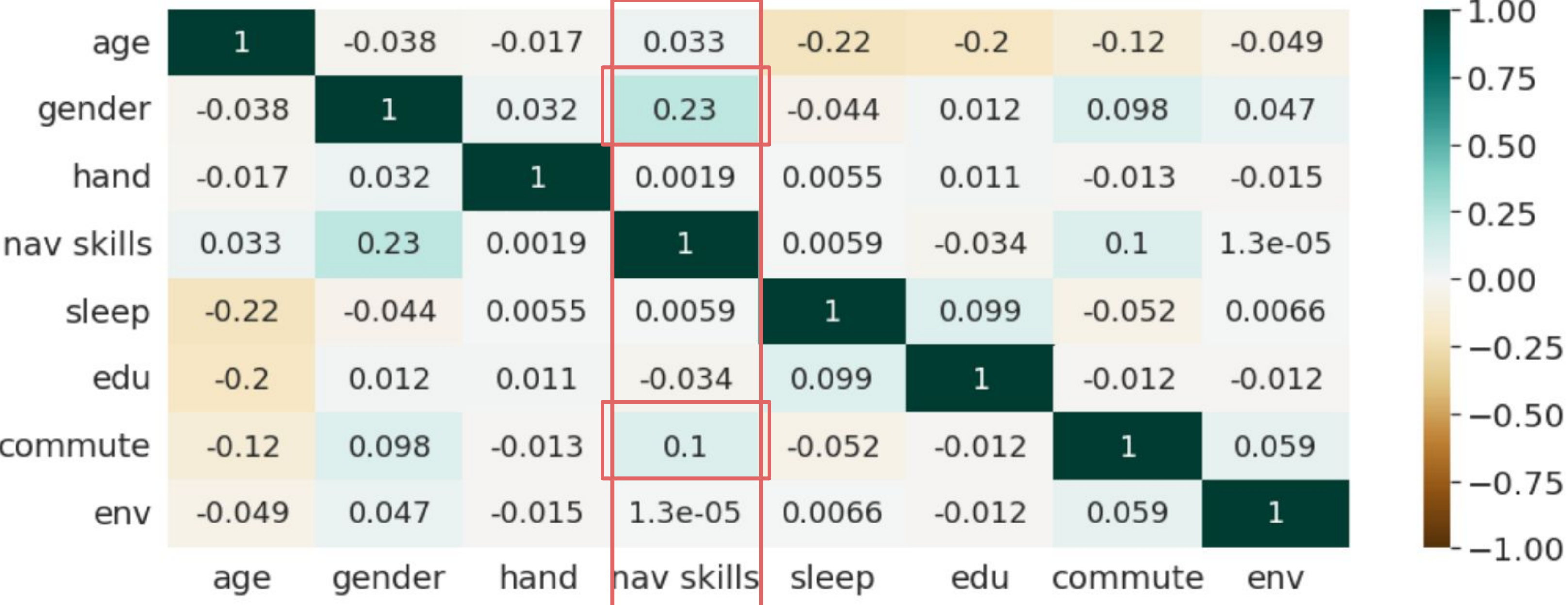
EXPLORATORY DATA ANALYSES

Descriptive Statistics

- N = 90,945



Bivariate Spearman Correlation



- All pairwise correlations among variables are low (< 0.3)

FACTOR ANALYSIS

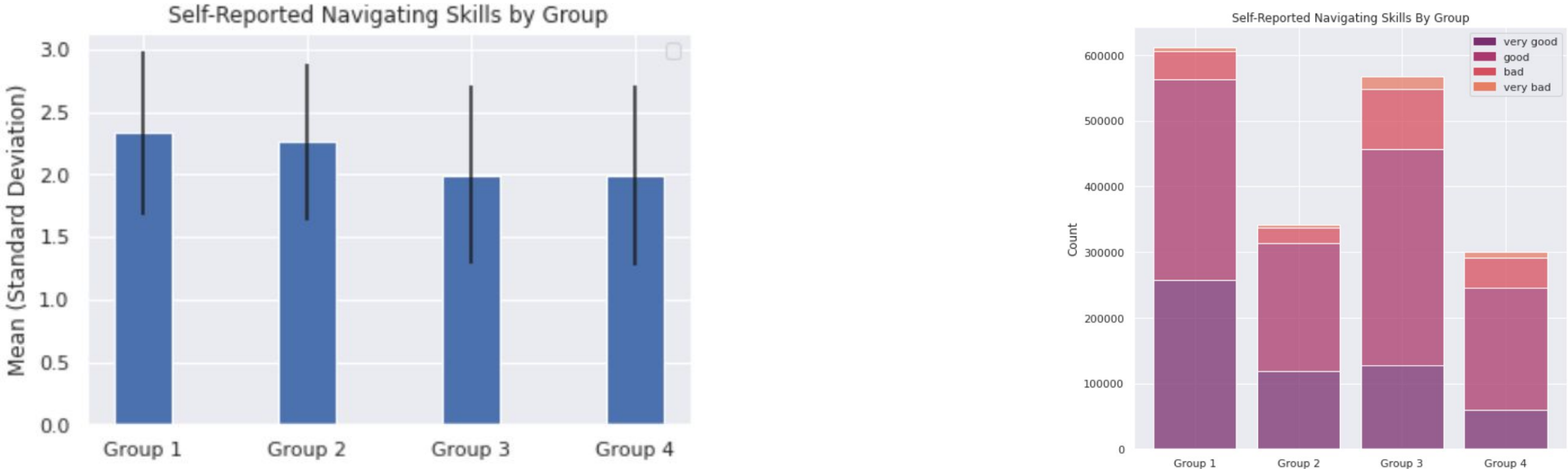
- We conducted factor analysis to see if there are latent factors underlying independent variables
- Bartlett Sphericity Chi Square Test tests whether there is a pattern among variables
Score = 153861 ($p < 0.001^{***}$), indicating there is a pattern among variables
- Kaiser-Meyer-Olkin (KMO) Test tests whether there is sufficient variance for factor analysis
Score = 0.518 < 0.6 (KMO criterion), indicating there is not sufficient variance for factor analysis
- Overall, there is a pattern among independent variables, but there is no latent factor, so the pattern directly comes from independent variables, which can be tested by cluster analyses

CLUSTER ANALYSES

- We conducted cluster analysis to see whether there are subpopulations in the dataset

- K-Means cluster analysis shows data can best be clustered into 4 groups

Group 1: Males secondary (edu level), all ages
Group 2: Males tertiary (edu level) primarily in their 20s
Group 3: Females, secondary (edu level) all ages
Group 4: Females tertiary (edu level) primarily in their 20s



- Means and standard deviations of self-reported navigating skills (SRNS) were similar across groups
- Chi-Square Test of Independence revealed that groups and SRNS were significantly associated $\chi^2 = 50509.76$, $p < 0.001$
- Post-hoc multiple tests (Bonferroni correction) shows the frequency of each SRNS level in each group was significantly different from the expected frequency

ORDINAL LOGISTIC REGRESSION

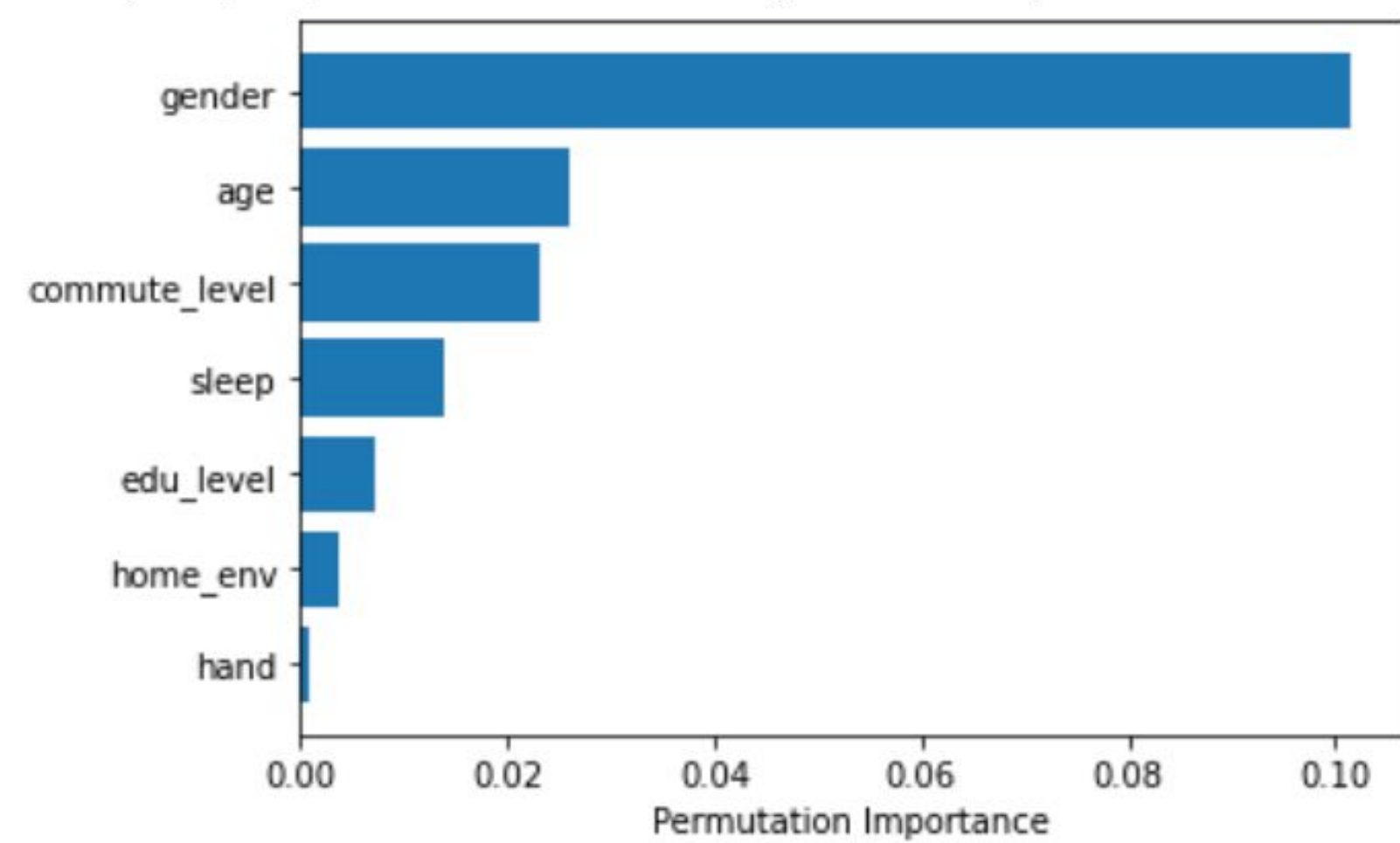
Factor	Coefficient	Standard Error	95%CI	Interpretation on predicting self-reported navigating skills
age	0.024***	0.003	[0.019, 0.030]	Older people +
gender (male = 1)	0.899***	0.004	[0.891, 0.908]	Male > Female
age by gender	0.148***	0.004	[0.140, 0.156]	Males have a stronger age effect
handedness (left = 1)	-0.021**	0.007	[-0.034, -0.008]	Left-handed < Right-handed
sleep	0.067***	0.002	[0.062, 0.071]	Longer daily sleep time +
education level	0.130***	0.004	[0.121, 0.138]	Higher education +
commute level	0.187***	0.002	[0.182, 0.191]	Longer commute time +
home environment	-0.031***	0.002	[-0.035, -0.026]	More city-like home environment -

*** $p < 0.001$ ** $p < 0.01$

- Gender is the strongest predictor, followed by commute level and age by gender interaction

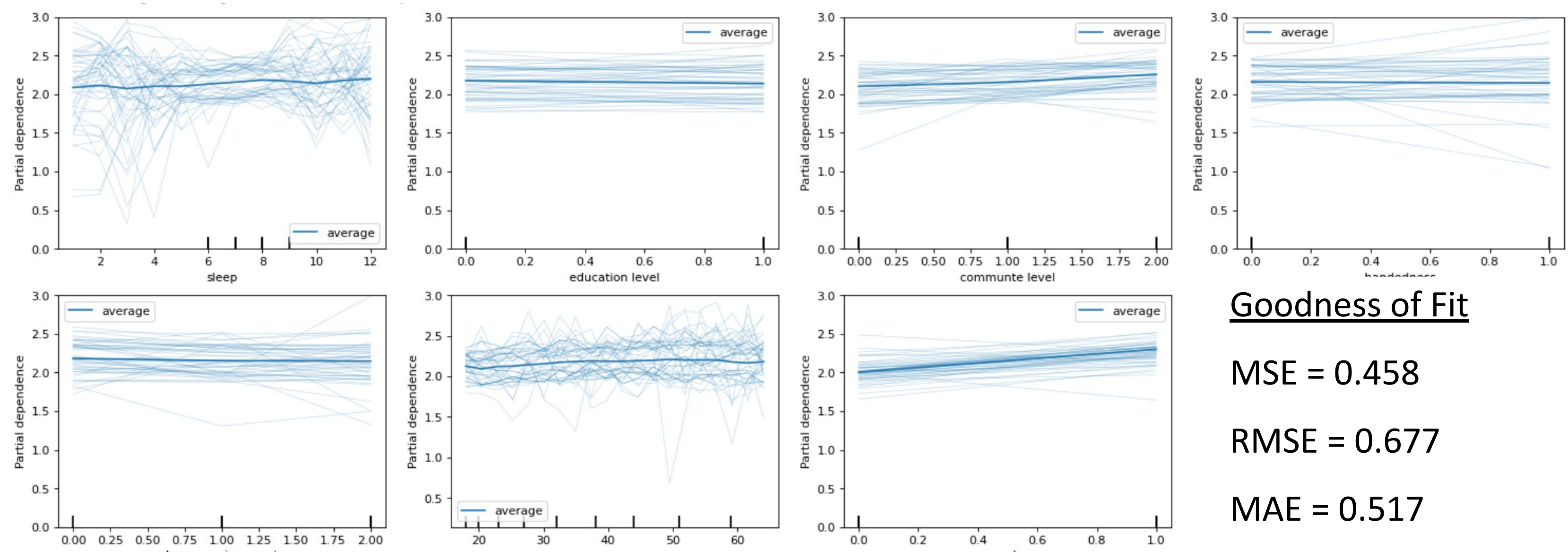
RANDOM FOREST REGRESSION

Feature Importance



- Gender is the most important factor in predicting self-reported navigating skills, followed by age, commute level, sleep, education level, home environment, and handedness
- The feature importance ranking is consistent with the effect size ranking in logistic regression analyses except that the rankings of sleep and education were swapped

Partial Dependence Plots with Individual Conditional Expectation



Goodness of Fit

MSE = 0.458

RMSE = 0.677

MAE = 0.517

- Partial Dependence Plot shows the marginal effect a featurn has on the predicted outcome
- Marginal Effect: the effect of a predictor by marginalizing over all other predictors
- The marginal effect of
 - sleep on self-reported navigating skills (SRNS) is positive: more hours of sleep was related to reporting higher scores
 - age on SRNS is positive: older people reported higher scores
 - gender on SRNS is positive: males reported higher navigating skills than females
 - commuting time on SRNS is positive: longer commuting times were related to reporting higher scores

DISCUSSION & CONCLUSIONS

Conclusions:

- Primary - Relations between demographic characteristics and self-reported navigating skills (SRNS)
 - Gender, age and average commuting time affects SRNS
 - Gender is the strongest predictor but all predictors are weak
 - Men reported greater skills than women
 - Older people, especially men, reported greater skills than younger people (aging trend was more evident in men than in women)
- Secondary - structure of demographic variables or subpopulations:
 - No latent factor
 - People can be clustered into 4 groups that mainly differ by age, gender, and education
 - Clusters and SRNS were significantly associated, consistent with the regression and random forest analyses

Limitations:

- Individual Level Prediction Accuracy is Low (Variation is large)
 - People with the same demographic characteristics can have different self-report navigating skills
 - People with the same SRNS can have various demographic characteristics
 - Thus, the demographic information in this dataset is insufficient to conduct individual level predictions on SNRS

Future Direction:

- Including objective measures to see if people with different demographic characteristics underestimated or overestimated themselves.

REFERENCES

- van der Ham, et al. (2021). Quality of self-reported cognition: Effects of age and gender on spatial navigation self-reports. Aging & mental health, 25(5), 873-878.
- Nazareth, A., et al. (2019). A meta-analysis of sex differences in human navigation skills. Psychonomic bulletin & review, 26(5), 1503-1528.
- Coughlan, G., et al. (2019). Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. Proceedings of the National Academy of Sciences, 116(19), 9285-9292.
- Coutrot, A., et al.(2018). Global determinants of navigation ability. Current Biology, 28(17), 2861-2866.
- Coutrot, A., et al. (2020). Cities have a negative impact on navigation ability: evidence from 38 countries. Biorxiv.