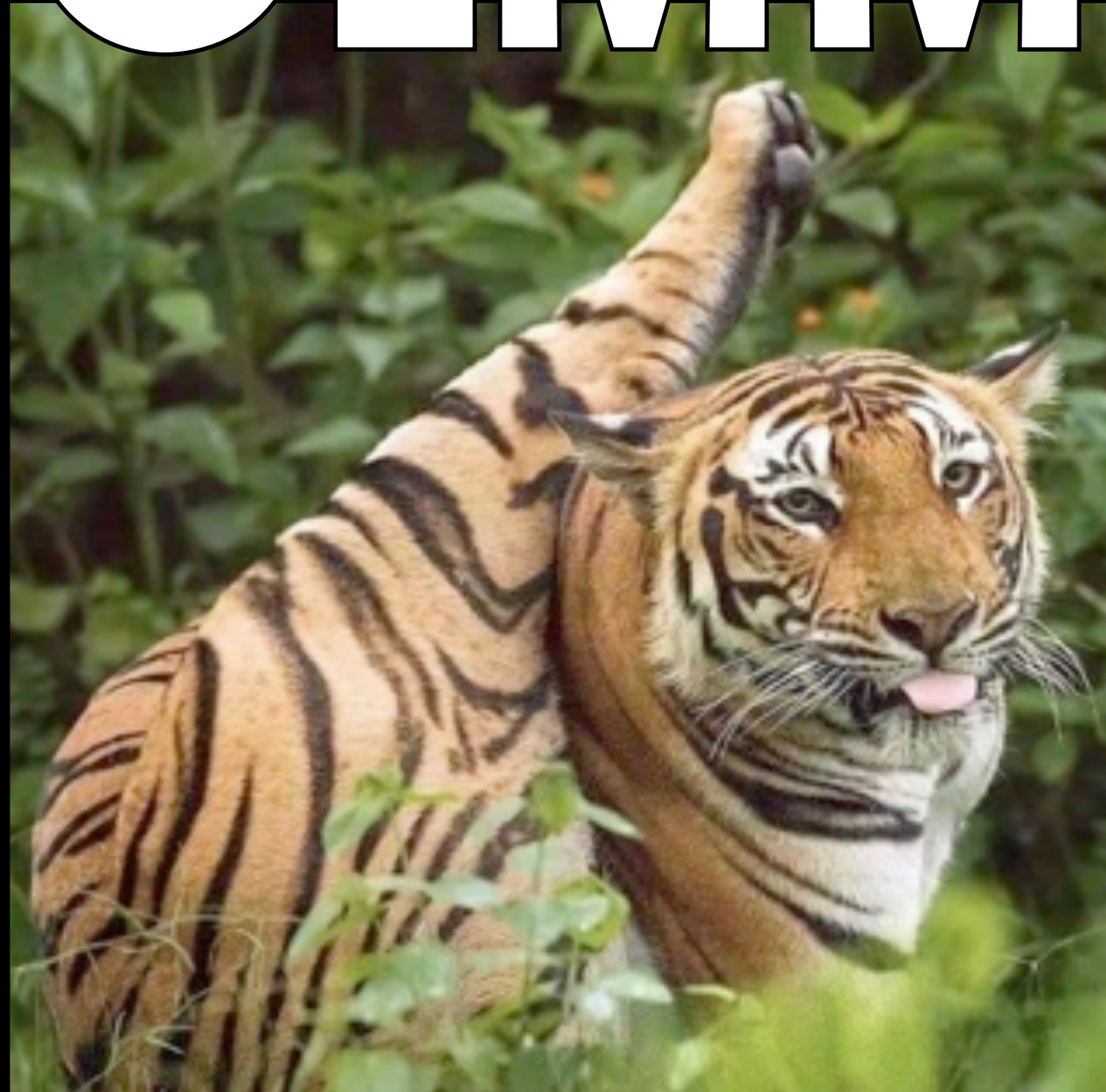


GLM



GLMM



More than regression

- Sometimes the causal model implies you need a regression model
- But this is not true in general
- Need a general approach for estimating arbitrary causal models

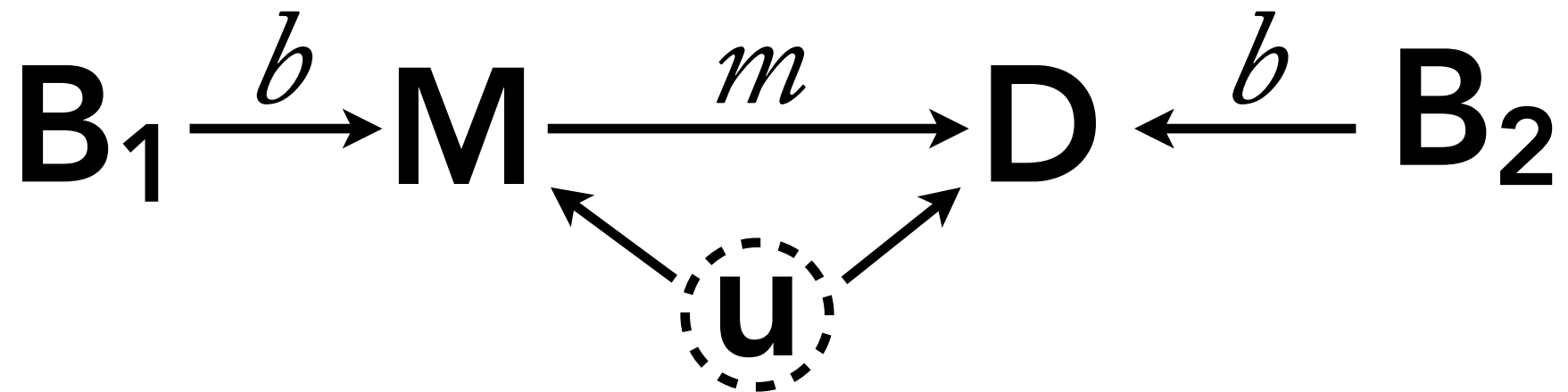
Bayesian Inference

- Bayesian inference is just probability theory
- Given a generative model, Bayesian inference can extract information from data
- If there is no information in the data, it will show
- The generative model is neither Bayesian nor anything else

Bayesian Causal Inference

- It can be hard to analyze an arbitrary causal model
- Full Luxury Bayesian Inference™
 - Program the structural causal model as a Bayesian network
 - Let probability theory analyze the model for you
- “Luxury” is not free: Computation can be challenging; analyzing the model yourself explains things that Bayes will not speak

Bayesian Moms



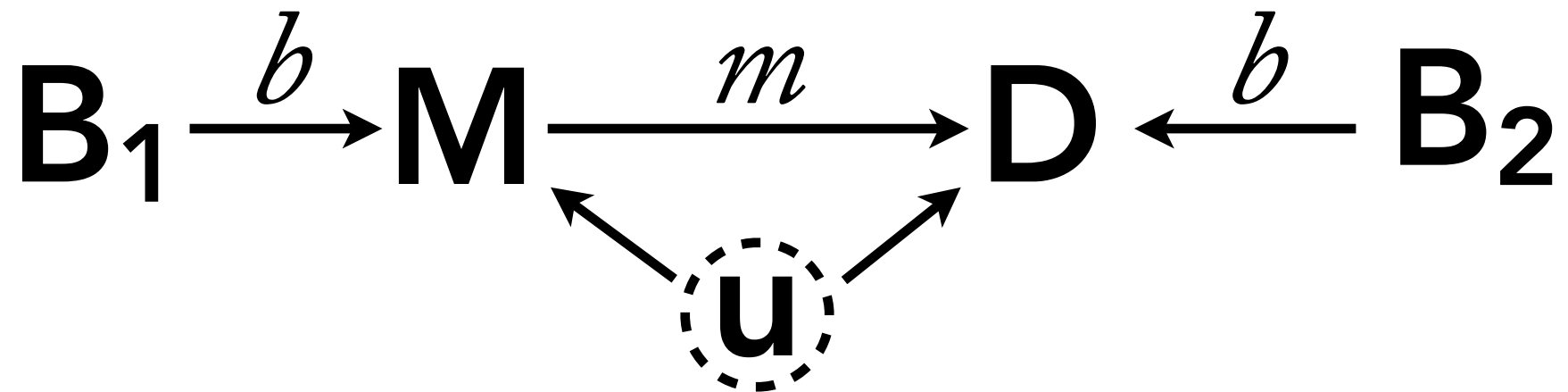
$$M \sim f_M(B_1, u)$$

$$D \sim f_D(M, B_2, u)$$

$$B_1, B_2 \sim f_B(\cdot)$$

$$u \sim f_u(\cdot)$$

Bayesian Moms



```
# non model
|   M ~ normal( mu , sigma ),
|   mu <- a1 + b*B1 + k*U[i],
# daughter model
|   D ~ normal( nu , tau ),
|   nu <- a2 + b*B2 + m*M + k*U[i],
# B1 and B2
|   B1 ~ bernoulli(p),
|   B2 ~ bernoulli(p),
# unmeasured confound
|   vector[N]:U ~ normal(0,1),
# priors
|   c(a1,a2,b,m) ~ normal( 0 , 0.5 ),
|   c(k,sigma,tau) ~ exponential( 1 ),
|   p ~ beta(2,2)
```

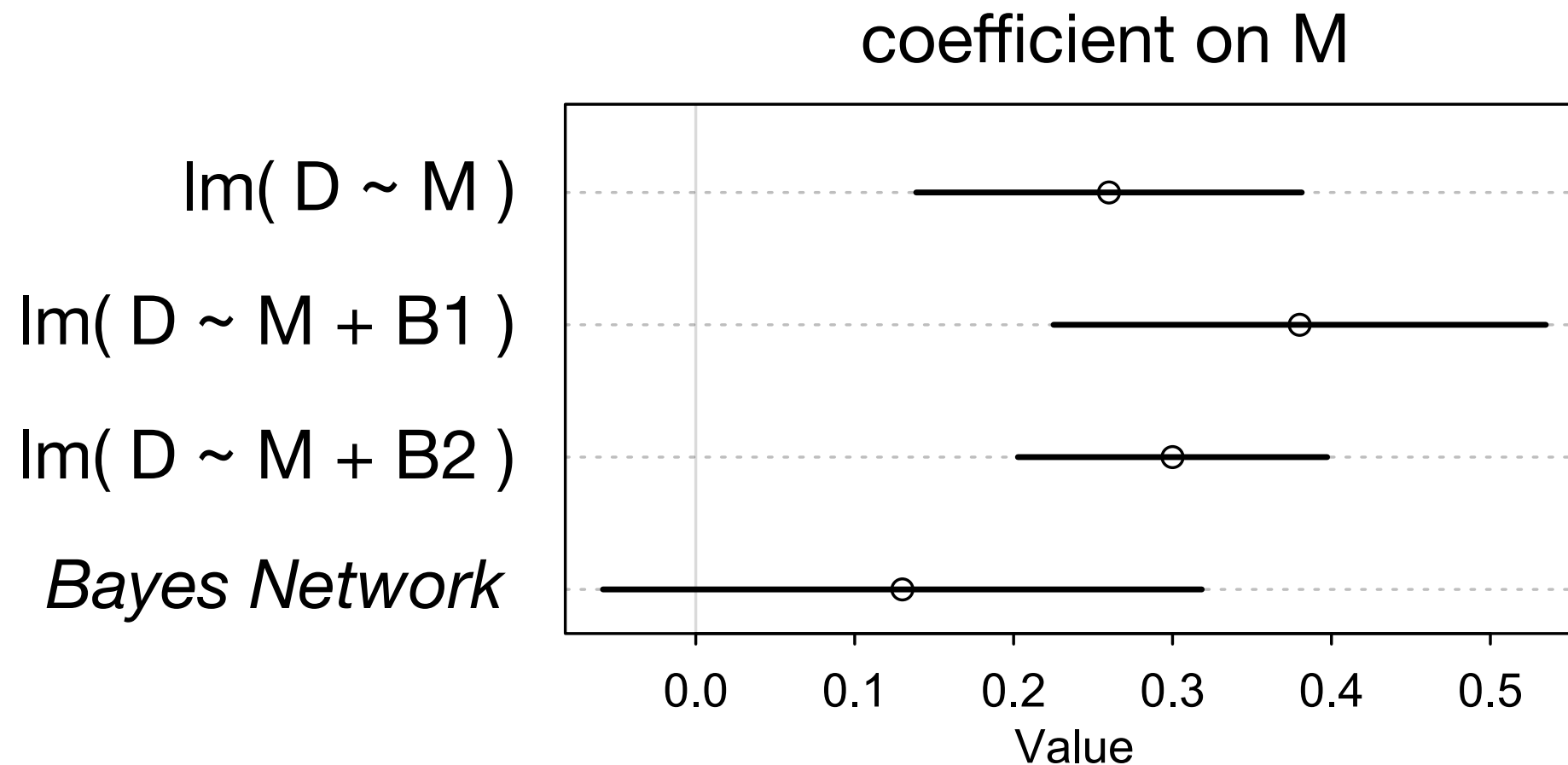
$$M \sim f_M(B_1, u)$$

$$D \sim f_D(M, B_2, u)$$

$$B_1, B_2 \sim f_B(.)$$

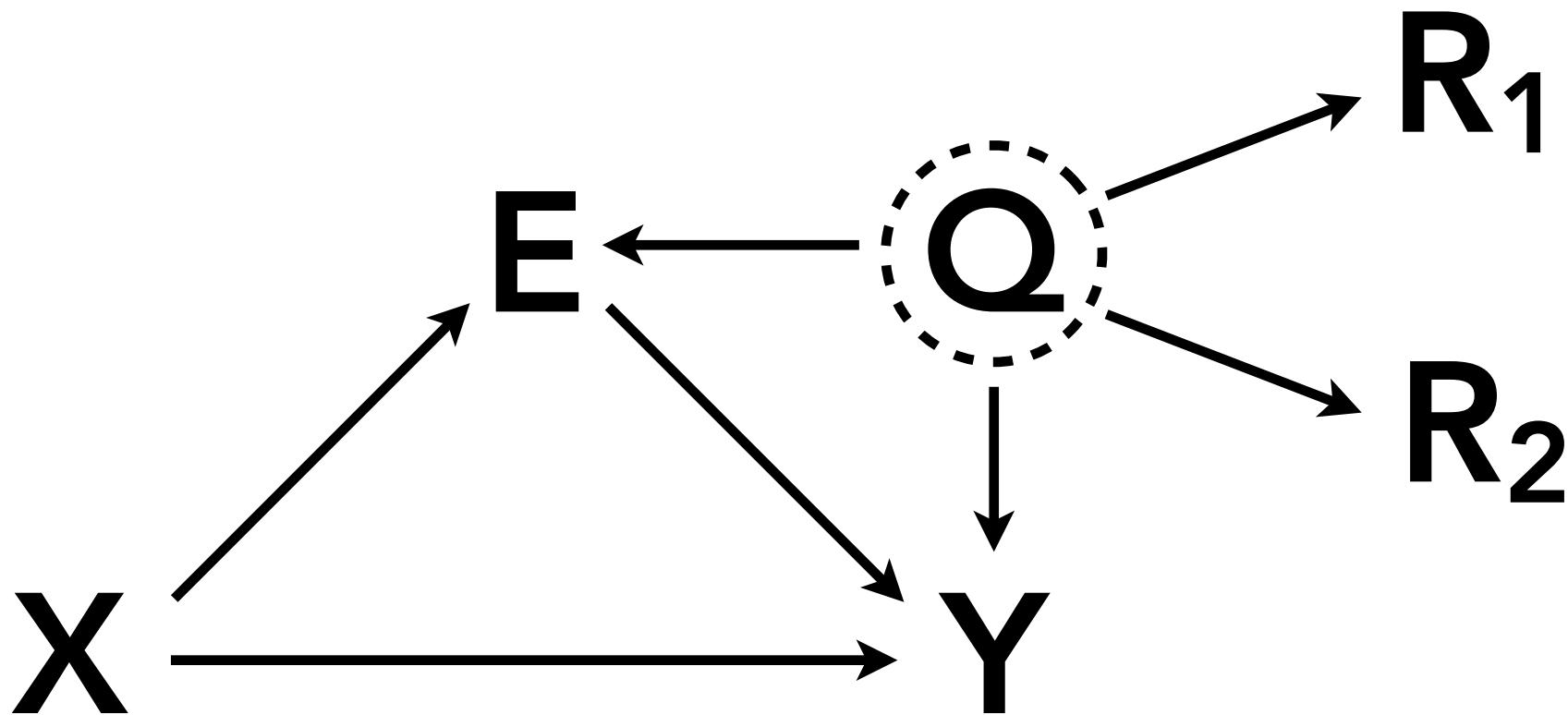
$$u \sim f_u(.)$$

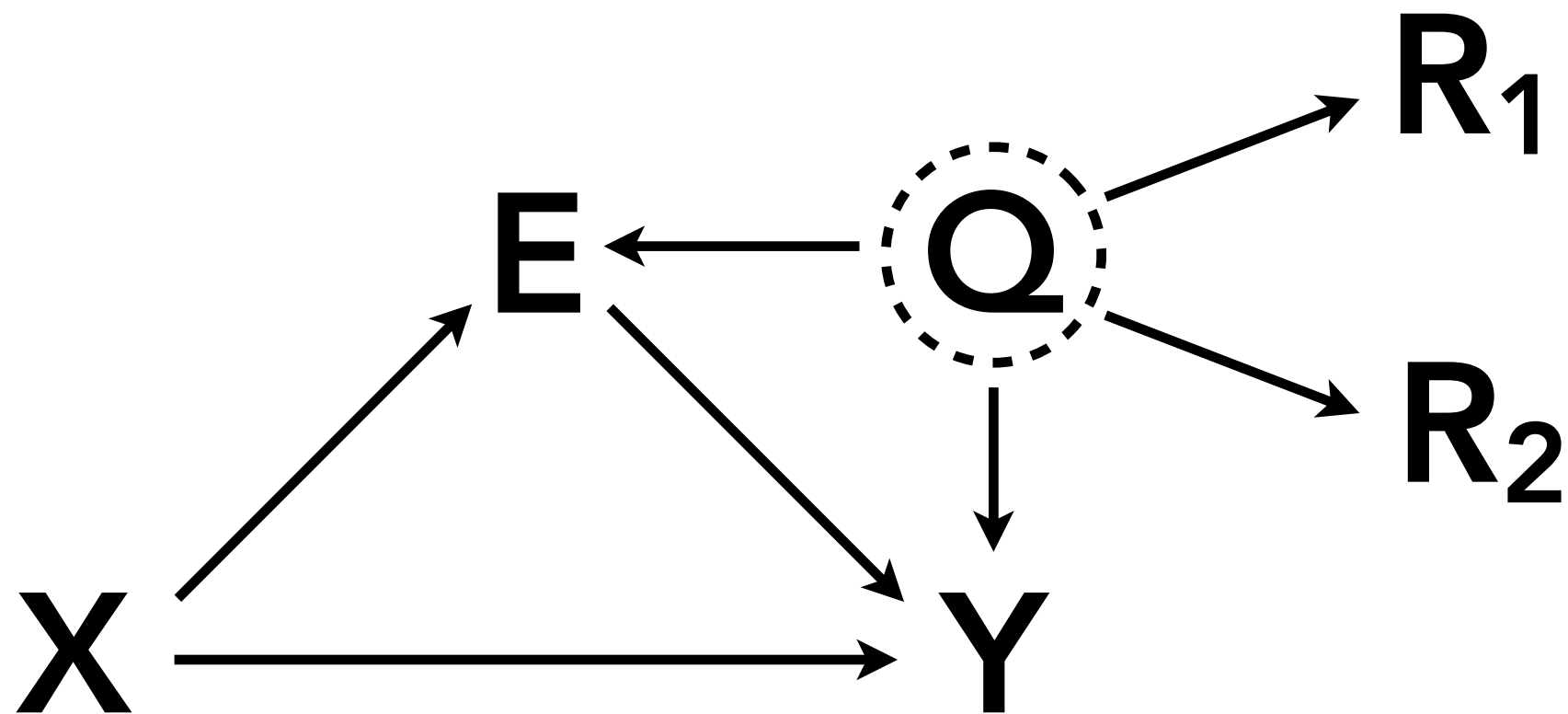
Bayesian Moms



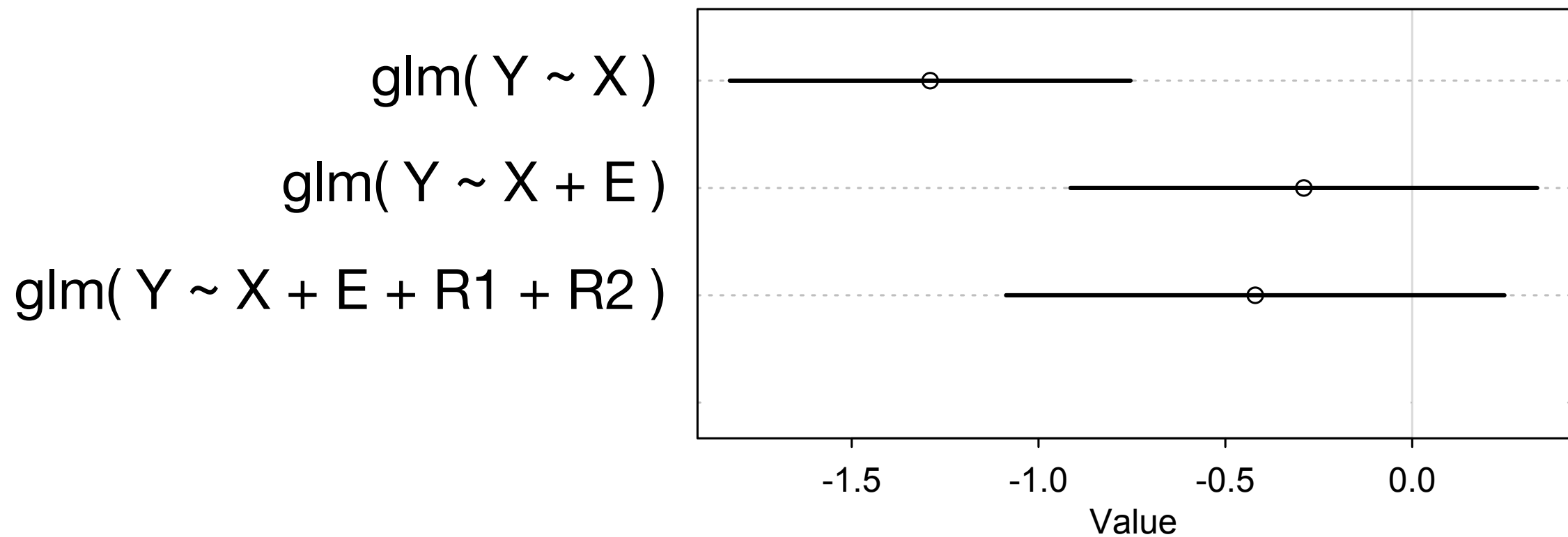
Peer Bias

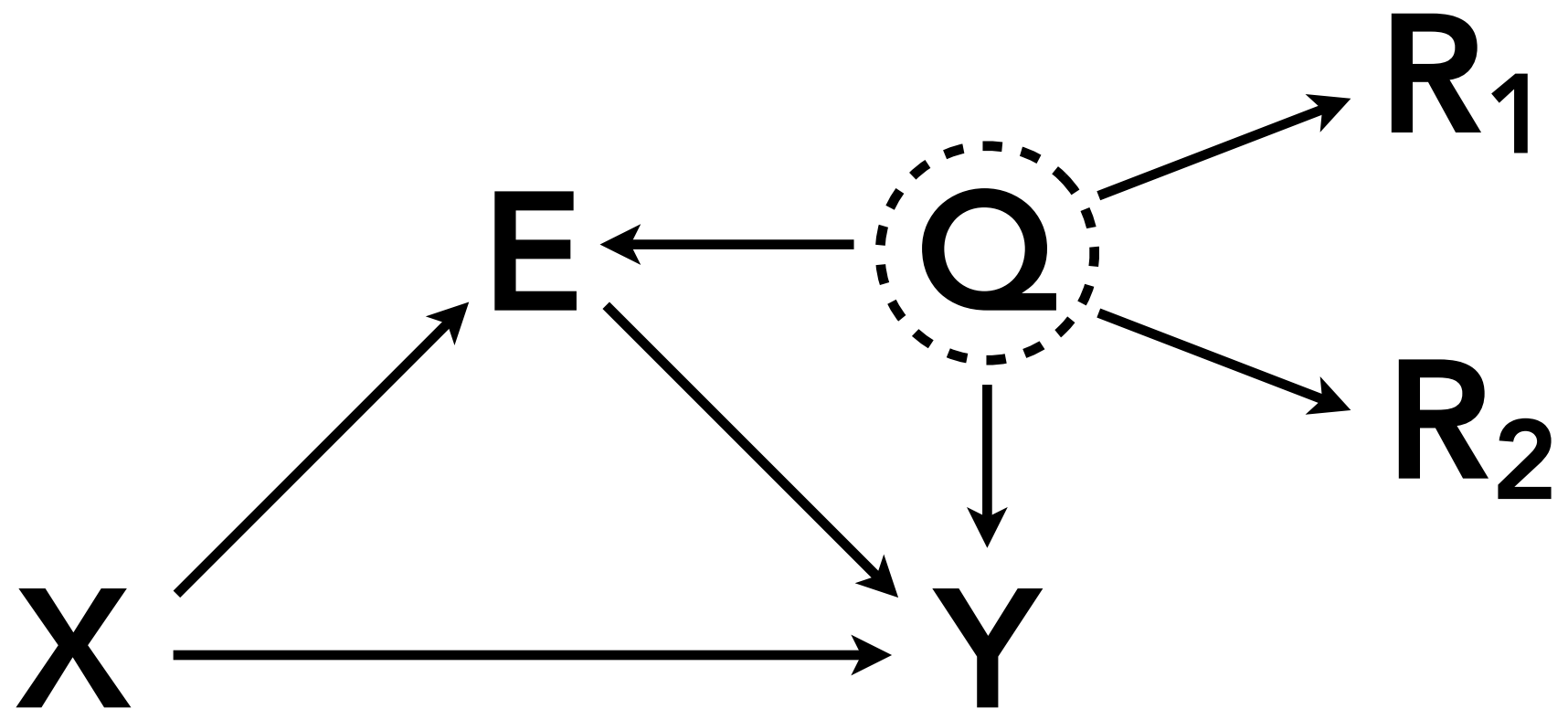
- We don't know \mathbf{Q} , but what if we have some proxies of it?



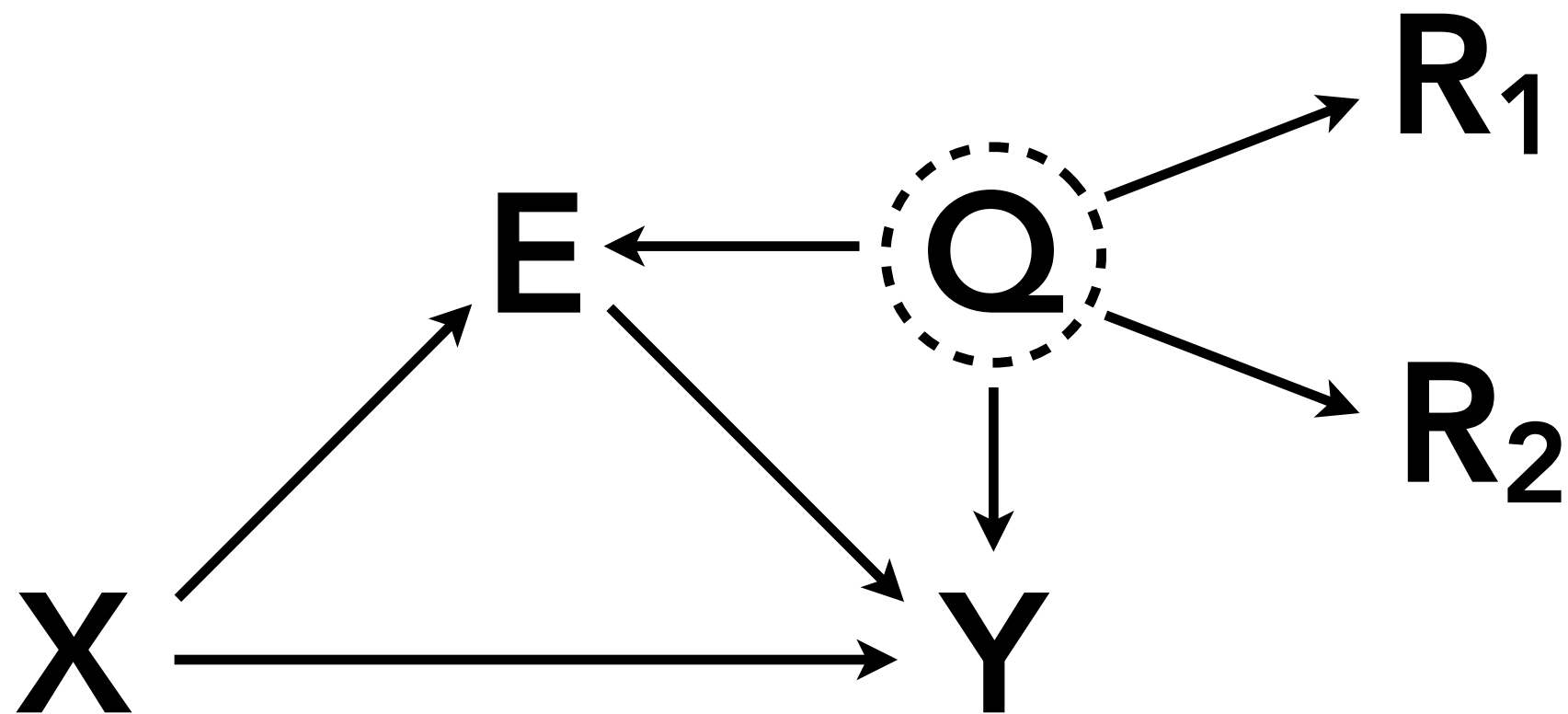


coefficient on X

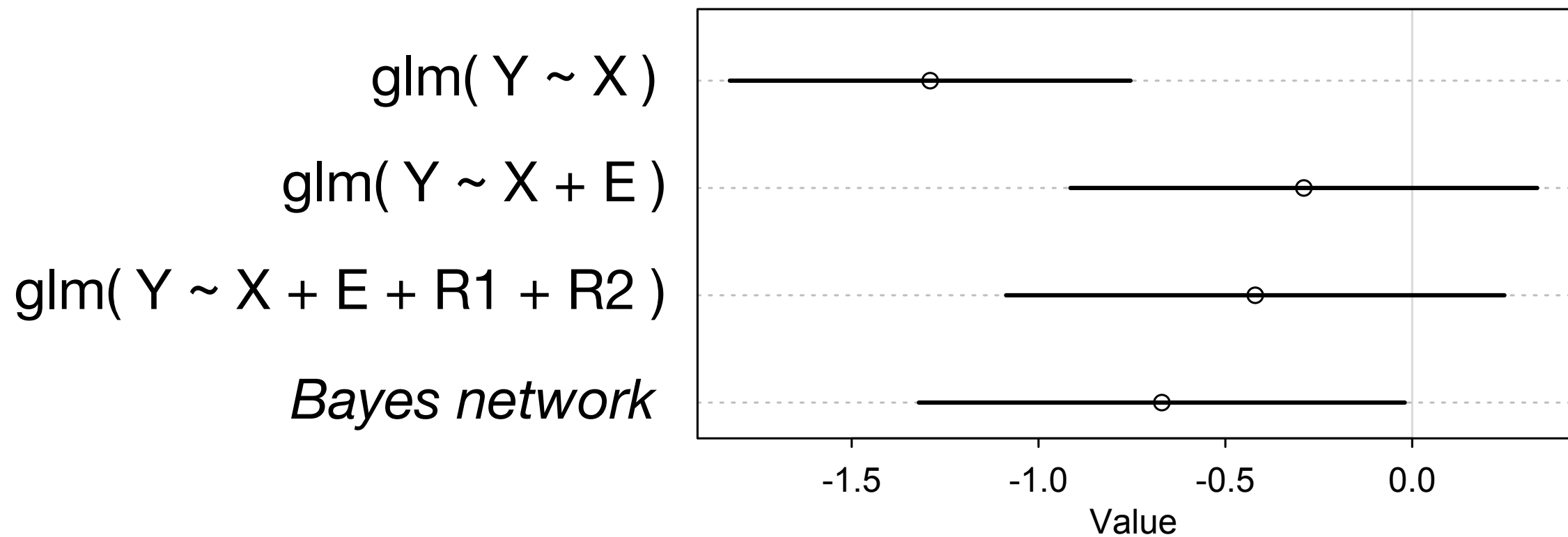




```
# Y model
Y ~ bernoulli(p),
logit(p) <- a[E] + X*XX + h*Q[id],
a[E] ~ normal(0,1),
X ~ normal(0,1),
h ~ half_normal(0,1),
# Q model
vector[id]:Q ~ normal(0,1),
R1 ~ normal(Q,1),
R2 ~ normal(Q,1)
```



coefficient on X



Not Magic

- Bayesian inference is only as good as
 - The generative model fed into it
 - The sample fed into it
 - The numerical algorithm
- But conditional on the model, you cannot do better
- Provides ways to handle missing data, measurement error — **add cause** of missingness and error to model
- Things like **social networks** & **phylogenies** are never “observed”. Must be **inferred** from often bad data. Need a causal model of observations.

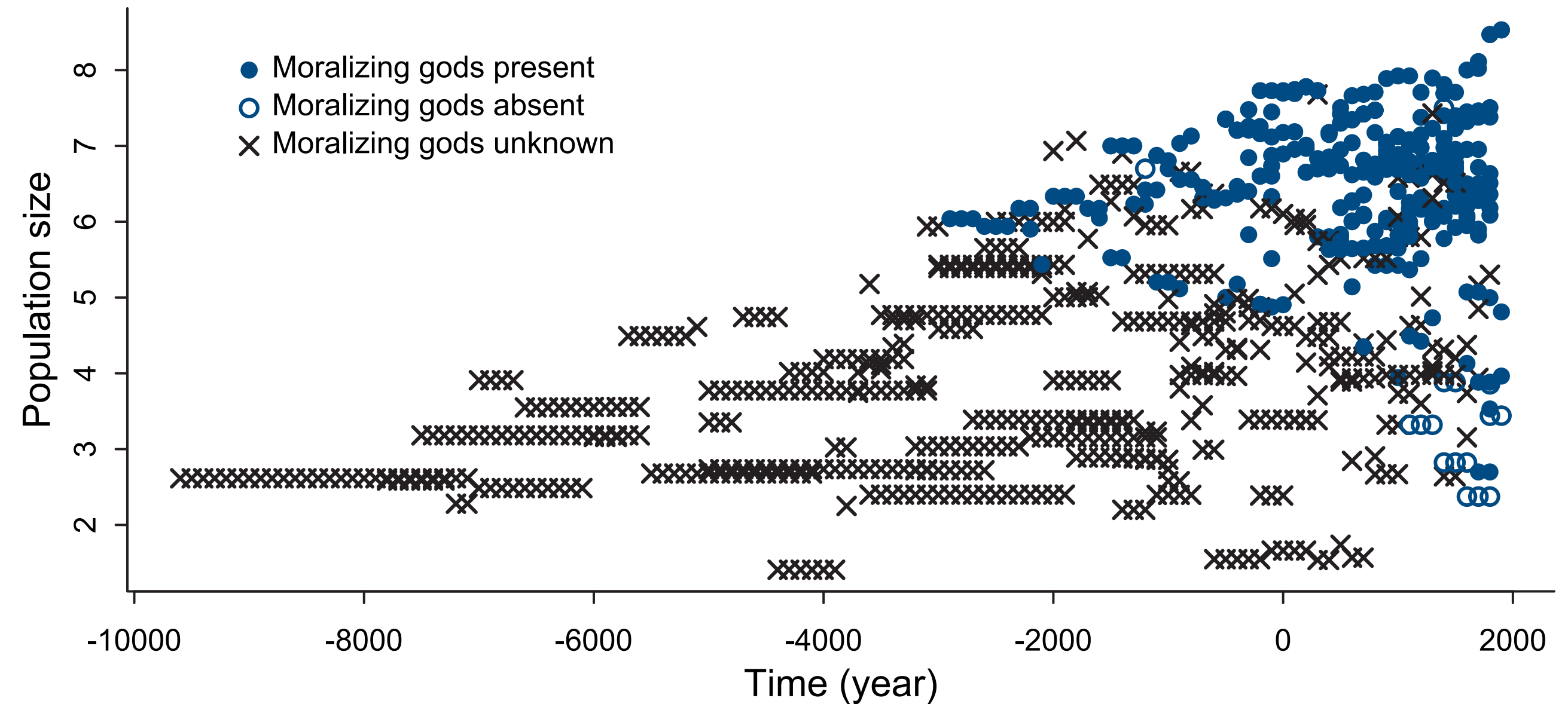


FIGURE 15.7. Missing values in the `Moralizing_gods` data. The blue points, both open and filled, are observed values for the presence of beliefs about moralizing gods. The `x` symbols are unknowns, the missing values.

Science Before Statistics

CAUSAL INFERENCE



Science Before Statistics

- Causal inference requires
 - A causal model distinct from any statistical models
 - Analysis of the implications of the causal model for (1) research design (2) testing (3) estimation strategy
 - Some way to perform estimation
- Descriptive and experimental research no exceptions

Much much more

- Computation of treatment effects: not a single parameter
- Post-stratification: effect for the population, not the sample
- Partial-identification: confounded but learning
- Research design

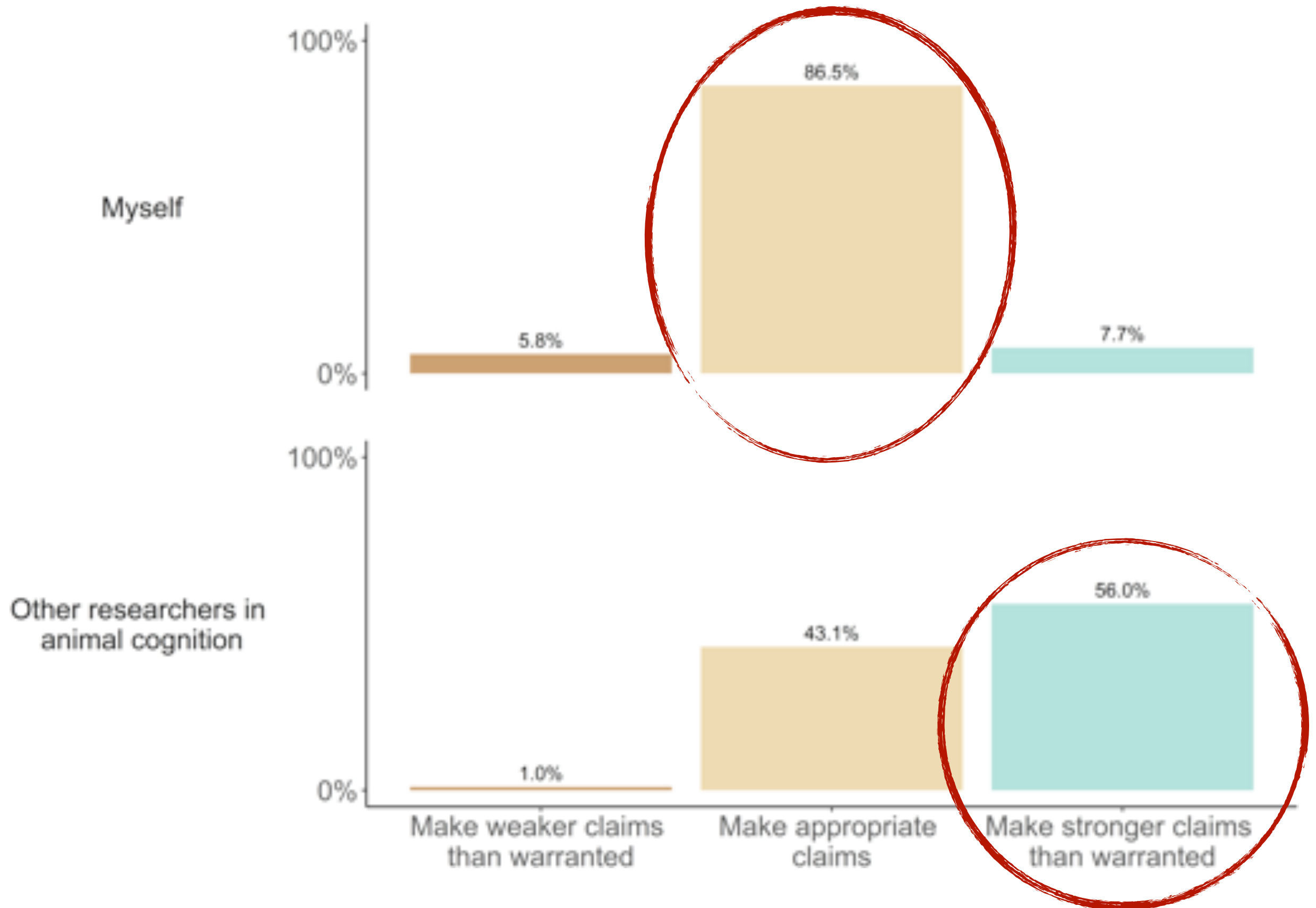
Full Luxury Causal Workflow

1. Derive candidate causal model using “science”
2. Program model as a generative simulation
3. Design research & validate statistical analysis using (2)
4. Confront model with data; celebrate wins and losses equally
5. Revise and repeat

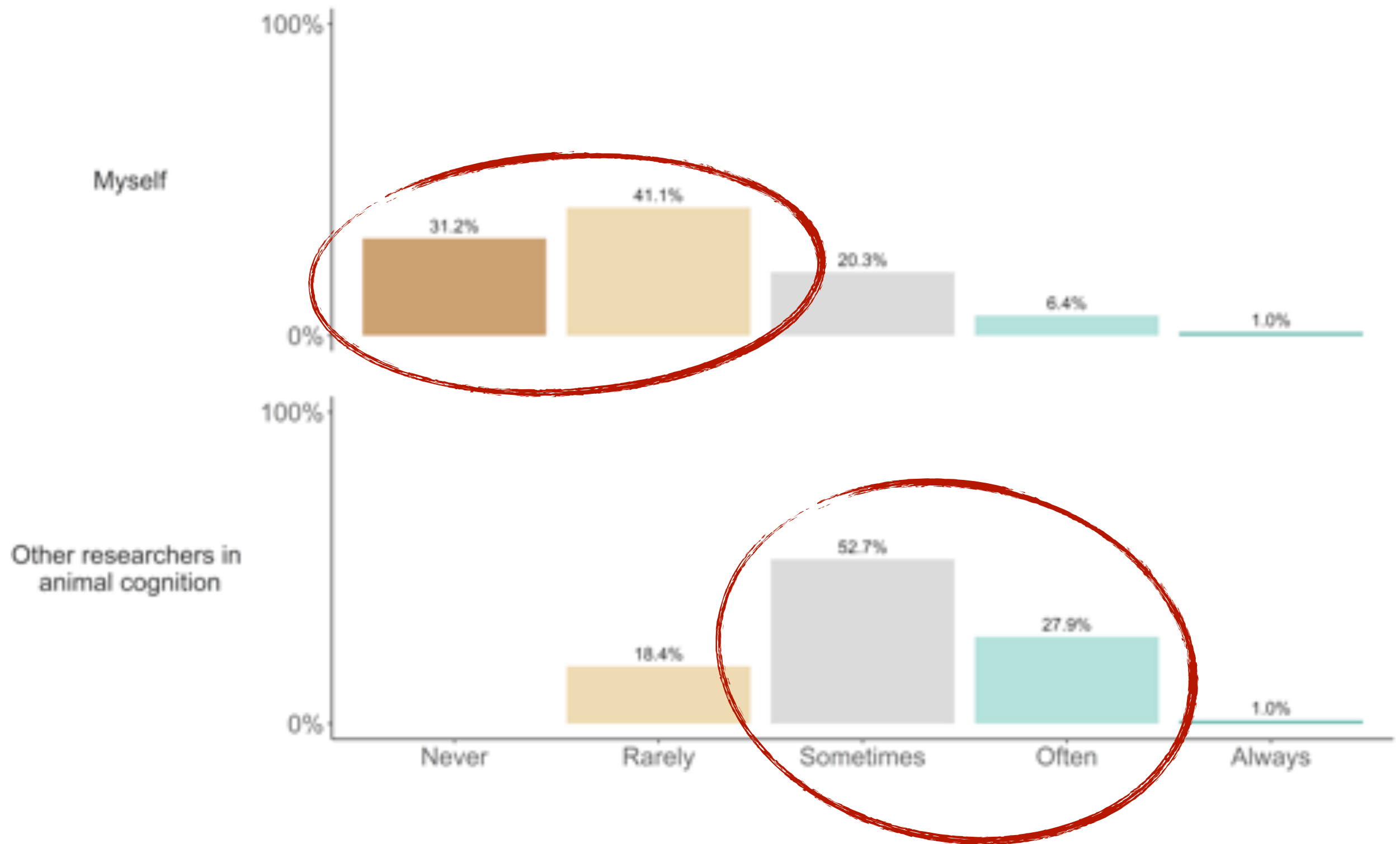
Full Sadness Non-Causal Workflow

1. Find or collect some variables that are **conceptually**, but not **logically**, relevant to phenomenon
2. Probe the data anyway you can to reveal **asterisks**
3. Tell a hopeful, **causal story** about what these asterisks exist
4. Never state the **assumptions** that license your story
5. Revel in your magnificent h-index

When submitting a paper, I/others...



How often are QRPs performed by?



Much much more

