

MODELING HIPPOCAMPUS VOLUME IN CAPTIVE AND WILD CHIMPANZEES

RICHARD MCELREATH

1. PURPOSE

The hippocampus is associated with spatial memory and navigation in humans (Maguire et al. 2000) and other animals. Chimpanzees who develop in the wild appear to learn a large number of food locations and their seasonality, and then use this knowledge to navigate the environment (Janmaat et al. 2013). Chimpanzees who develop in captivity lack the opportunities or need for this knowledge. It is unknown whether the difference between wild and captive environments leads to anatomical differences in the chimpanzee hippocampus. However if hippocampus differences exist, this supports the proposition that wild chimpanzees rely upon large amounts of spatial memory for strategic navigation. The age at which differences are detectable would also help us understand how the knowledge develops.

The goal here is to use a new sample of chimpanzee brain scans to investigate the anatomical differences between the hippocampi of captive and wild resident chimpanzees. We investigate both overall differences in volume as well as differences in sub-regions. Our specific causal estimand is the direct effect of residence environment (captive versus wild) on hippocampus volume. The direct effect is needed, because hippocampi may differ through indirect effects on overall brain size.

2. PRIOR WORK

x

3. THE SAMPLE

The sample comprises scans [scanner language here!] of post-mortem brains of 16 chimpanzees. Seven were residing in captivity at time of death. Nine were residing in the wild. Ages at time of death range from 1.7 years to approximately 56 years of age. Later ages are only approximate, because wild individuals cannot be precisely aged unless their birth was observed.

DEPARTMENT OF HUMAN BEHAVIOR, ECOLOGY AND CULTURE, MAX PLANCK INSTITUTE FOR EVOLUTIONARY ANTHROPOLOGY, DEUTSCHER PLATZ 6, 04103 LEIPZIG, GERMANY

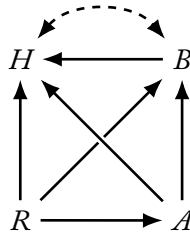
E-mail address: richard_mcelreath@eva.mpg.de.

Date: April 16, 2025.

4. MODELING STRATEGY

We consider two model families. The first treats the problem as a general measurement problem and adjusts for individual brain volume as a covariate. The second builds a simple brain growth model and treats hippocampus volume as an age and residence specific proportion of total brain volume.

Before developing the statistical models, it is necessary to make clear what we believe about the causal structural relations among the variables. In the minimal model, we want to interpret gross size differences in hippocampi between captive and wild individuals as the causal effect of developing in captive versus wild environments. Causal inference requires structural causal assumptions. In the simplest form, these assumptions can be represented as directional causal relations among the variables. Let H be an individual's hippocampus volume at time of death, B the same individual's brain volume, A the individual's age, and R the residential environment (captive or wild) of the individual. Then the structural relations, which variables "listen" to changes in others, can be drawn as:



The treatment of interest R potentially influences age at death A , total brain size B , and the outcome of interest H . Total brain size B is a strong causal influence on H . Age A certainly influences B and may directly influence H as well. Finally, there may be unmeasured common influences on H and B as well, represented by the dashed double-headed arrow between them.

For the estimand of the direct effect of R on H , by the back-door criterion (Pearl 1993), there are no back-door paths into R . However since we want the direct effect only, the minimum adjustment set is $\{A, B\}$ and the unmeasured confounds between H and B . We must stratify by B to close the indirect causal path $R \rightarrow B \rightarrow H$. And we must stratify by A to close a possible indirect path $R \rightarrow A \rightarrow H$.

When we stratify by B , this opens a biasing path through the unmeasured confounds of H and B (the dashed curve at top). Unfortunately we know of no strategy for eliminating these confounds, so any estimates must be interpreted as potentially biased. This bias is common to nearly all studies of this kind, since individual genetic and environmental common causes of brain anatomy are rarely measured.

Note that the total causal effect of R on H is not biased and requires no stratification, despite the unmeasured confounds. There are no biasing paths. But the total effect is not of direct interest, since it would include the influence of overall larger brains.

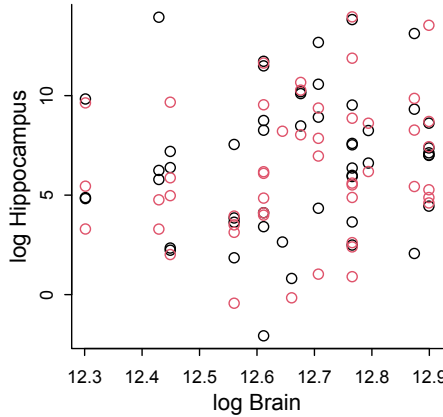
4.1. Measurement model. Each individual contributes multiple hippocampus measurements across two hemispheres. These measurements are correlated within individuals, across hemispheres. Our goal is to model this covariance and the mean values as functions of residence and total brain volume.

$$\begin{aligned} \begin{bmatrix} \log H_{Li} \\ \log H_{Ri} \end{bmatrix} &\sim \text{MVNormal} \left(\begin{bmatrix} \mu_{Li} \\ \mu_{Ri} \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right) \\ \mu_{Li} &= \alpha_{R[i]} + \beta \log B_i + \gamma A_i \\ \mu_{Ri} &= \mu_{Li} + \delta \end{aligned}$$

65 If hippocampus volume scales directly with total brain volume, then β should be 1. Then the meaning of the parameter α is the logarithm of the hippocampus' proportion of total brain volume. We don't want to set this prior based upon our sample of chimpanzees. But using humans as a reference, a typical adult hippocampus is about 3500 mL (Woolard & Heckars 2012). If we use 1100 cm³ as a reference
70 adult human brain volume, then the hippocampus proportion is about 0.0032, so α should be approximately $\log 0.0032 = -5.74$. Using these guidelines, we assign a set a weakly-informative priors for the data model above.

$$\begin{aligned} \alpha_R &\sim \text{Normal}(-5.7, 2) \\ \delta &\sim \text{Normal}(0, 0.5) \\ \beta &\sim \text{Normal}(1, 0.5) \\ \gamma &\sim \text{Normal}(0, 0.5) \\ \rho &\sim \text{LKJCorr}(4) \\ \sigma &\sim \text{Exponential}(1) \end{aligned}$$

And we can simulate from these priors to visualize the prior predictive distribution of implied hippocampus volumes. Showing captive in black and wild in red:



75 The prior predictive distribution for these priors is surely too wide—it includes many very large hippocampus volumes. However it contains no bias for a difference

between wild and captive, and the wide priors help us to notice model misspecification when it is present. The code base contains a module for reproducing and plotting this prior predictive distribution.

4.2. Growth model. x

Hippocampus volume H is necessarily an age-specific proportion p of total brain volume B .

$$H = p_A B \quad (1)$$

On the log scale, which we'll use for modeling, this is:

$$\log H = \log p_A + \log B$$

The proportion p_A could increase or decrease with age, because while the volume of the hippocampus increases with age, and the volume of the brain increases with age, the relative volume could increase or decrease.

4.3. Dimensionless analysis. Often it is revealing to divide out all measurement scales and express the relations among measurements as dimensionless ratios.

5. IMPLEMENTATION

We implement the models above in the probabilistic programming language Stan, through the `cmdstanr` interface in R (Gabry et al. 2025) and the `rethinking` package (McElreath 2024). The `ulam` model specification for the base covariance measurement model is:

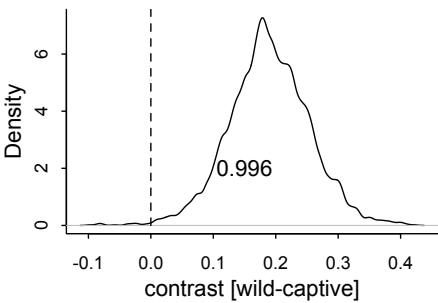
```
c(logHL, logHR) ~ multi_normal(c(muL, muR), Rho, Sigma),
muR <- muL + delta,
muL <- a[R] + bB*logB + g*A,
a[R] ~ normal(-5.7, 2),
delta ~ normal(0, 0.5),
bB ~ normal(1, 0.5),
g ~ normal(0, 0.5),
Rho ~ lkj_corr(4),
vector[2]:Sigma <- rep_vector(sigma, 2),
sigma ~ exponential(1)
```

We drew 4000 total samples from 8 chains with 500 warmup steps for each chain. There were no problems in exploring the posterior distribution, as indicated by the R-hat convergence diagnostic, trace rank plots, and bulk effective sample size estimates. See Vehtari et al. (2021) for definition and discussion of convergence diagnostics. The marginal posterior summaries and diagnostics for each parameter are given below.

	mean	sd	5.5%	94.5%	rhat	ess_bulk
a[1]	-4.86	1.33	-6.98	-2.75	1	1206.96
a[2]	-4.67	1.32	-6.76	-2.57	1	1205.32
delta	0.02	0.03	-0.02	0.07	1	2022.12
bB	0.94	0.11	0.77	1.11	1	1206.12

g	0.00	0.00	-0.01	0.00	1	3902.55
Rho[2,1]	0.57	0.16	0.30	0.78	1	1992.20
Rho[1,2]	0.57	0.16	0.30	0.78	1	1992.20
sigma	0.13	0.02	0.10	0.16	1	1828.53

120 The first two parameters are the log-intercepts for captive, $a[1]$, and wild, $a[2]$, individuals. We require the contrast between these two distributions, which is given below.



125 Approximately 99.6% of the posterior mass is above zero, indicating strong evidence for larger average proportional hippocampi in wild-resident individuals.

REFERENCES

- Gabry, J., Češnovar, R., Johnson, A., and Bröder, S. (2025). *cmdstanr: R Interface to 'CmdStan'*. R package version 0.9.0, <https://discourse.mc-stan.org>.
- 130 Janmaat, K. R., Ban, S. D., and Boesch, C. (2013). Chimpanzees use long-term spatial memory to monitor large fruit trees and remember feeding experiences across seasons. *Animal Behaviour*, 86(6):1183–1205.
- Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., and Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97(8):4398–4403.
- 135 McElreath, R. (2024). *rethinking: Statistical Rethinking book package*. R package version 2.42.
- Pearl, J. (1993). [bayesian analysis in expert systems]: Comment: Graphical models, causality and intervention. *Statistical Science*, 8(3).
- 140 Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved \hat{r} for assessing convergence of mcmc (with discussion). *Bayesian Analysis*, 16(2).