Sex difference in the predictability of mobility

Clint D. Kelly1,✉

1 Département des Sciences biologiques, Université du Québec à Montréal, Montréal, Canada

✉ Correspondence: [Clint D. Kelly <kelly.clint@uqam.ca>](mailto:kelly.clint@uqam.ca)

# Abstract

# Introduction

# Methods

### Statistical analysis

We removed individuals from the final dataset that died during the study (n = 14) or that escaped during testing (n = 2). Our final sample sizes (and summary statistics) are given in . Statistical analyses were performed within the R v. 4.2.1 statistical environment (R Core Team 2024). We assessed the impact of our treatments on body mass by using an analysis of covariance (ANCOVA) wherein body mass after treatment was entered as the response variable and initial body mass was entered as a covariate. ANCOVA is the preferred method for analyzing pre-post data rather than ANOVA on change scores or repeated measures ANOVA, which can be biased when regression toward the mean is present (Bland and Altman 1994; Dimitrov and Rumrill Jr 2003). Sex, age, and treatment were also entered into the model as fixed factors.

The variable total distance travelled (cm) was Yeo-Johnson transformed to approximate a Gaussian error distribution (Yeo and Johnson 2000). We adjusted the latency to visit all quadrant scores by subtracting an individual’s time from the maximum time available (600 s) so that higher values represent more exploratory individuals. This variable was orderNorm transformed to approximate a Gaussian error distribution (Peterson and Cavanaugh 2019). All data transformations were performed using the R package *bestNormalize* (Peterson 2021).

We report posterior means with 95% credibility intervals (CrI) from Bayesian generalized linear mixed-effects models (*brms* package: Bürkner 2017). Inference was based on CrIs that did not overlap zero. Models were run for 8000 iterations (500 warmups) on 4 chains, using relatively uninformative, default priors and a thinning interval of 2 (total post-warmup samples = 9000). Posterior predictive checks were performed to ensure adequate model fits and trace plots confirmed that models converged with low among-chain variability (Rhat = 1.00).

We investigated the average effects of treatment, age, and sex on travel distance and latency to visit all quadrants using multivariable general linear mixed-effects models. We constructed two models for each of our two dependent behavioural variables and compared them using leave-one-out cross-validation (LOO) (Vehtari et al. 2017). First, we built a full model having age (two levels), sex (two levels), treatment (five levels), test number (two levels) and their interactions as fixed effects (model 1). We then removed the interactions if non-significant and re-ran the model (model 2). In all models, individual ID was included as a random intercept separately for each age, sex and treatment combination. An age-by-sex-by-treatment interaction in the residual part of the models was included to estimate the within-individual (residual) variance for each combination of age, sex, and treatment. This model structure permitted us to test not only for the effects of age, sex, and treatment on average behaviours but to also quantify their effects on among-individual (VA) and residual within-individual behavioural variance (VW) separately for each age, sex, and treatment combination (see Chapple et al. 2022; Polverino et al. 2023). Both dependent variables were scaled (mean=0, SD=1) before analysis to aid model fitting and we assumed Gaussian error structure for both variables. Test number was left-centered (i.e., test 1 = 0) to set the model intercept at the first trial.

We calculated magnitude differences (V) in both among- (VA) and within- (VW) individual variances to test whether these components of variation differed between treatments in young and old female and male beetles (Royauté and Dochtermann 2021). We leveraged our Bayesian framework to directly estimate the distribution of V by taking the difference in the posterior distribution of each combination of interest. We interpret the posterior mode of ΔV as the estimated strength of ΔV, with 95% credible intervals representing the precision around our estimates (Royauté and Dochtermann 2021).

We calculated the unpartitioned phenotypic correlation between distance travelled and latency to visit all quadrants for each age, sex, and treatment combination using Pearson’s product-moment correlation (*r*p ± 95% confidence intervals). We then partitioned the correlations into their components among (*r*ind) and within (*r*e) individuals by running separate bivariate linear mixed-effects models for each age, sex, and treatment combination. Each model included distance travelled and latency to visit all quadrants as response variables and individual ID as a random intercept. Like the univariate models above, we ran the bivariate models on four chains for 8000 iterations (1000 warmups), using weakly informative, default priors.

All data were visualized using the R packages *ggplot2* (Wickham 2016) and *tidybayes* (Kay 2020).

# Results

The population level mean residual standard deviation for nightly travel distance was estimated to be 7.66 m (back-transformed intercept of the dispersion model exp(y0) = -0.47; Supp Table 1). As predicted, females were more predictable in their nightly travel distance than males (ysex = 0.32 [0.04, 0.57]). The predicted standard deviation from the mean residual standard deviation (rIIV) for movement varied across individuals (ω2ID = 0.12 [0.01, 0.32]), which suggests that predictability of nightly movement differs across individuals.

Discussion

# Declarations

**Competing interests** The authors have no financial or non-financial competing interests to declare.

**Authors’ contributions** Both authors contributed to the conception and design of the study. AB collected the data. CDK analyzed the data and wrote the first draft of the manuscript. AB and CDK reviewed and approved the manuscript.

**Funding** Natural Sciences and Engineering Research Council of Canada (NSERC).

**Availability of data and materials** Data will be made available on Open Science Framework at the time of publication.

# References

Bland J, Altman D. 1994. Some examples of regression towards the mean. BMJ. 309(6957):780.

Bürkner P-C. 2017. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software. 80(1):4308.

Chapple D, Naimo A, Brand J, Michelangeli M, Martin J, Goulet C, Brunton D, Sih A, Wong B. 2022. Biological invasions as a selective filter driving behavioral divergence. Nat Commun. 13(1):5996.

Dimitrov DM, Rumrill Jr PD. 2003. Pretest-posttest designs and measurement of change. Work. 20(2):159–165.

Kay M. 2020. tidybayes: Tidy Data and Geoms for Bayesian Models. R package version 305. <http://mjskay.github.io/tidybayes/>.

Peterson RA. 2021. Finding optimal normalizing transformations via best normalize. R Journal. 13(1):2312–2327.

Peterson R, Cavanaugh J. 2019. Ordered quantile normalization: A semiparametric transformation built for the cross-validation era. Journal of applied statistics.

Polverino G, Aich U, Brand J, Bertram M, Martin J, Tan H, Soman V, Mason R, Wong B. 2023. Sex-specific effects of psychoactive pollution on behavioral individuality and plasticity in fish. Behav Ecol. 34(6):969–978.

R Core Team. 2024. R: A language and environment for statistical computing.

Royauté R, Dochtermann NA. 2021. Comparing ecological and evolutionary variability within datasets. Behav Ecol Sociobiol. 75(9).

Vehtari A, Gelman A, Gabry J. 2017. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 27(5):1413–1432.

Wickham H. 2016. ggplot2: Elegant graphics for data analysis. Springer.

Yeo I, Johnson RA. 2000. A new family of power transformations to improve normality or symmetry. Biometrika. 87(4):954–959.