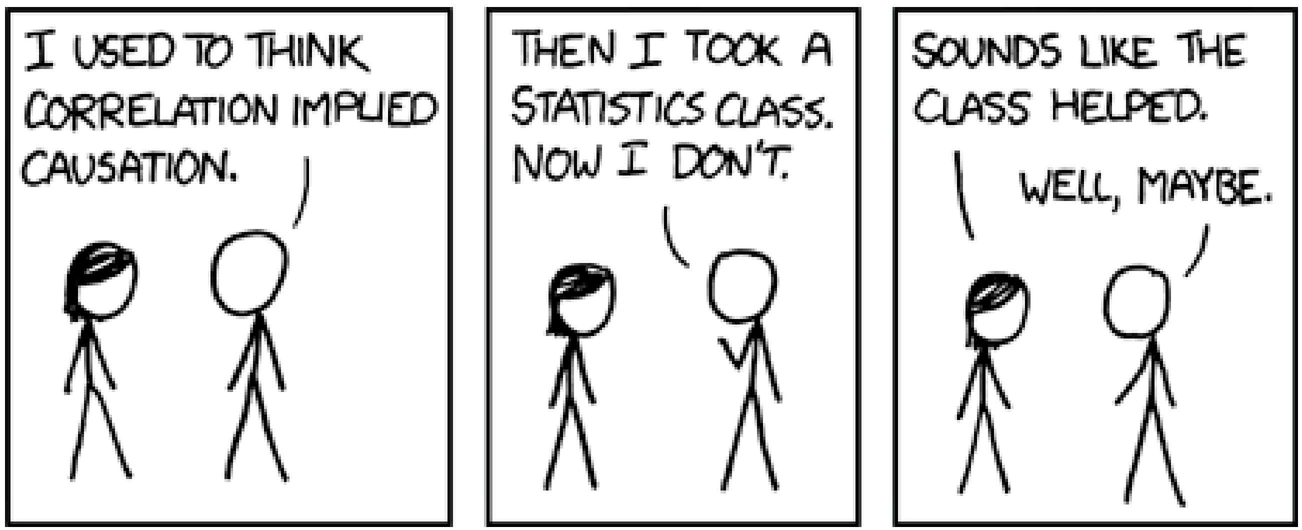
3. What is a causal model?

“Correlation does not equal causation… but where there’s smoke, there’s fire.“

-Jim Grace


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Overview

1. Inferring causality

2. An Example from Coral Reefs

3.1 Inferring Causality

Why?

3.1 Causality. Correlation

Why does the effect happen?
What is the cause?

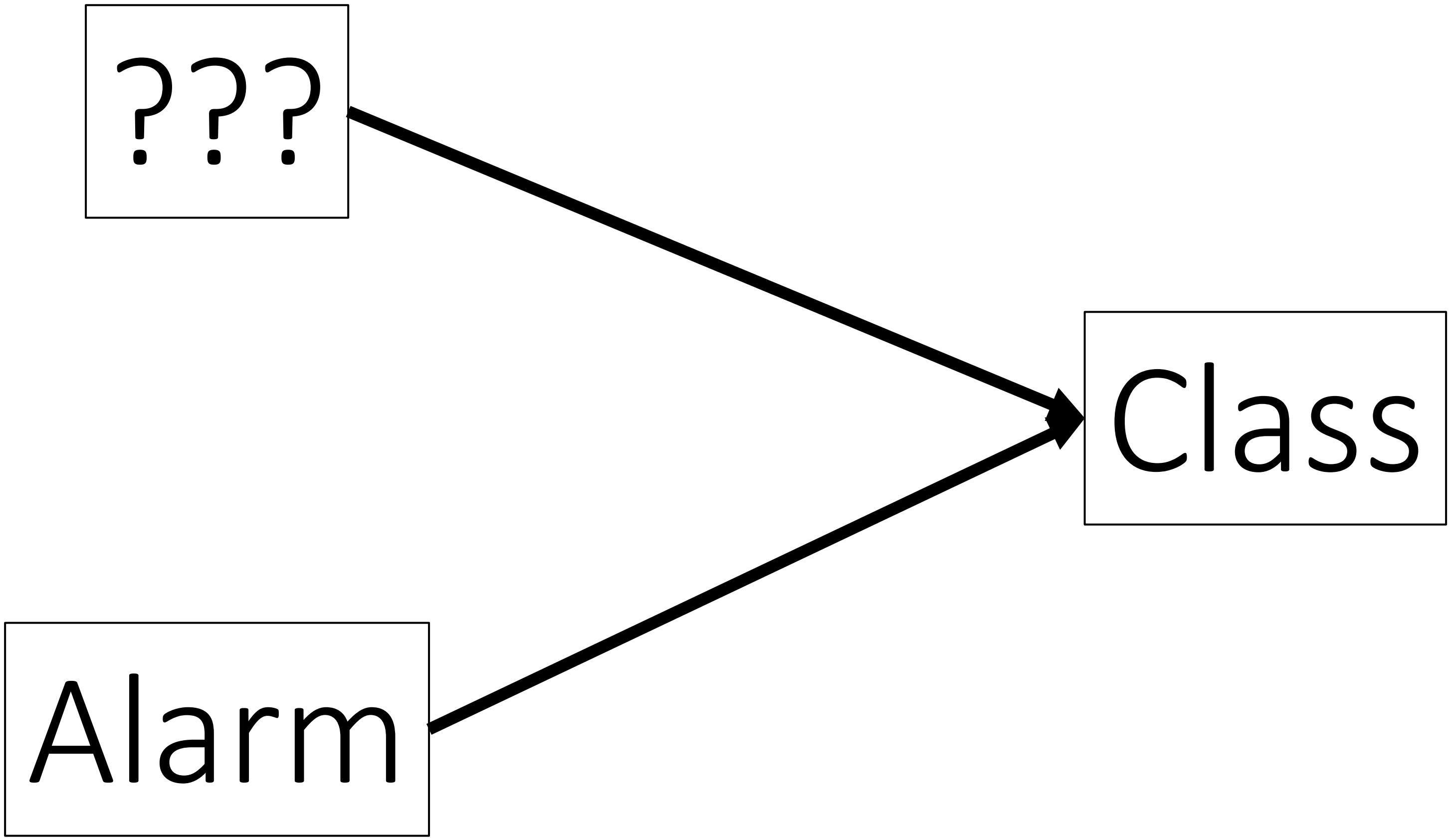
3.1 Causality. An example

Imagine you are perpetually late to this class…why?

What time the alarm goes off sets what time you arrive to class..

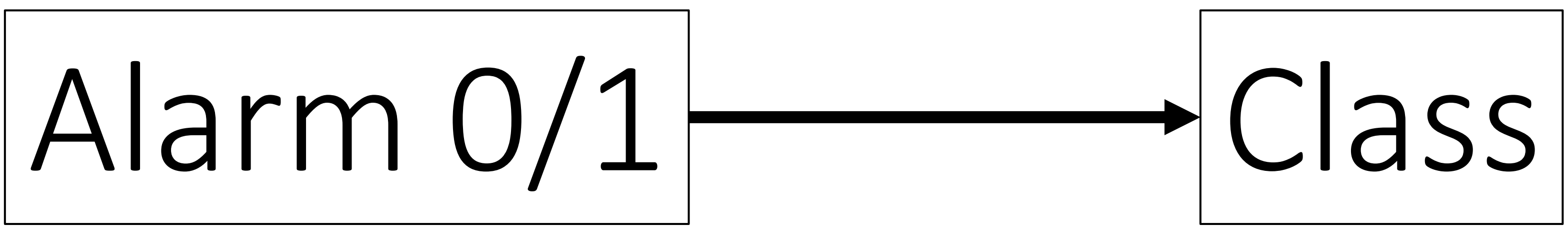
3.1 Causality. An example

A lot of other things can determine what time you arrive for class



<u>How do we determine what out of the entire *universe* of possible causes us to be late for class??</u>

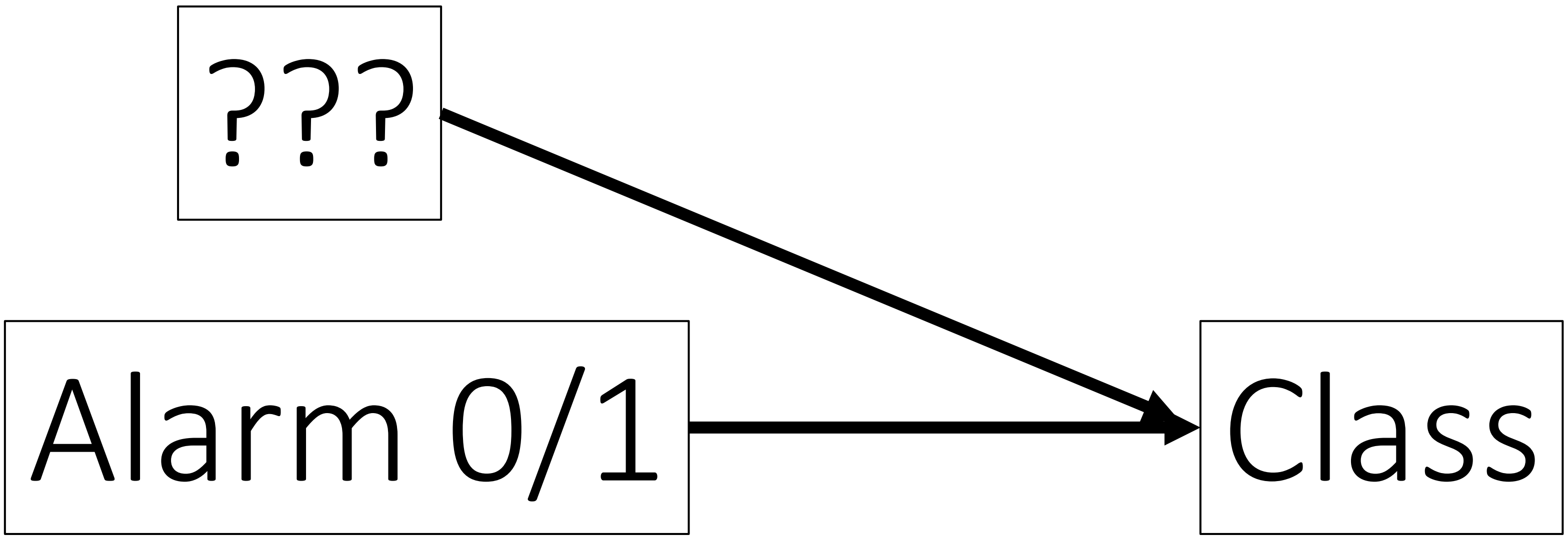
3.1 Causality. Solution #1: Perform an experiment



Try not setting the alarm and see if you are late…
(*not recommended)

Or try setting the alarm for different times…

3.1 Causality. Solution #1: Perform an experiment

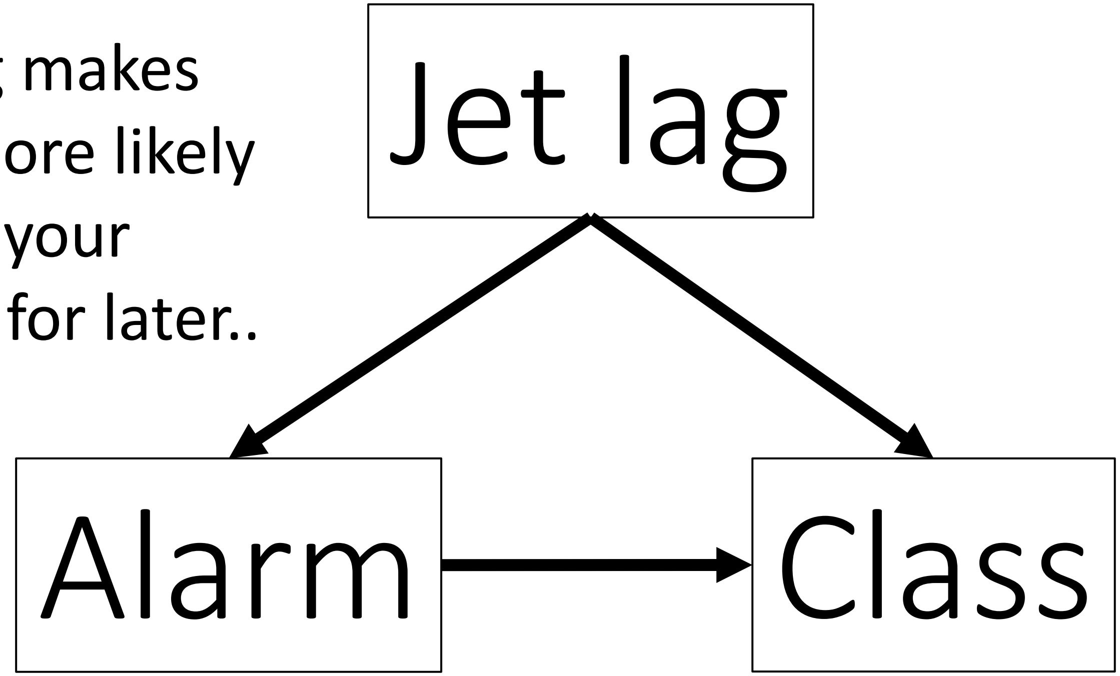


But you can't control the universe and there are still lots of things that can vary from day to day that can make you late… (aka, traffic, weather, "dog ate my homework")

3.1 Causality. Solution #2: Close the back-door

Imagine there is something that affects both what time you set your alarm, and what time I schedule class…

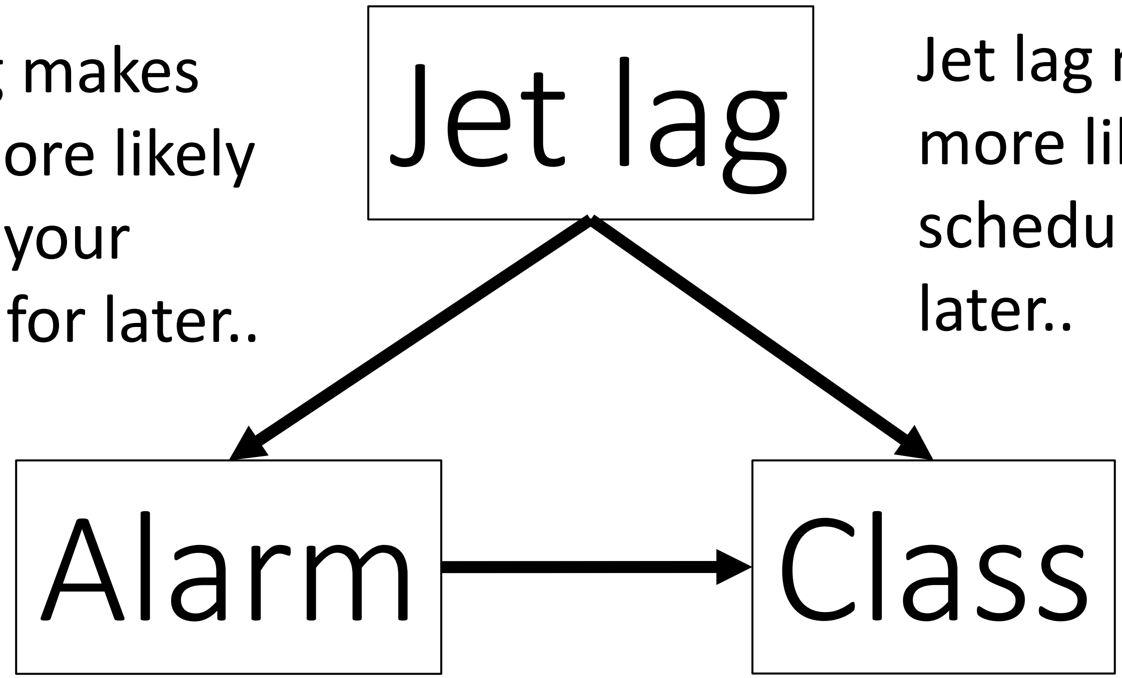
Jet lag makes you more likely to set your alarm for later..



3.1 Causality. Solution #2: Close the back-door

Imagine there is something that affects both what time you set your alarm, and what time I schedule class…

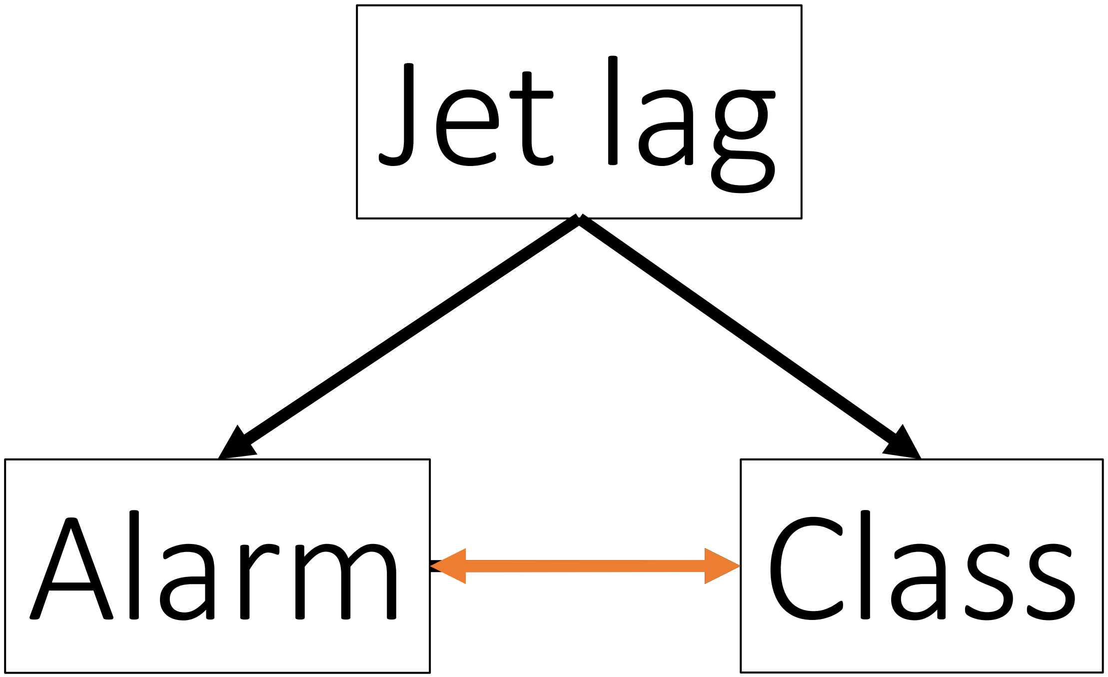
Jet lag makes you more likely to set your alarm for later..



Jet lag makes me more likely to schedule class later..

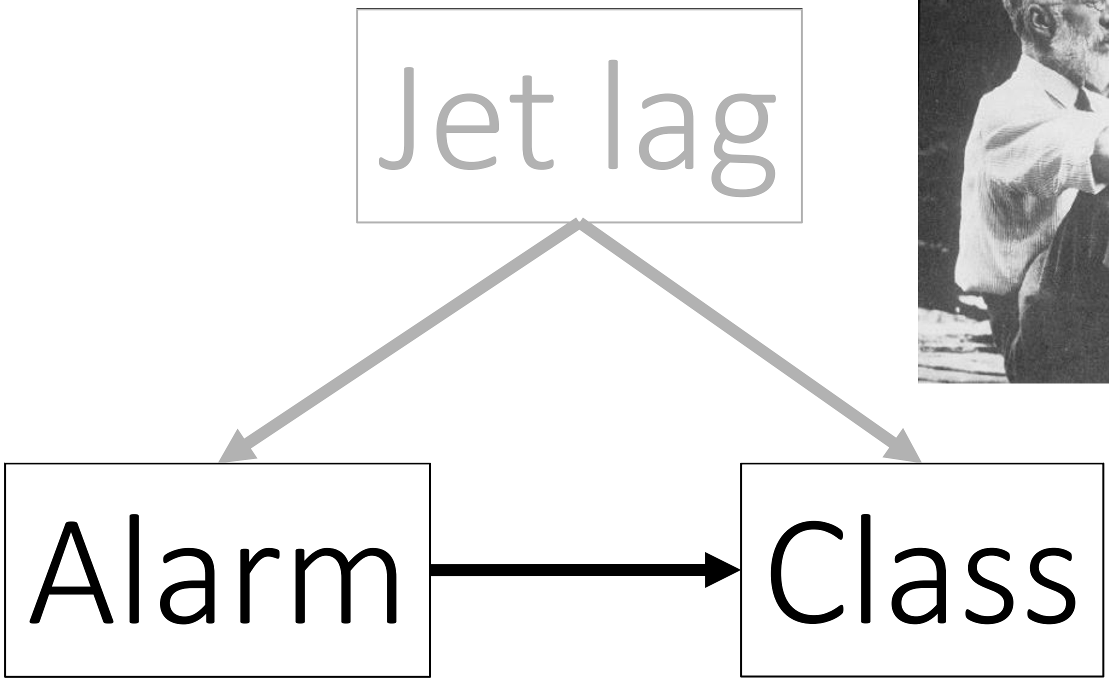
3.1 Causality. Solution #2: Close the back-door

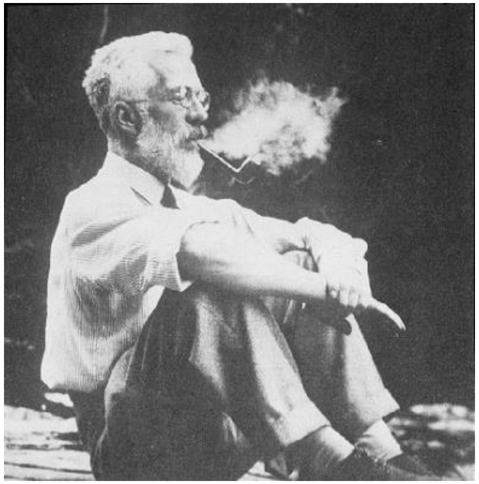
A “common cause” could generate an apparent correlation…



3.1 Causality. Solution #2: Close the back-door

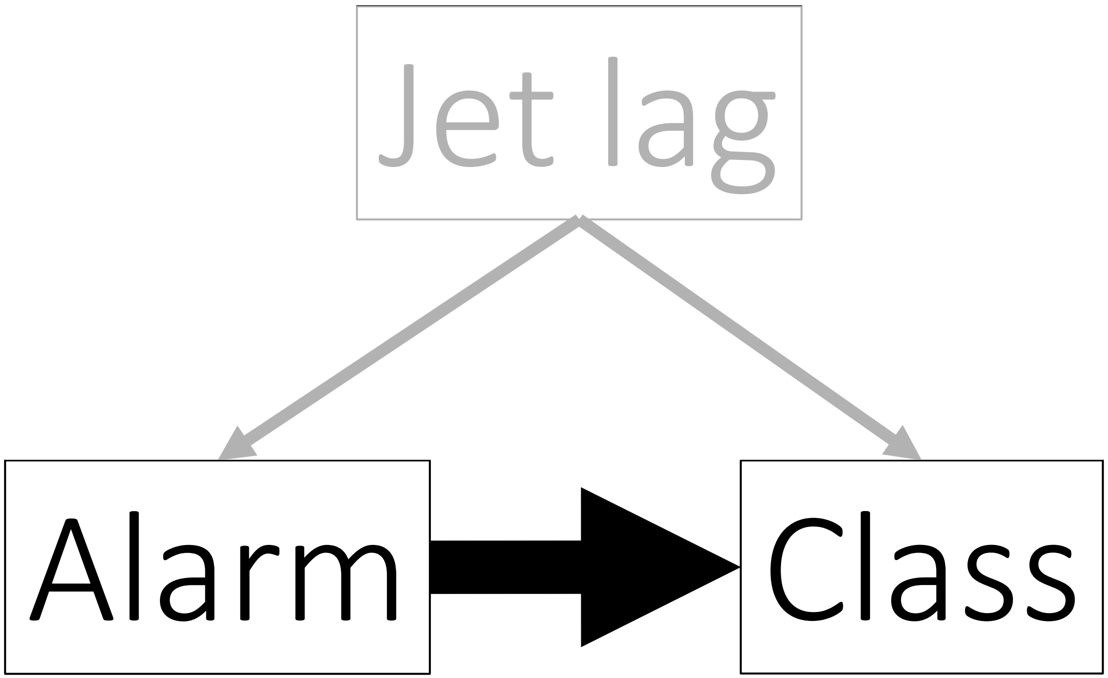
A “common cause” could generate an apparent correlation… that could *mistakenly* be interpreted as causal in its absence





3.1 Causality. Solution #2: Close the back-door

Furthermore, the estimates of the relationship are also biased! Alarm seems *more* important in getting to class on time because its incorporating some of the information from jet lag

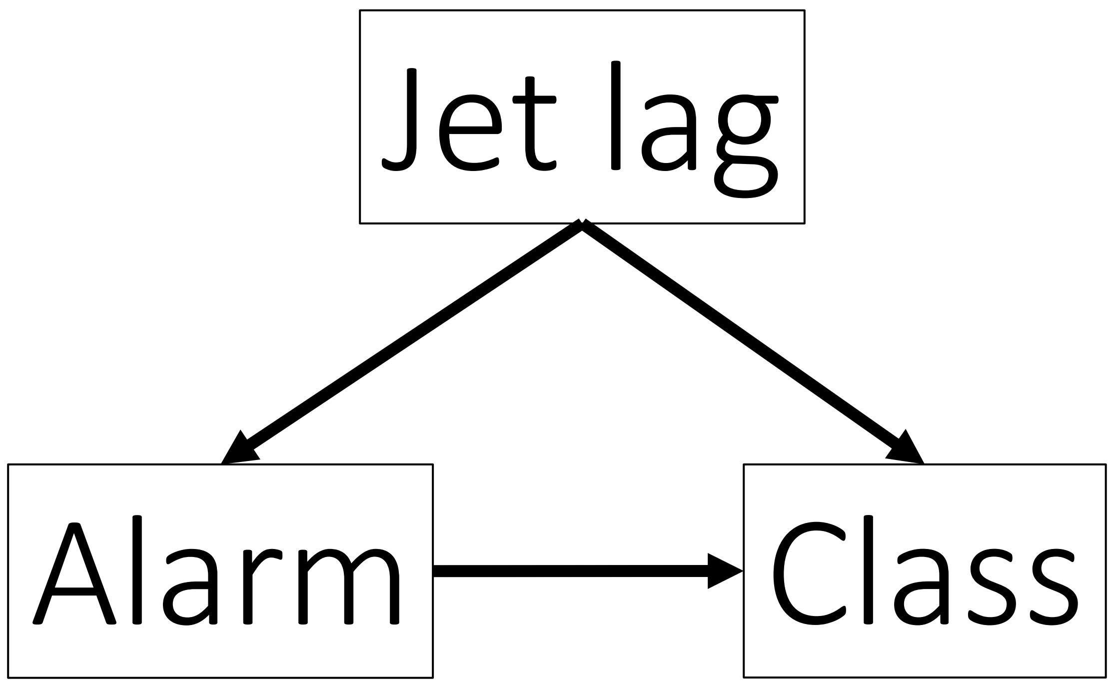


3.1 Causality. Solution #2: Close the back-door

```
# create effect y based on cause
set.seed(1)

y <- rnorm(100)

cause <- y + runif(100, 0, 5)

# set x to also correlated with cause
x <- cause + runif(100, 0, 5)

# apparent correlation
cor(y, x)

# significant effect of x
summary(lm(y ~ x))

# significant effect of cause
summary(lm(y ~ x + cause)) # true relationship
```

3.1 Causality. Solution #2: Close the back-door

But including the common cause in the model isolates the potentially causative effect… this is known as *<u>“shutting the back door”</u>*



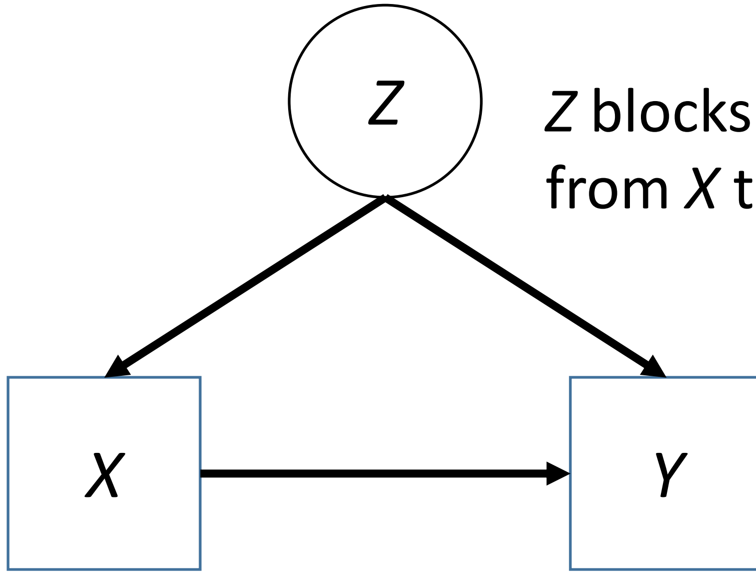
3.1 Causality. Solution #2: Close the back-door

If we want to know the causal effect of *X* on *Y* and have some set of variables *Z*, then *Z* satisfies the *back-door criterion* if:

1. *Z* blocks every path from *X* to *Y* that has an arrow *into* *X* – makes sure you condition on a confounder

2. No part of *Z* is a descendent of *X* (there is no path from *X* to any of the variables in *Z*) – stops from conditioning on the effect of the cause

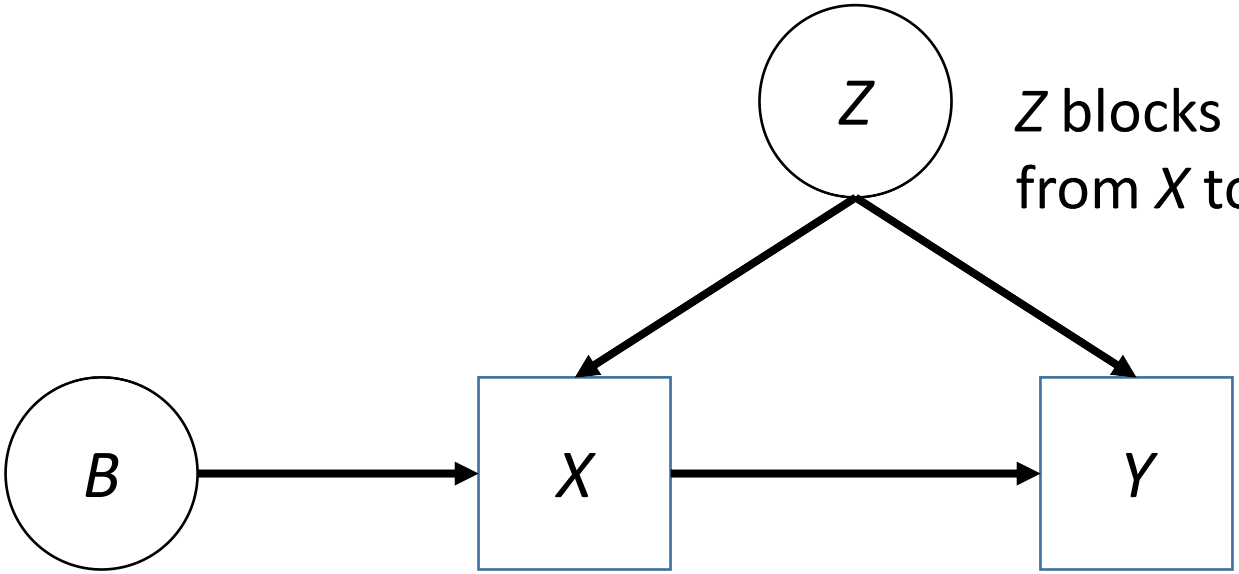
3.1 Causality. Solution #2: Close the back-door



Z blocks back-door path from *X* to *Y* (condition 1)

Z is not a descendent of *X* (condition 2)

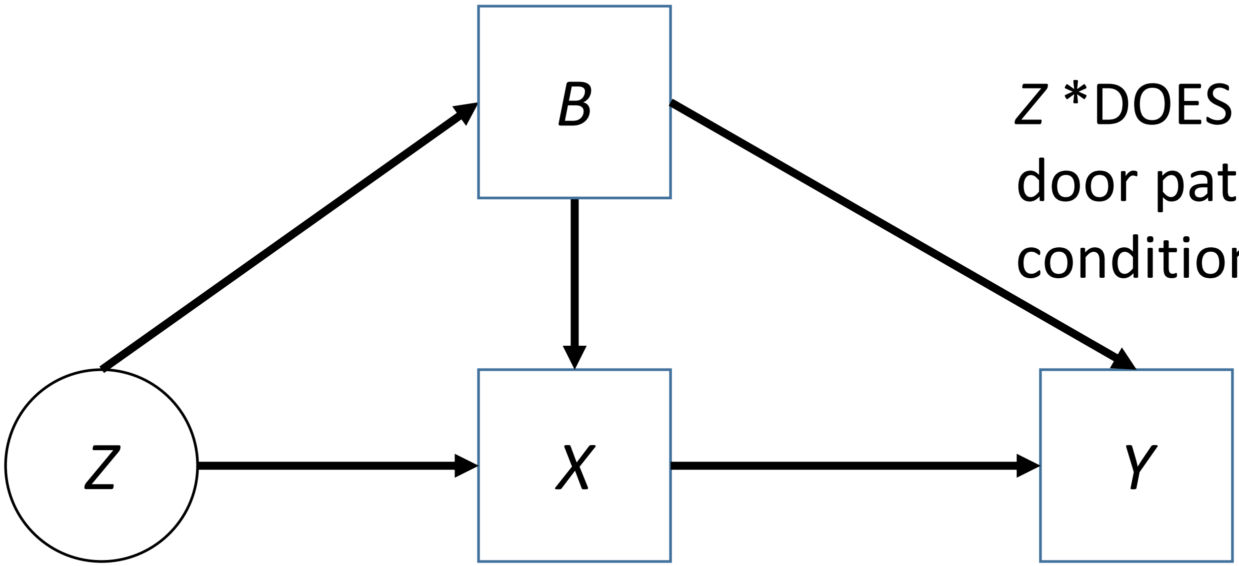
3.1 Causality. Solution #2: Close the back-door



Z blocks back-door path from *X* to *Y* (condition 1)

Z nor *B* is not a descendent of *X* (condition 2)

3.1 Causality. Solution #2: Close the back-door



Z *DOES not* block back-door path from *X* to *Y* (FAIL condition 1)

3.1 Causality. Solution #2: Close the back-door

- Paths do not imply *ultimate* causality

- Very rarely does an effect have a *single* cause or a *single* confounder

- Just have to include the *biggest* confounders
  - There will always be unexplained variability in the response
  - Need to minimize but not eliminate this error

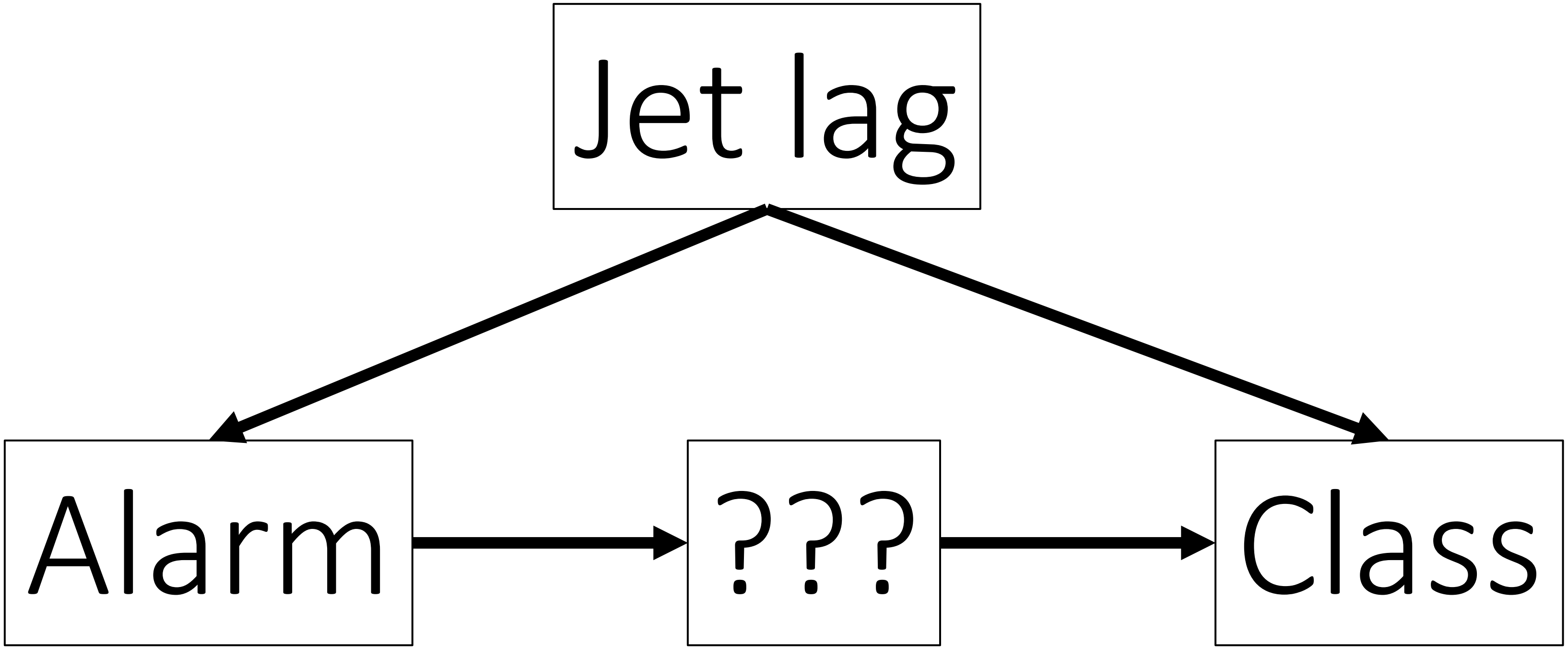
- Don’t forget, this is a *hypothesis* –we can always test it later with more data!

3.1 Causality. Solution #2: Close the back-door

- How do choose confounders?
  - Theory
  - Previous measurements
  - Statistical tests

3.1 Causality. Solution #3. Close the front-door

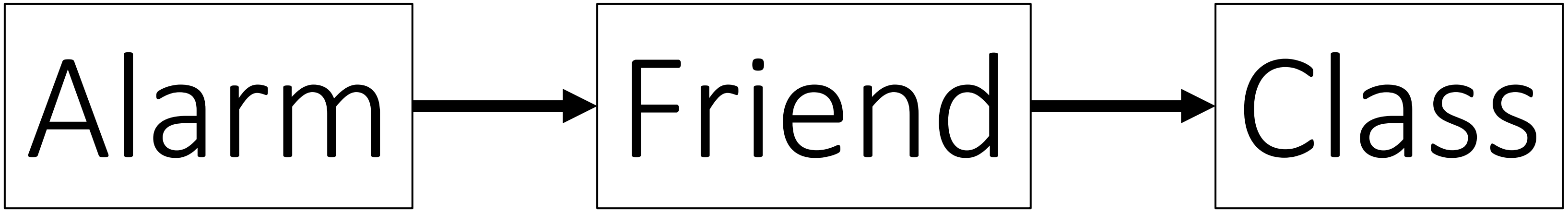
What if there were something else that completely mediated this relationship?



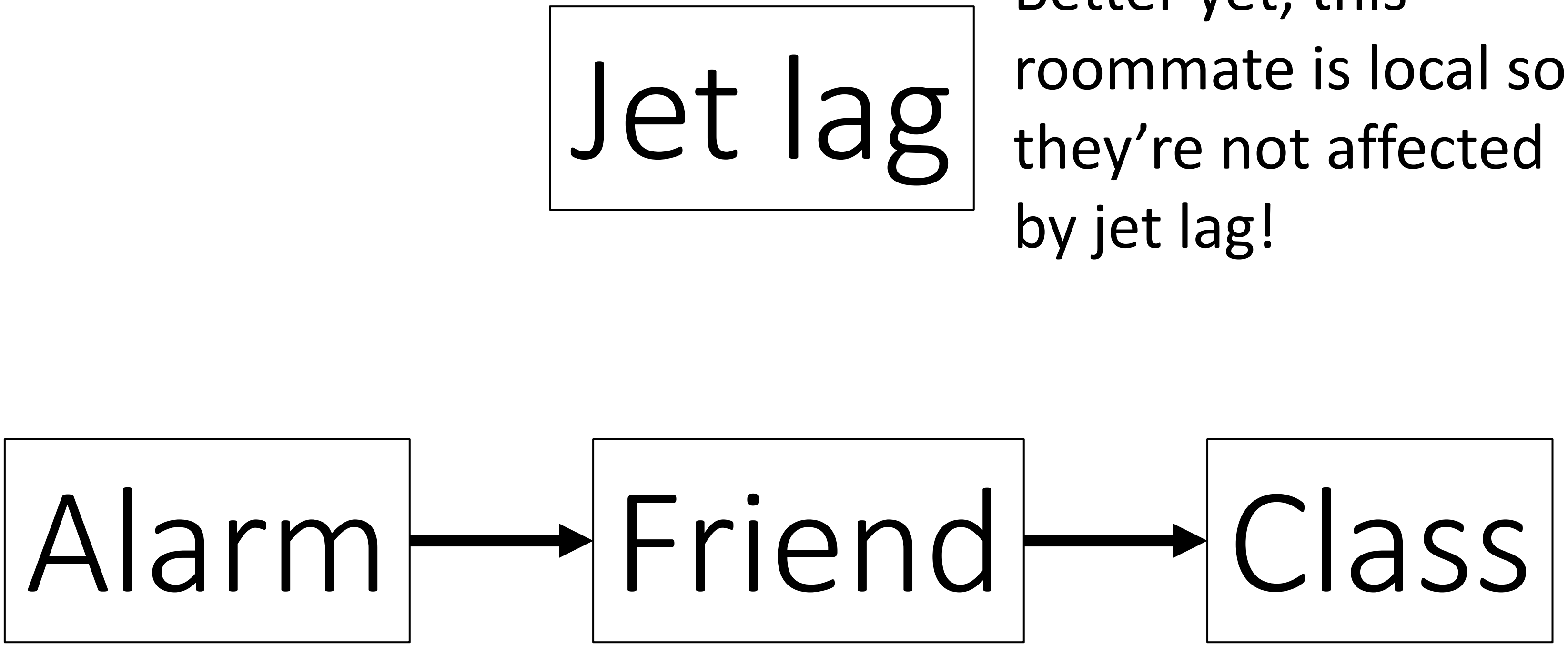
3.1 Causality. Solution #3. Close the front-door

Say you had a roommate who sometimes hears your alarm when it goes off early…and wakes you up so you're not late for class!

If it goes off too late your roommate might have already left and you get to sleep in

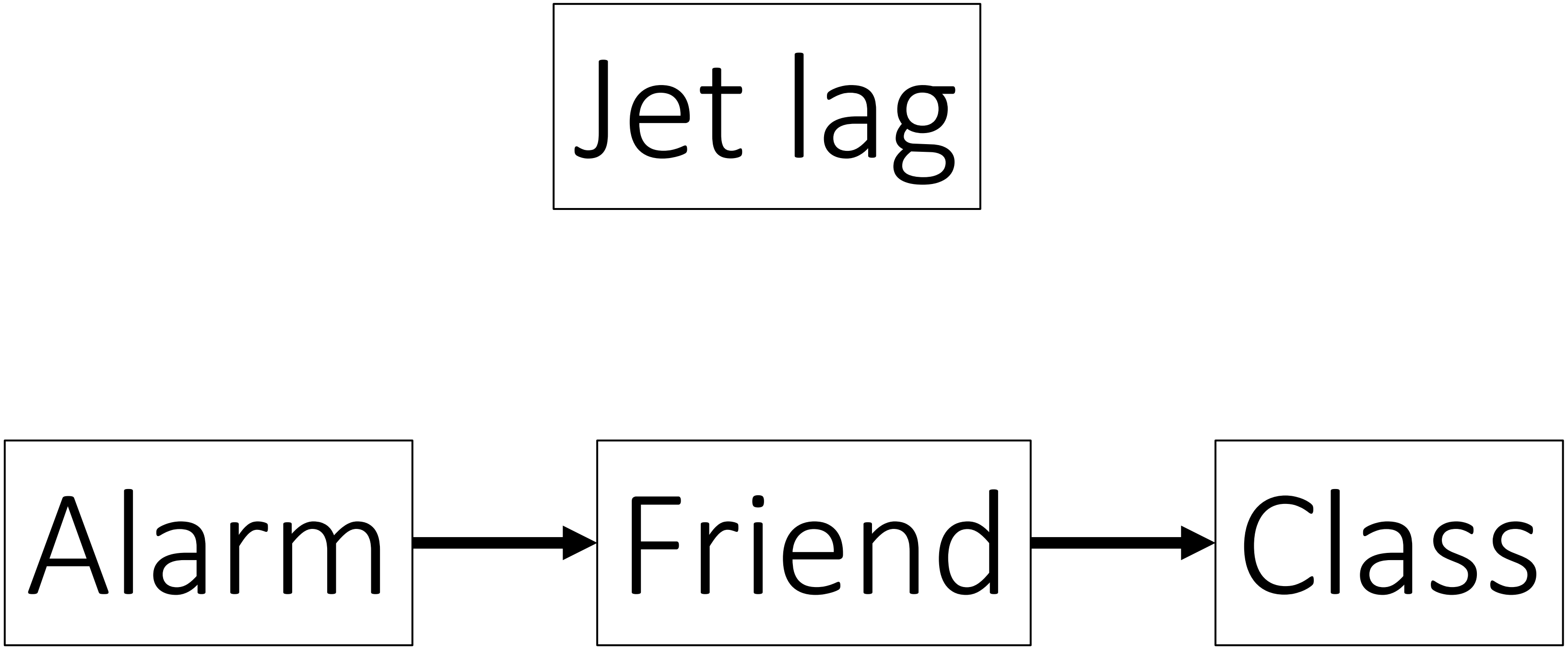


3.1 Causality. Solution #3. Close the front-door



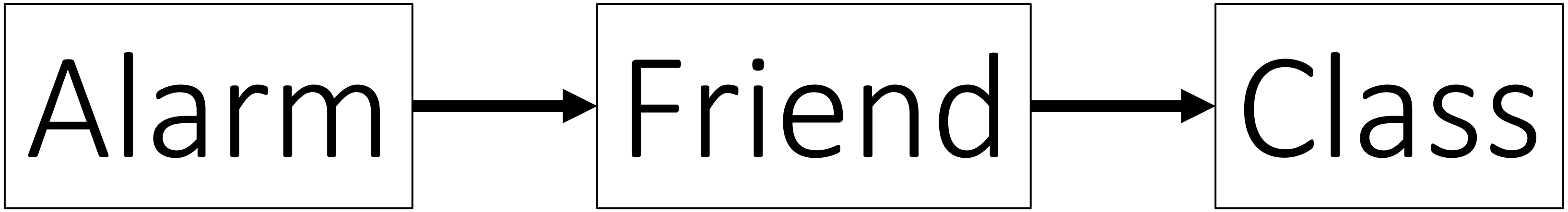
Better yet, this roommate is local so they're not affected by jet lag!

3.1 Causality. Solution #3. Close the front-door



This friend breaks the chain of causality between alarm and when you arrive for class…this is known as “*shutting the front door*”

3.1 Causality. Solution #3. Close the front-door



Shutting the front door removes the confounding influence of jet lag even when its not included

3.1 Causality. Solution #3. Close the front-door

If we want to know the causal effect of *X* on *Y* and have a mediator *S*, *S* satisfies the *front-door criterion* if:

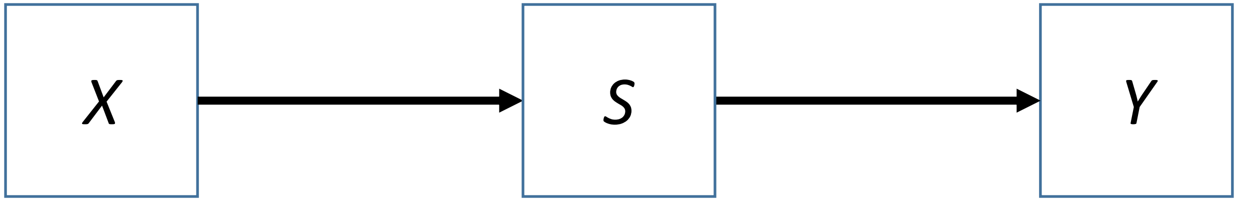
1. *S* blocks all directed paths from *X* to *Y* – intercepts full effects of *X* (breaks the chain)

2. There are no unblocked back-door paths from *X* to *S* – where cause and mediator have hidden influences (i.e., jat lag)

3. *X* blocks all back-door paths from *S* to *Y* – where signal from *X* can sneak through

3.1 Causality. Solution #3. Close the front-door

No backdoor paths from *X* to *S* (condition 2)

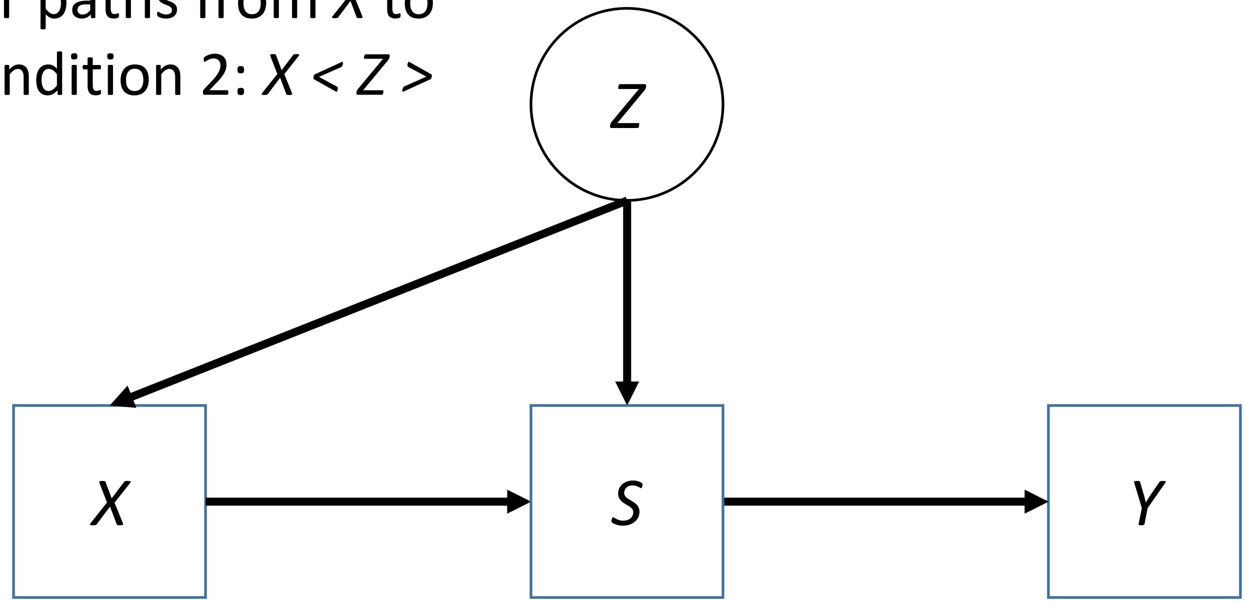


X blocks all backdoor paths from *S* to *Y* (condition 3)

S blocks all directed paths between *X* and *Y* (condition 1)

3.1 Causality. Solution #3. Close the front-door

Back-door paths from *X* to *S* (FAIL condition 2: *X* < *Z* > *S*)



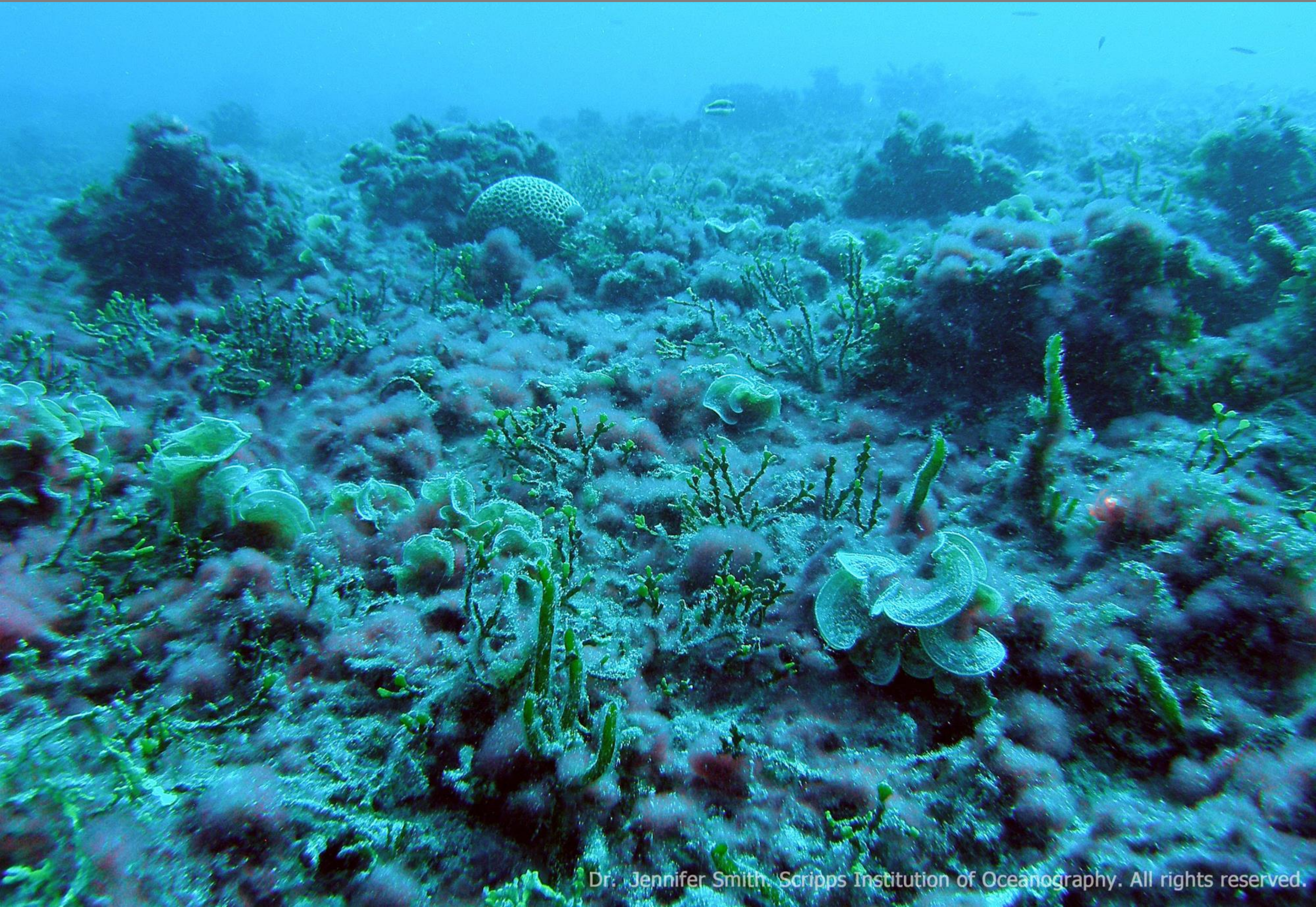
S blocks all directed paths between *X* and *Y* (condition 1)

3.1 Causality. Solution #2: Close the front-door

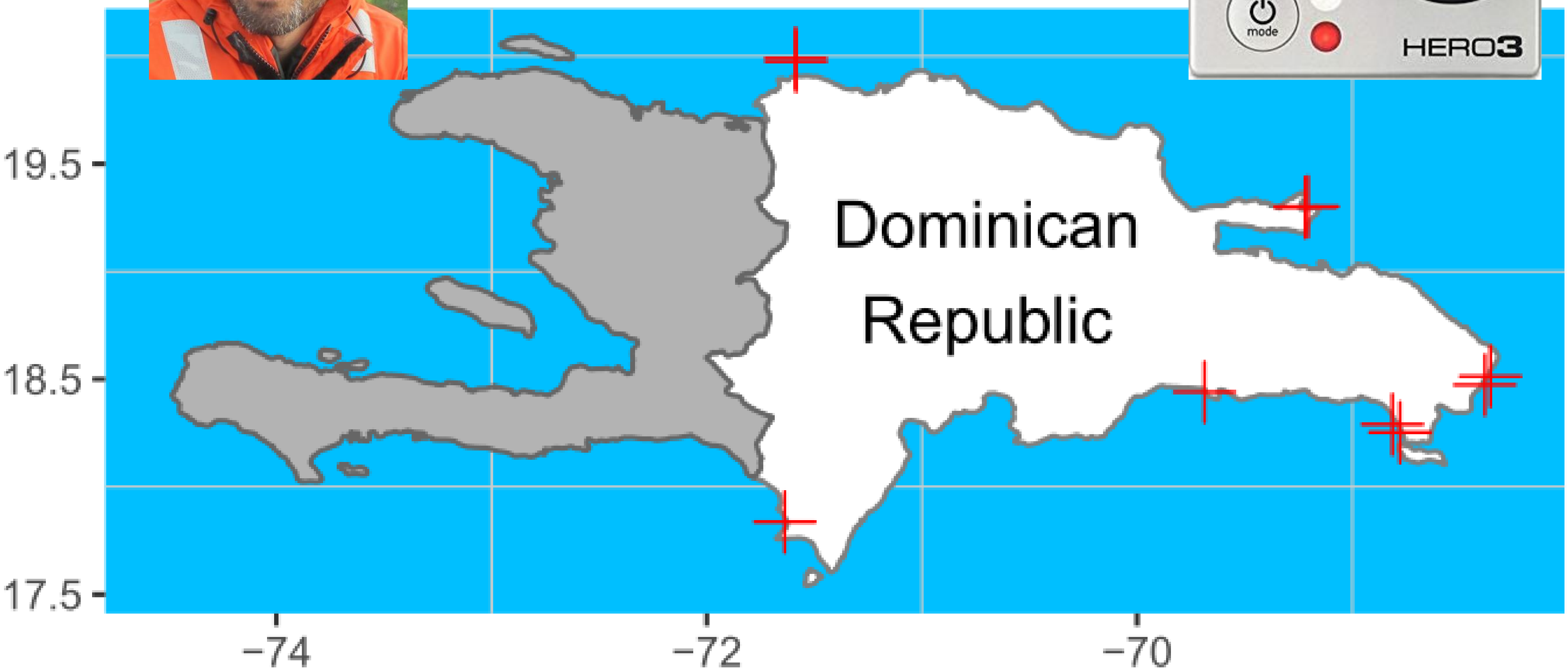
- Rare to find a mediator that is totally unaffected by confounding variables (condition 2)

3.2 Causality. An Example from Coral Reefs

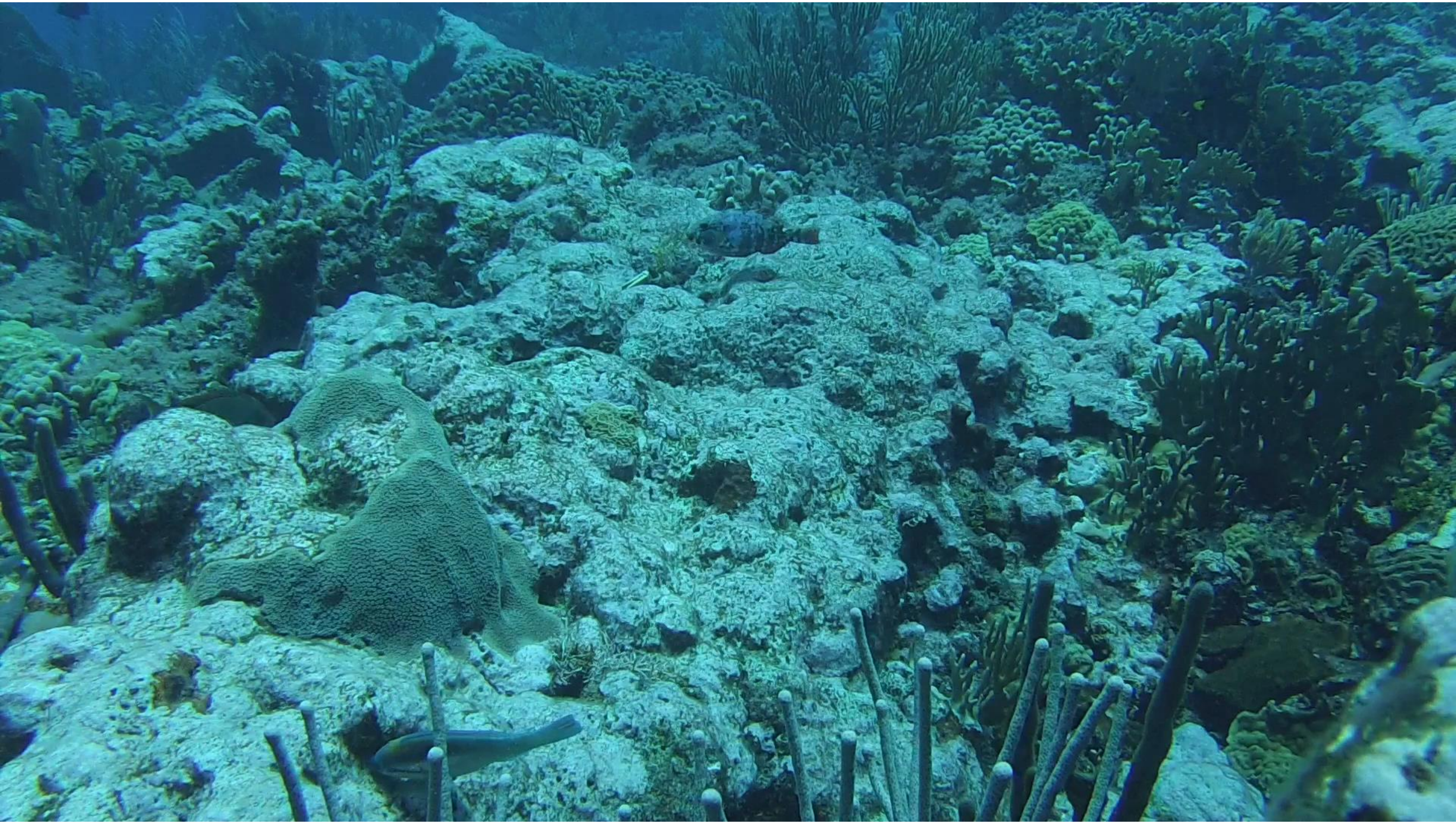
3.2 Causality. Algal phase shift



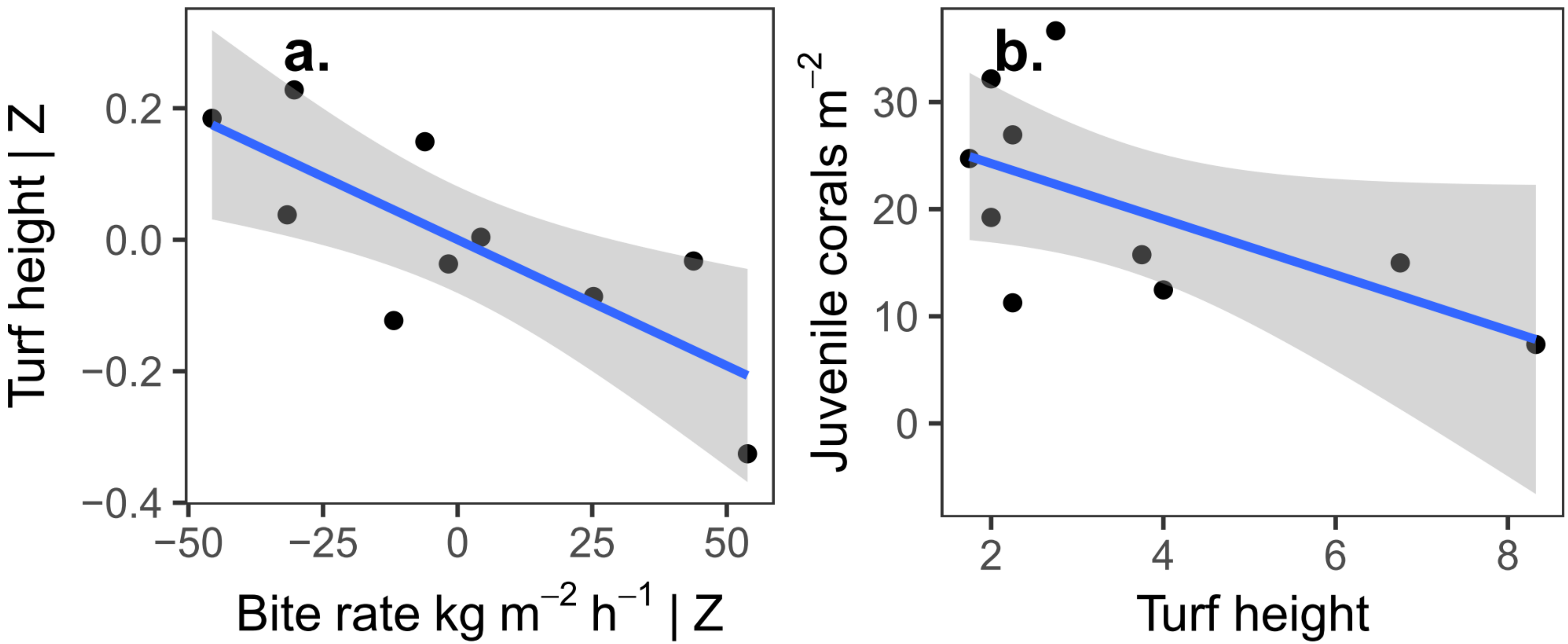
3.2 Causality. Observational study



3.2 Causality. Herbivores grazing



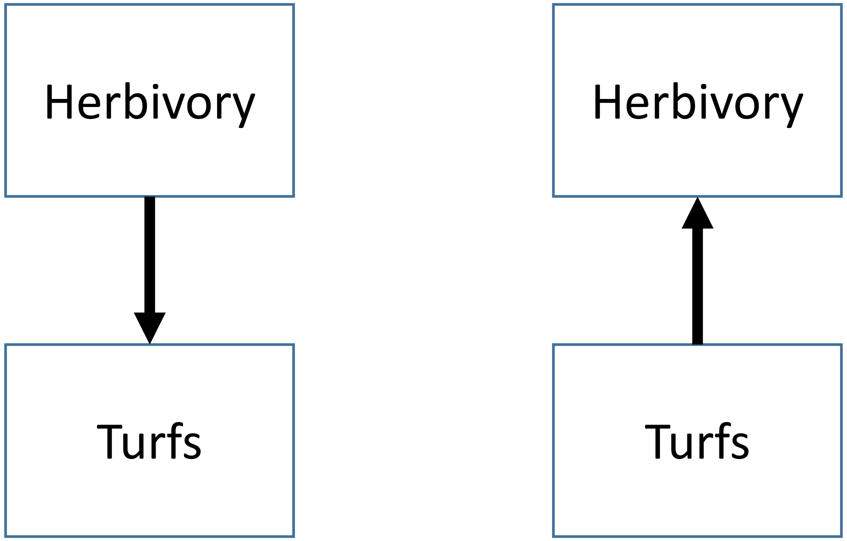
3.2 Causality. Herbivory helps!



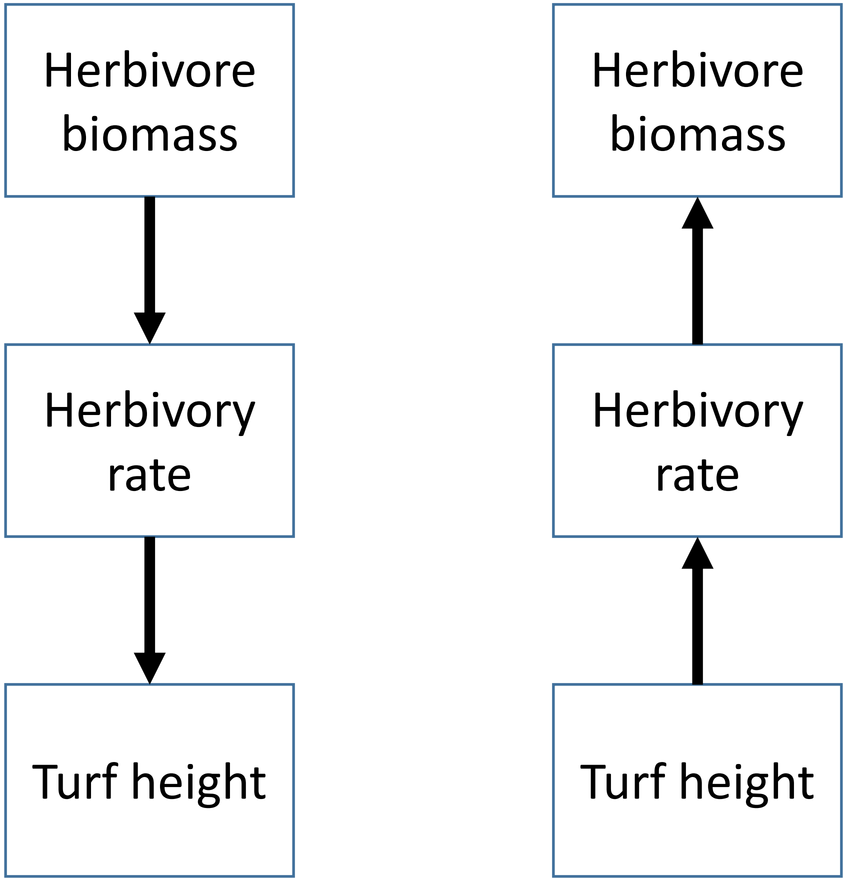
3.2 Causality. Reviewer 3

"The problem is that there is an equal probability that it is a bottom up response. That it is the benthic conditions (= low turfs) that drive increased herbivore biomass, diversity and coral recruitment." *–Reviewer #3*

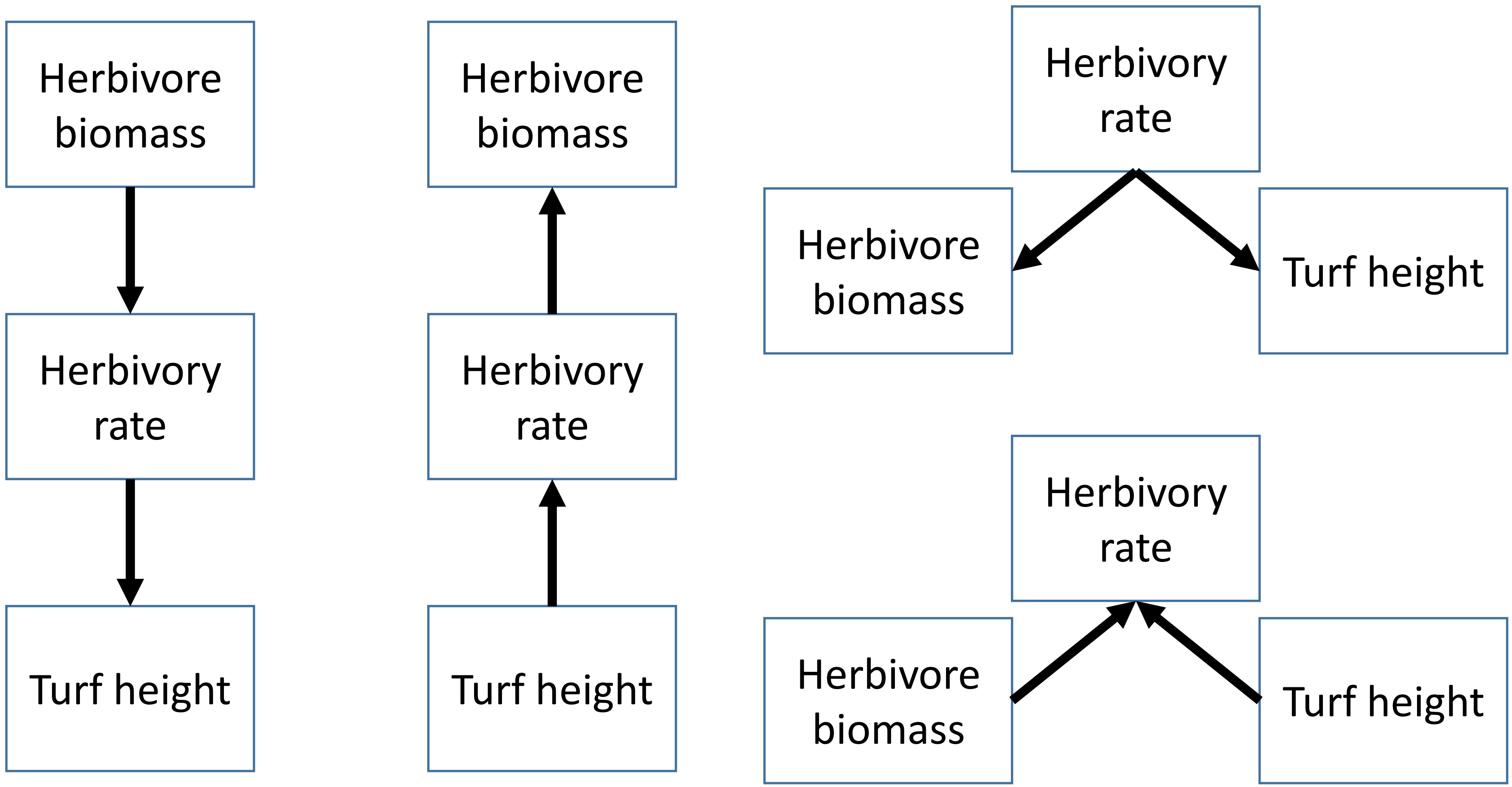
3.2 Causality. Alternate structures



3.2 Causality. Alternate structures

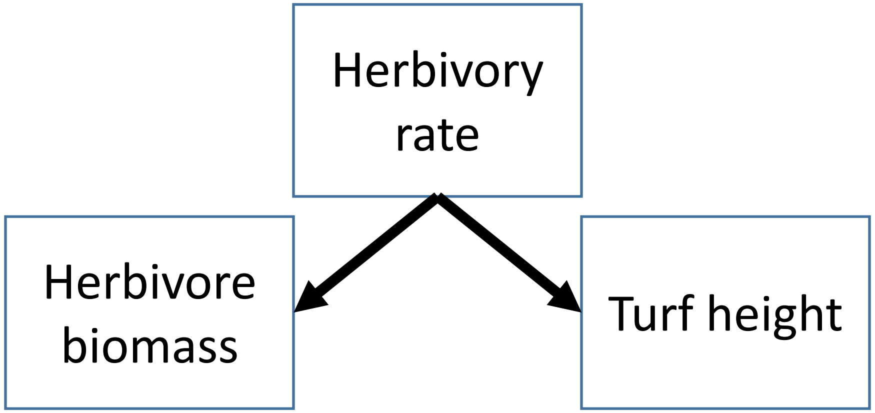


3.2 Causality. Alternate structures

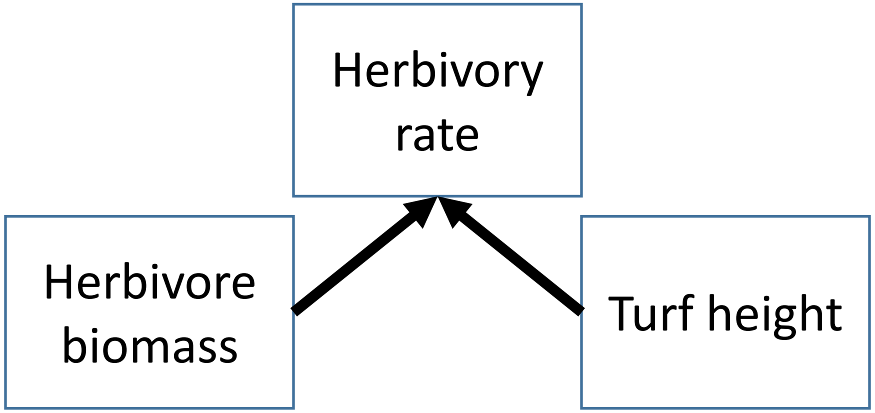


3.2 Causality. Alternate structures

<u>*Fork* (common cause)</u>



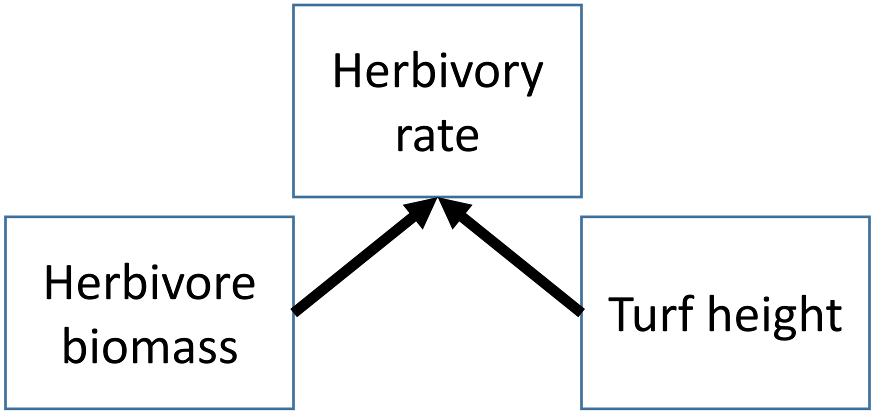
<u>*Collider* (common effect)</u>



3.2 Causality. Alternate structures

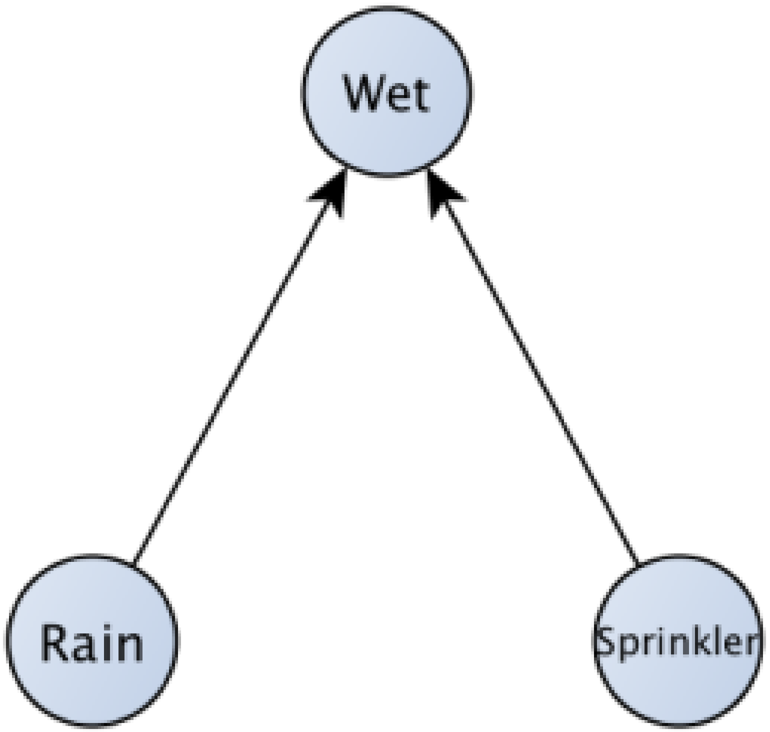
- If this graph is true: Herbivore biomass and turf height are independent *unless* conditioned on herbivory rate
- Knowing herbivory rate was high when turf height was low means herbivore biomass is likely high (because these are the two causes of high herbivory rates)

Collider (common effect)



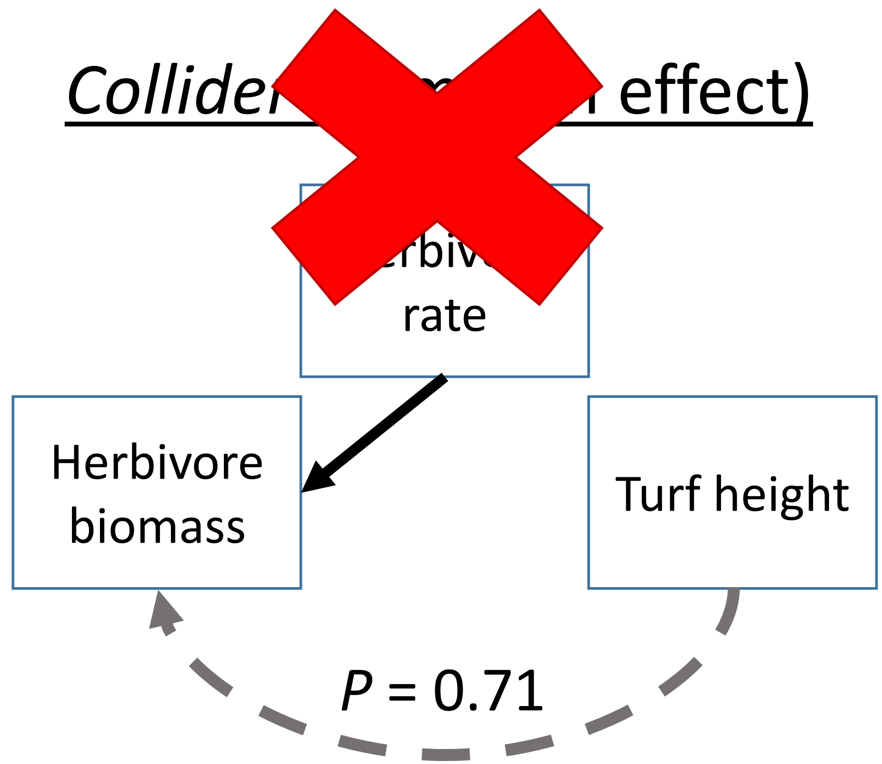
3.2 Causality. Alternate structures

Is the sidewalk wet?

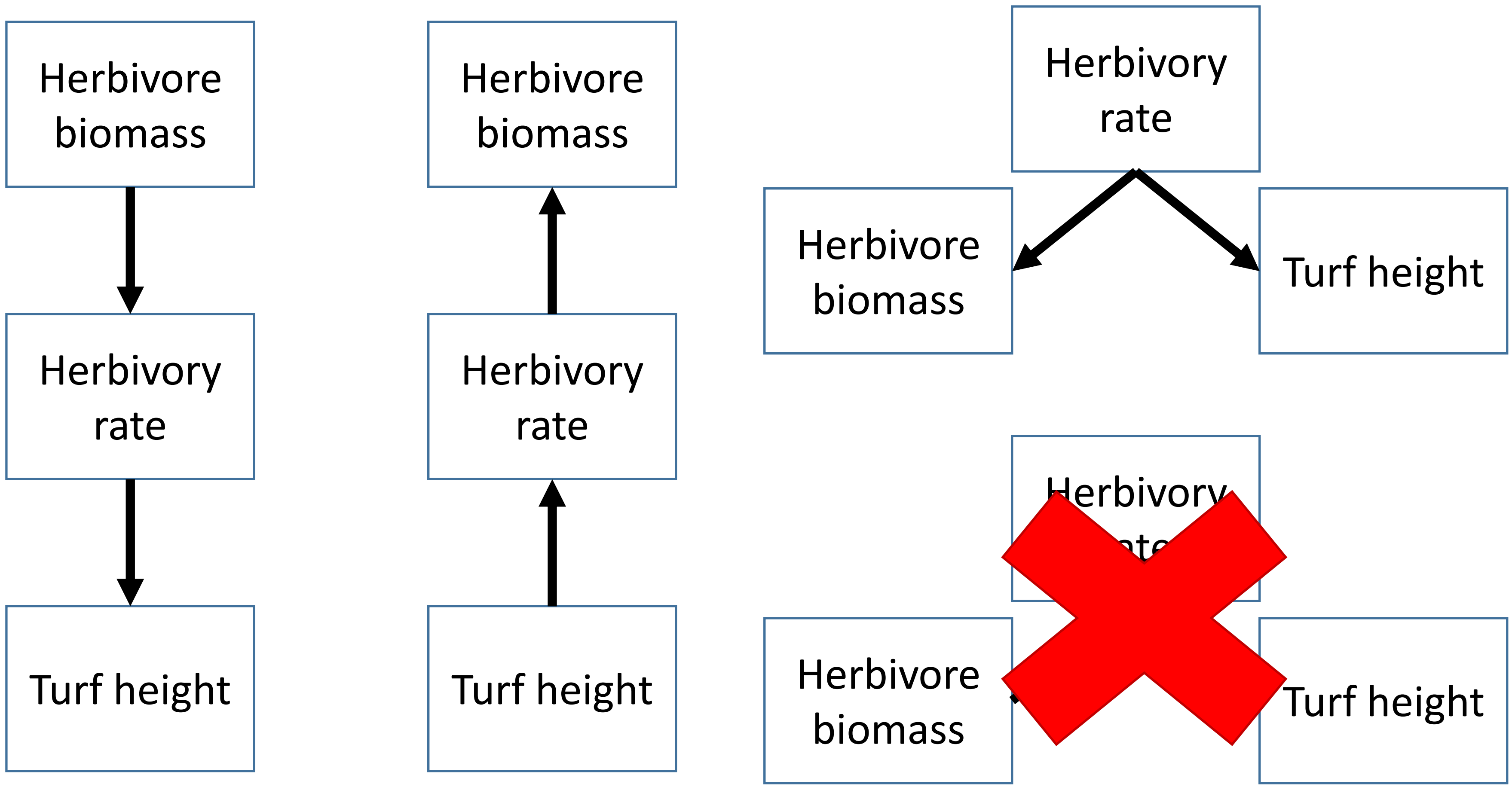


If yes and the sprinkler is off…it must be raining!

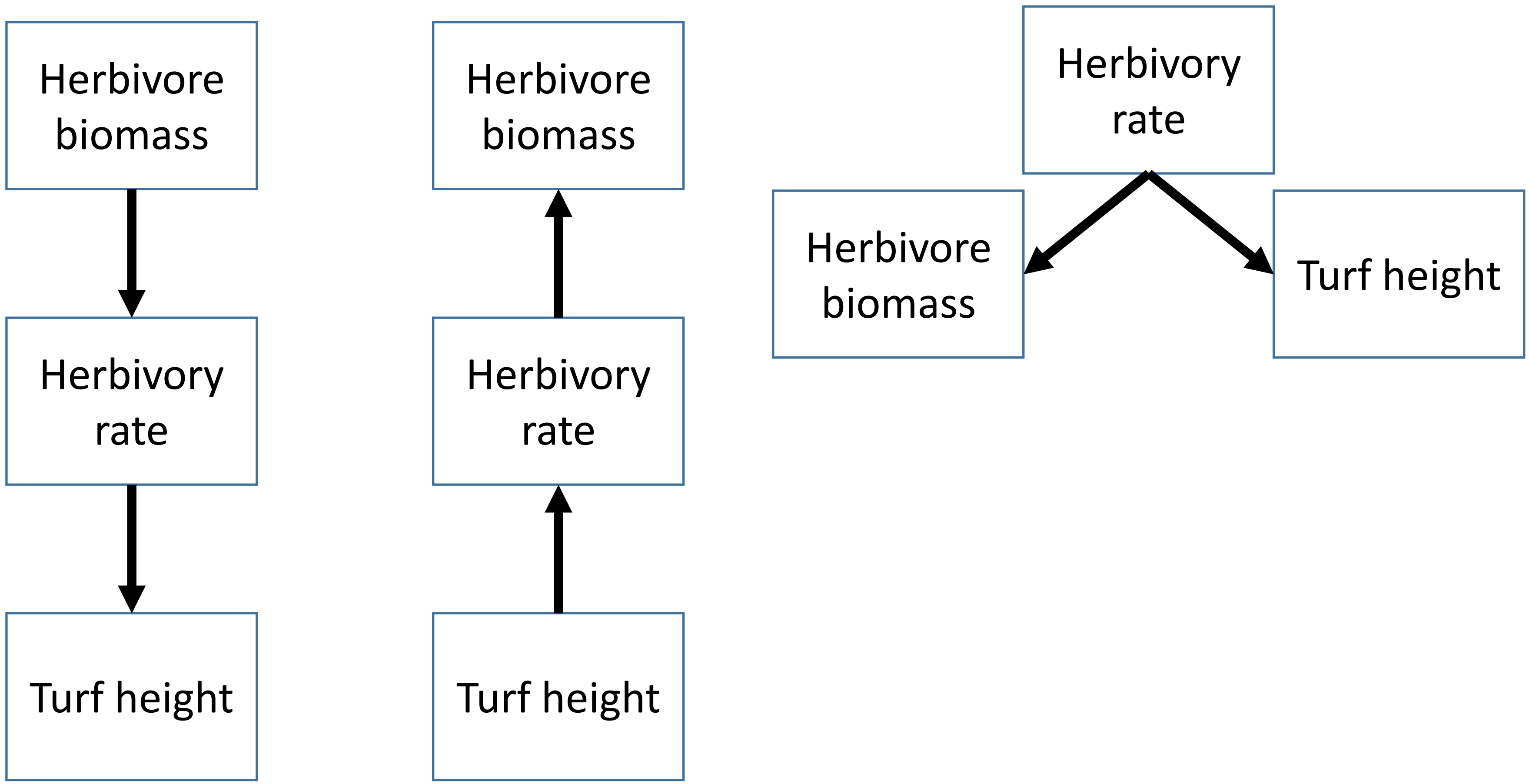
- Test for dependence of herbivore biomass and turf height given herbivory rate: if not significant, then the relationship is *not* a collider
- Knowing something about turf height tells you nothing about herbivory biomass…



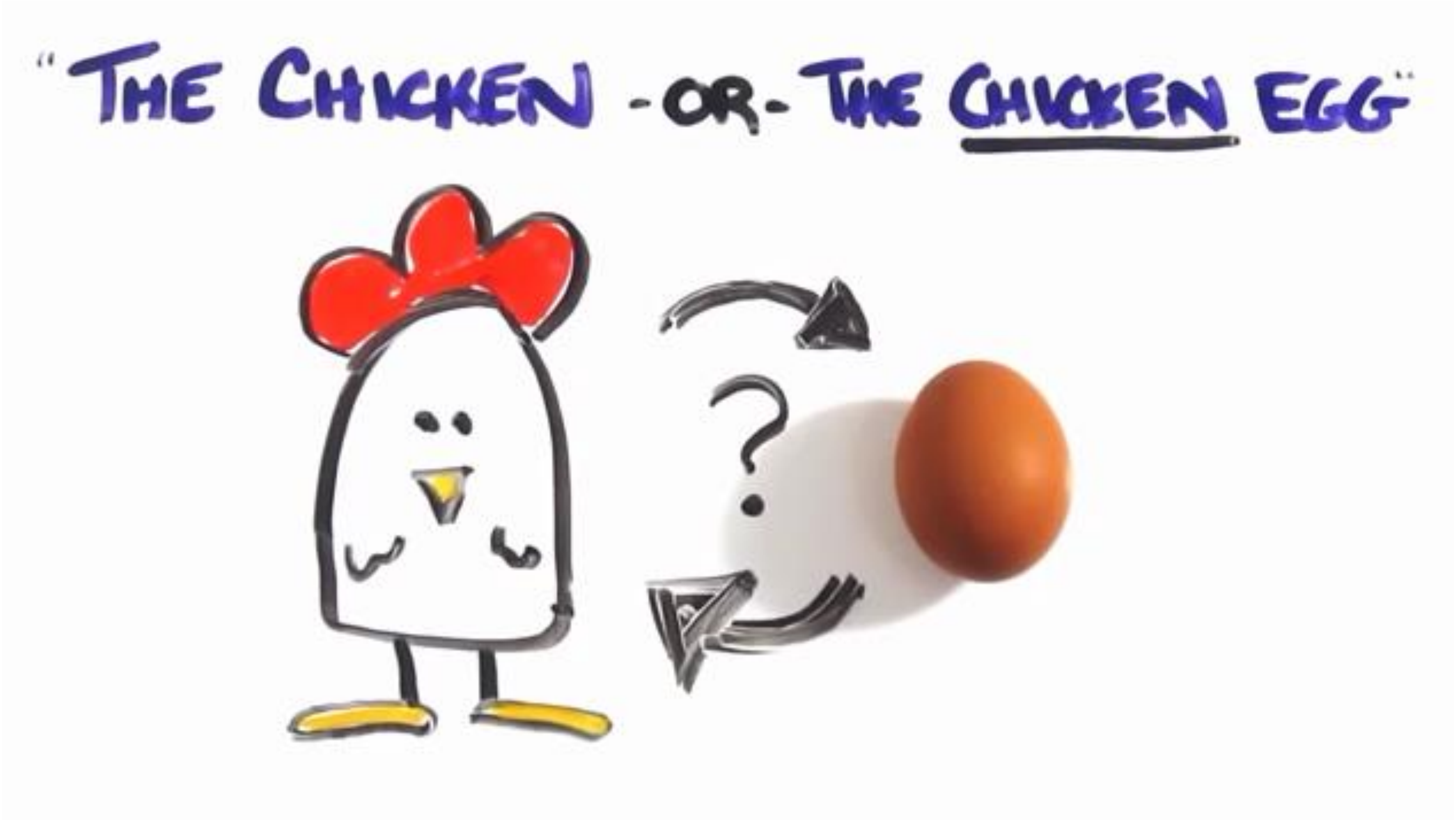
3.2 Causality. Alternate structures



3.2 Causality. Logical deduction



3.2 Causality. Logical deduction



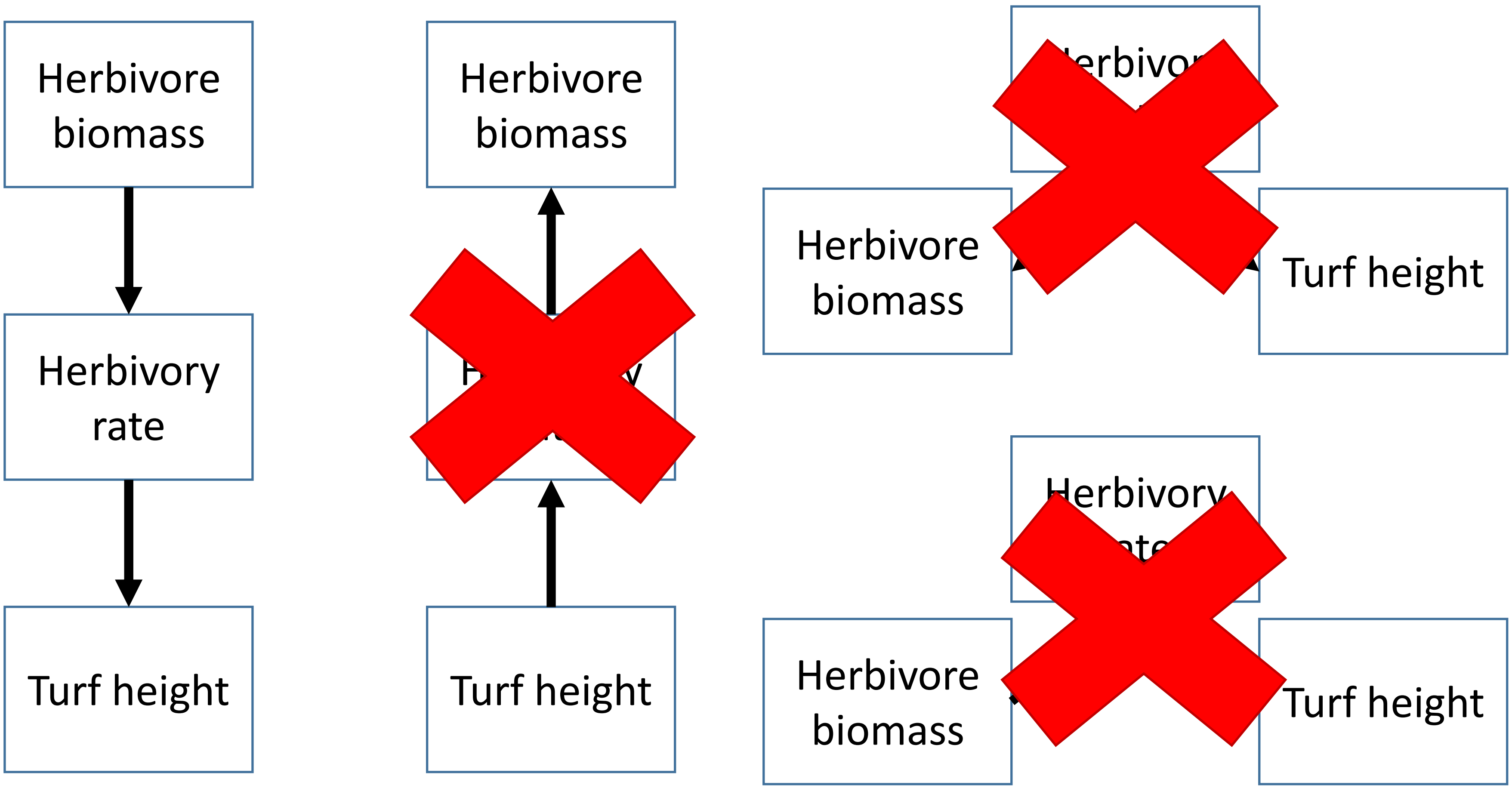
3.2 Causality. Logical deduction

Do herbivores cause herbivory rates?

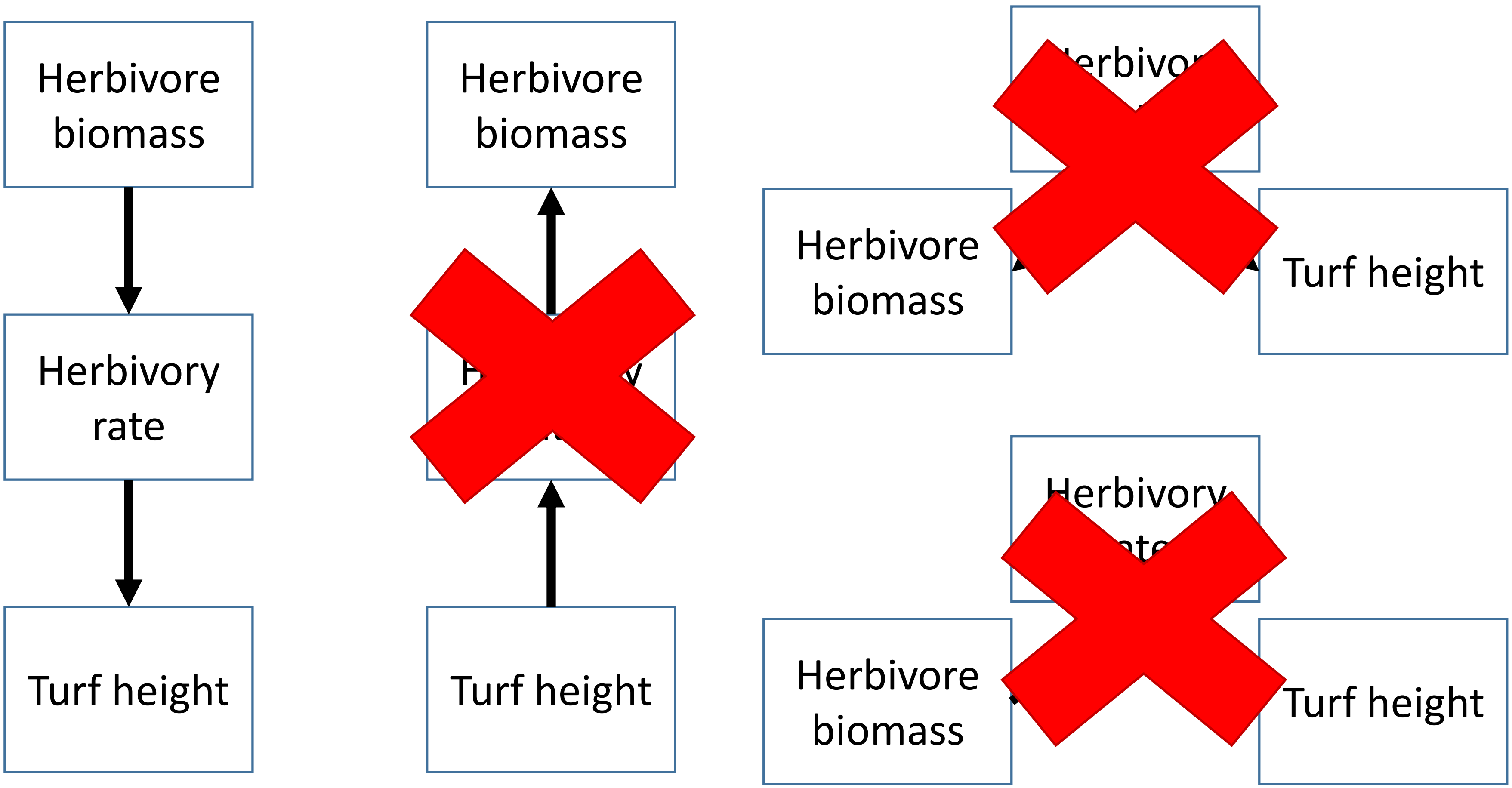
Or do herbivory rates cause herbivores?

<u>Can we measure herbivory rate without first observing herbivores?</u>

3.2 Causality. Alternate structures



3.2 Causality. Alternate structures



3.2 Causality. Other evidence

Ecosystems (2009) 12: 1316–1328
DOI: 10.1007/s10021-009-9291-z

ECOSYSTEMS

Limited Functional Redundancy in a High Diversity System: Single Species Dominates Key Ecological Process on Coral Reefs

Andrew S. Hoey* and David R. Bellwood

Australian Research Council Centre of Excellence for Coral Reef Studies and School of Marine and Tropical Biology, James Cook University, Townsville, Queensland 4811, Australia

Abstract

Herbivory is a key process structuring plant communities in both terrestrial and aquatic ecosystems, with variation in herbivory often being related to shifts between alternate states. On coral reefs, regional reductions in herbivores have underpinned shifts from coral to dominance by leathery macroalgae. These shifts appear difficult to reverse as these macroalgae are unpalatable to the majority of herbivores, and the macroalgae suppress the recruitment and growth of corals. The removal of macroalgae is, therefore, viewed as a key ecological process on coral reefs. On the Great Barrier Reef, *Sargassum* is a dominant macroalgal species following experimentally induced coral–macroalgal phase-shifts. We, therefore, used *Sargassum* assays and remote video cameras to directly quantify the species responsible for removing macroalgae across a range of coral reef habitats on Lizard Island, northern Great Barrier Reef. Despite supporting over 50 herbivorous fish species and six macroalgal browsing species, the video footage revealed that a single species, *Naso unicornis*, was almost solely responsible for the removal of *Sargassum* biomass across all habitats. Of the 42,246 bites taken from the *Sargassum* across all habitats, *N. unicornis* accounted for 89.8% (37,982) of the total bites, and 94.6% of the total mass standardized bites. This limited redundancy, both within and across local scales, underscores the need to assess the functional roles of individual species. Management and conservation strategies may need to look beyond the preservation of species diversity and focus on the maintenance of ecological processes and the protection of key species in critical functional groups.

Key words: *Naso unicornis*; functional redundancy; phase-shift; macroalgae; *Sargassum*; coral reef; herbivory.

Introduction

Herbivory is widely acknowledged as a key process structuring plant communities in both terrestrial and aquatic ecosystems (Scheffer and others 2001). Whilst there are fundamental differences among ecosystems in the nature of herbivory and its importance relative to other processes (Shurin and others 2006; Gruner and others 2008), areas of

Received 30 July 2009; accepted 27 September 2009; published online 23 October 2009

Electronic supplementary material: The online version of this article (doi:10.1007/s10021-009-9291-z) contains supplementary material, which is available to authorized users.

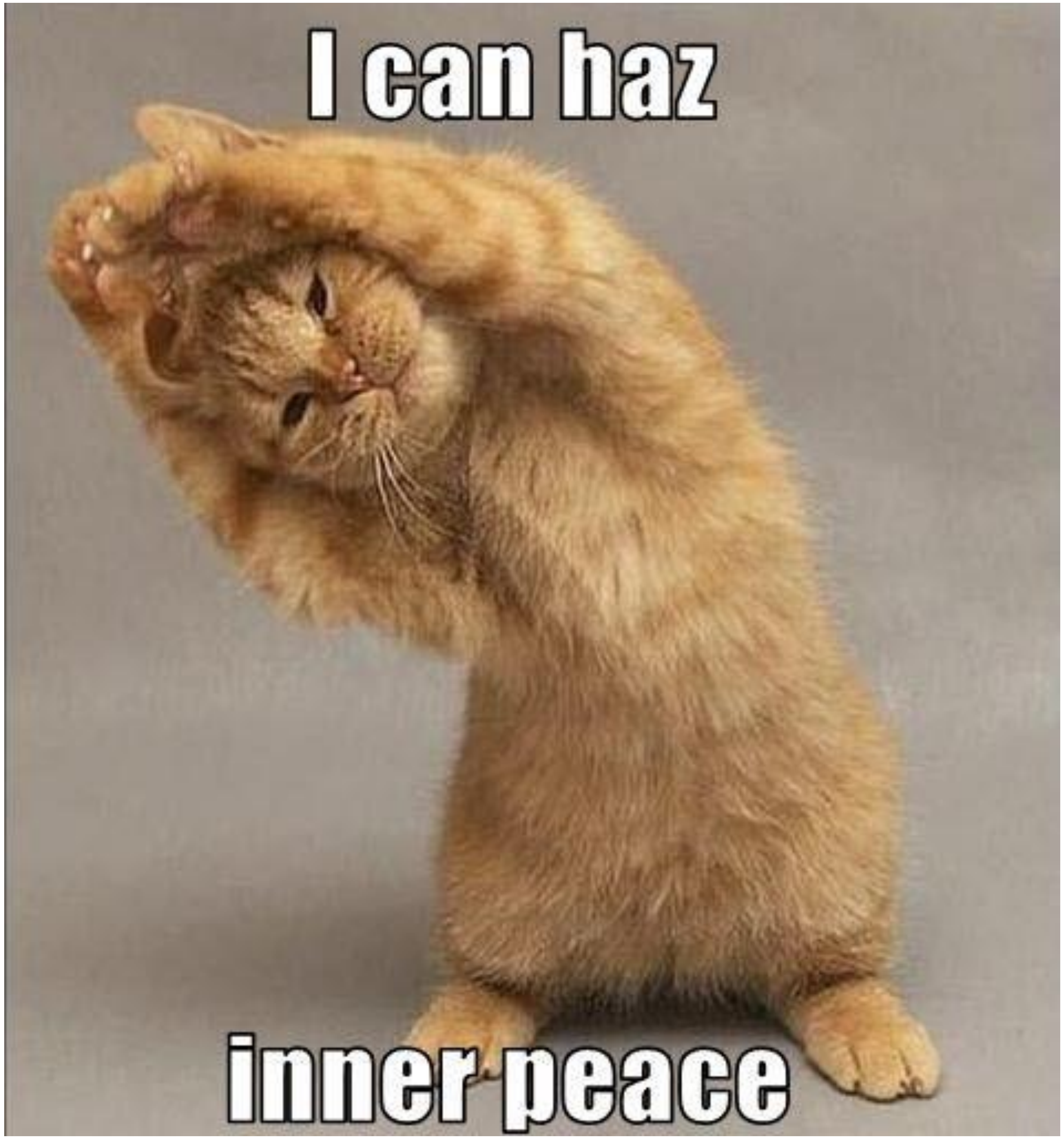
**Corresponding author; e-mail:* andrew.hoey@jcu.edu.au

3.2 Causality. How?

- Statistical tests of conditional independence
- Application of logical “induction” about the direction of causality
- But may still fail condition 2 (herbivore biomass and herbivory rate might have hidden external drivers, e.g., temperature) → may not be an actual front door if not independent of exogenous drivers

3.2 Causality. Conclusions

- Thoughtful construction of your path diagram can lead to stronger causal inferences
- BE TRANSPARENT: say why you thinks paths are or aren’t causal and provide the evidence or justification
- May not be ultimately causal or may later be revealed to be spurious
- BE HONEST: science advances by others noticing what you left out


I can haz
inner peace