

# Assignment 01

Intro

Prof. Dr. Ralf Lämmel & M.Sc. **Johannes Härtel**  
([johanneshaertel@uni-koblenz.de](mailto:johanneshaertel@uni-koblenz.de))

Important

Do experiment/  
collect data and  
then analyze.

Important

Focus of next *data  
collection*  
assignment.

Question	Design	Study Type	Critical for the Study Type	Data Sources
<i>Write 1 sentences covering a research question.</i>	<i>Write 1-2 sentences describing the study design.</i>	<i>Classify the study design into the experimental and observational study type.</i>	<i>Write 1-2 sentences to explain why you associate it with the particular study type and not the other.</i>	<i>Add existing digital data sources (if not experimental study).</i>
Does IQ of a mother influence the IQ of a child?	We collect data on mother and child IQ and check the relation.	Observational	We can not modify the IQ of the mother.	We may analyze written text in social networks and ancestor relations.

# Summary

- **Questions**

- Try to form clean questions on causation between X and Y.

- **(Critical for) Study Type**

- List reasons for experimental/observational: “missing data”, “ethical or practical problem with modification”, “modification impossible” and “environmental issues”.
- The essential property of an experiment is that some property X is modified at random by doing the experiment. If we just sample subjects with the intended properties, we have an observational study.

# Summary (Cont)

- **Data Source**

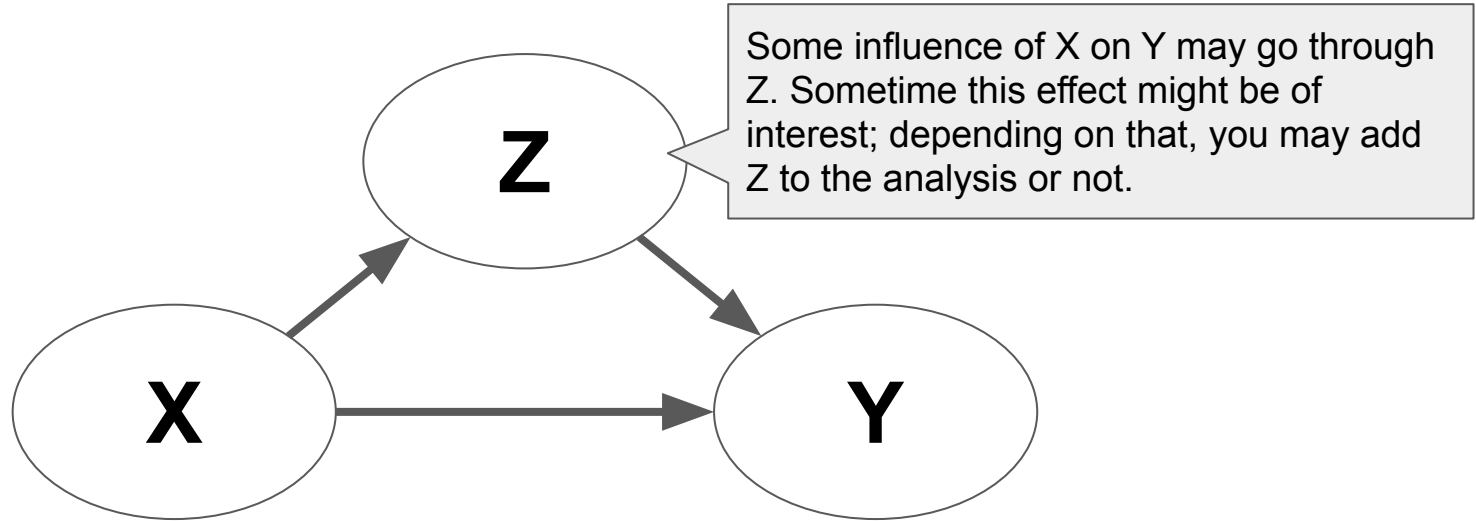
- Simulated/fake data can not be used for answering real research questions. It is just a way of testing your methodology.

- **Study Design:**

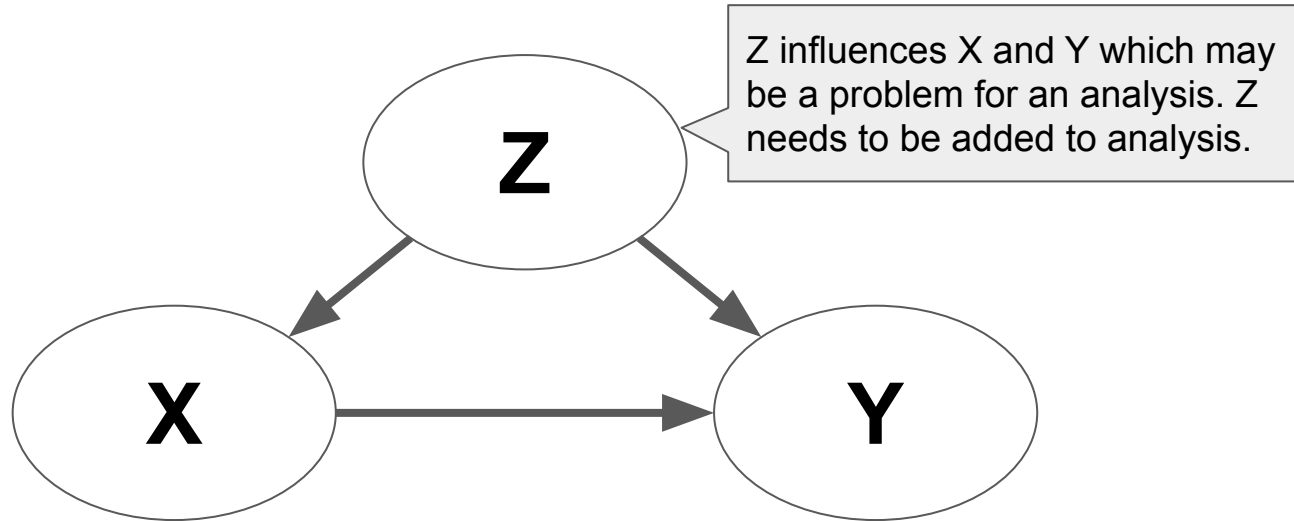
- Basic answers on do experiment/collect data then analyze are sufficient. There are so many variations and pitfalls that we cannot systematically cover at this point in time.

# Some more background (not asked for in assignment 1)

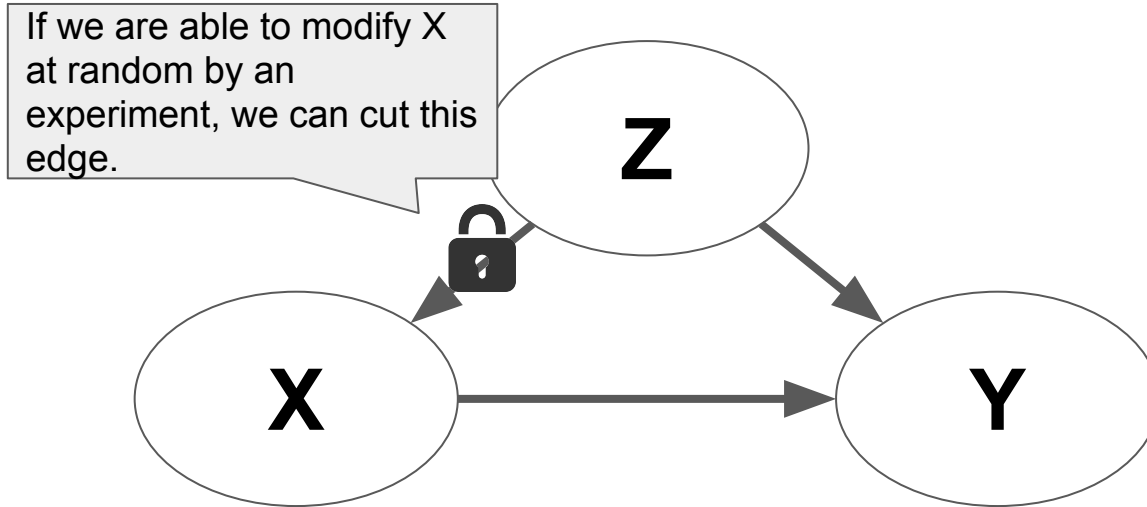
## More Background: A **Mediator** Variable Z



# More Background: A **Classical Confounder** Variable Z

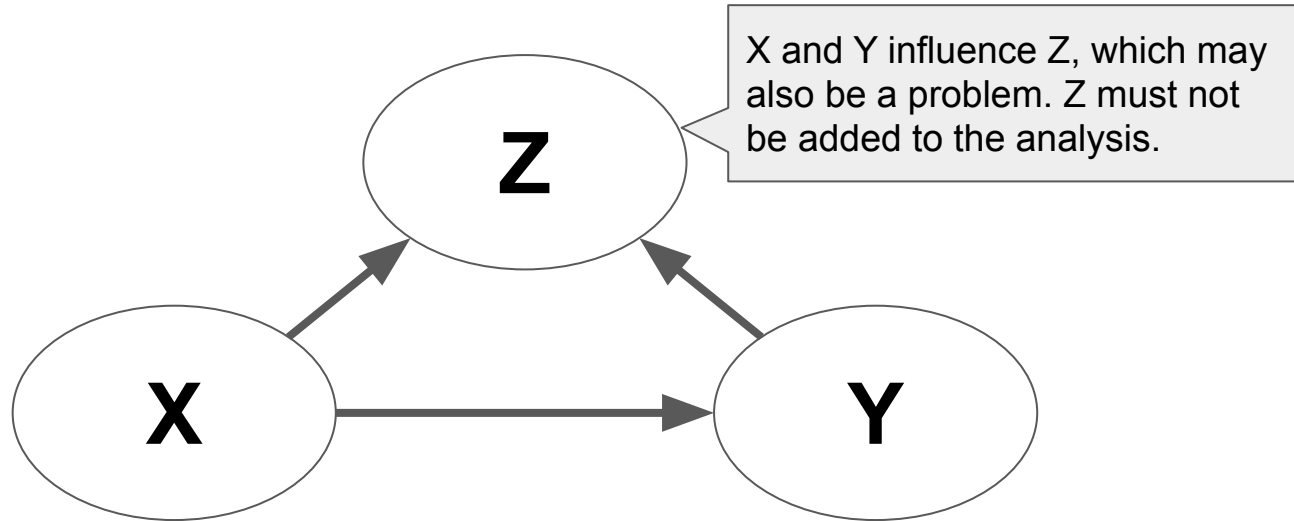


# More Background: Benefit of an **Experiment**





## More Background: A **Collider** Variable Z



# Some of your Questions

# What if we change the design of our web shop?



# What if we change the design of our web shop, does it increase sales for a certain product?

Don't miss the **effect of your cause!**  
(This is also a problem in the lecture slides, sorry for that)



# Does the design of our web shop influence the sales for a certain product?

More refactoring.



Does the design of our web shop (**Des**) influence the sales for a certain product (**Sal**)?



# Does the design of our web shop (**Des**) influence the sales for a certain product (**Sal**)?



If we are not interested in a commonly applied design changes, but in a very specific change for our shop, we need to run **an experiment**, since corresponding **data is missing**.



One example of a commonly applied design change: **Flat design** is a user interface design style that uses simple, two-dimensional elements and bright colors

Does age of a person affect the use of products because of familiarity with these types of products?

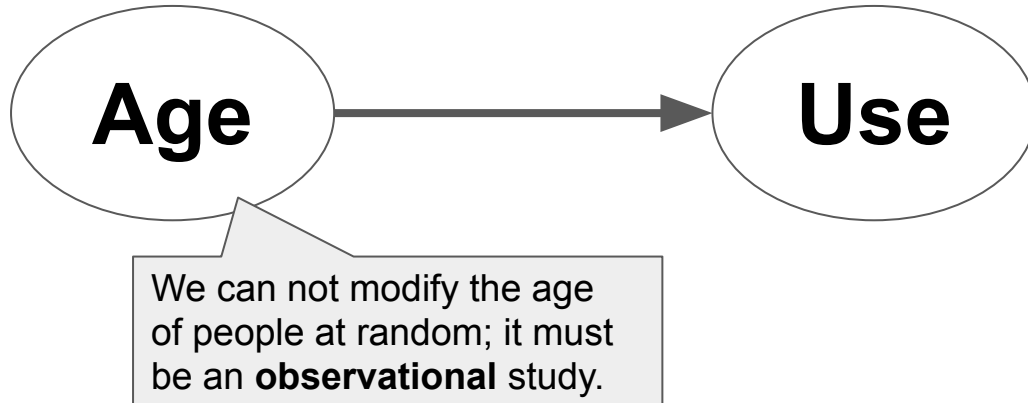




Does age of a person (**Age**) affect the use of products (**Use**) because of familiarity with these types of products?



Does age of a person (**Age**) affect the use of products (**Use**) because of familiarity with these types of products?

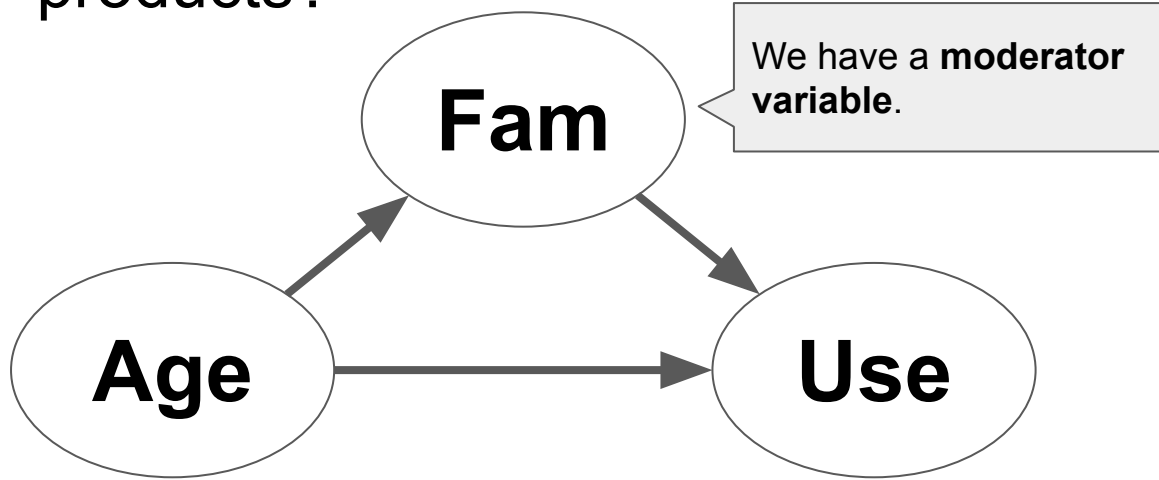


# Does age of a person (**Age**) affect the use of products (**Use**) because of familiarity with these types of products?

Some of you stated that we can sample two groups (with different age). This is **no experiment**, since we do not modify the age at random.



Does age of a person (**Age**) affect the use of products (**Use**) because of familiarity (**Fam**) with these types of products?



# Does the consumption of alcohol during pregnancy cause premature birth?



Does the consumption of alcohol during pregnancy (**Alc**) cause premature birth (**PrB**)?



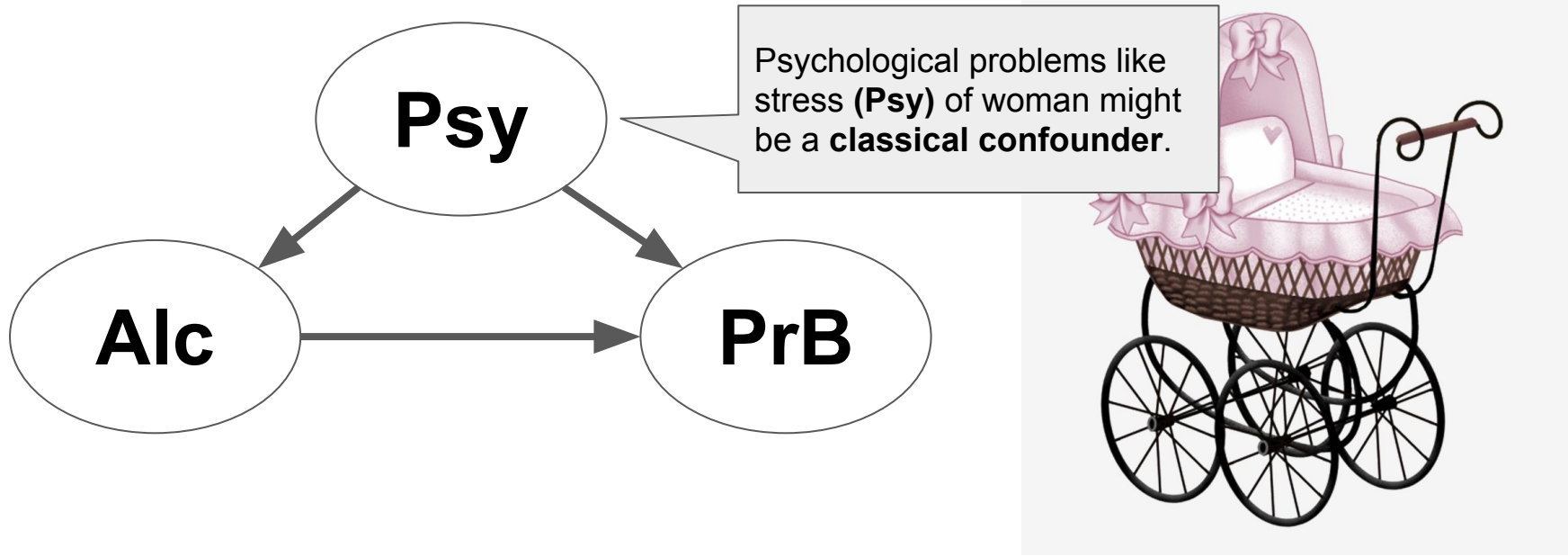
# Does the consumption of alcohol during pregnancy (**Alc**) cause premature birth (**PrB**)?



It is **ethically problematic** to modify drinking behavior of pregnant women, must be **observational**.

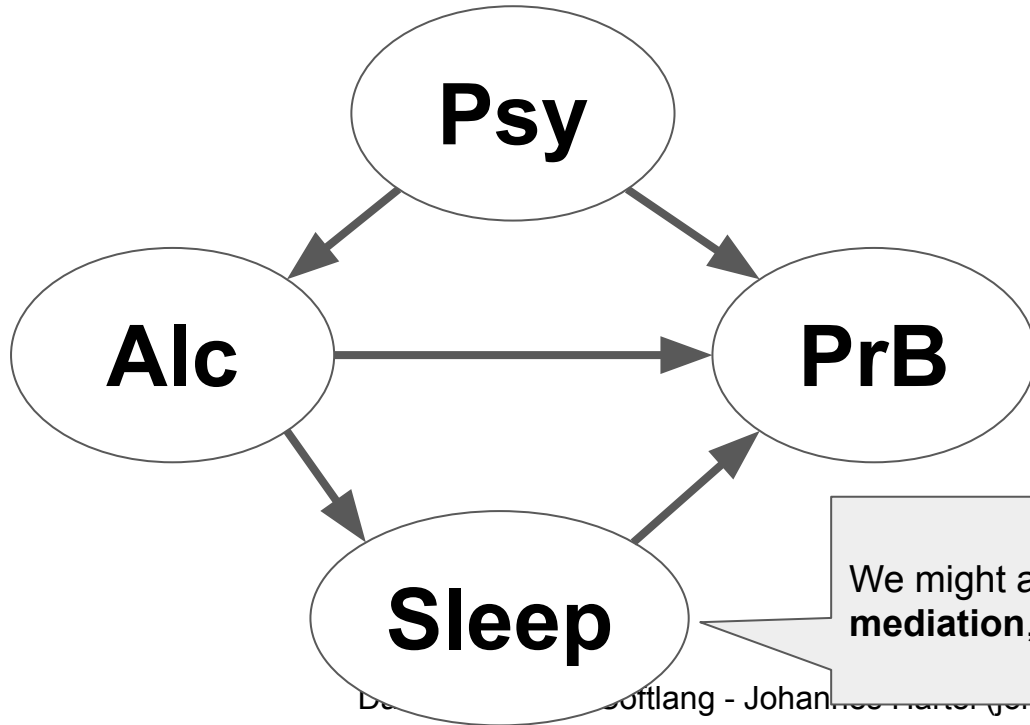


# Does the consumption of alcohol during pregnancy (**Alc**) cause premature birth (**PrB**)?





Does the consumption of alcohol during pregnancy (**Alc**) cause premature birth (**PrB**)?



We might also have **mediation**, e.g., by sleep.

# Does the time it takes to reach your workplace affect your job satisfaction?



Does the time it takes to reach your workplace (**T**)  
affect your job satisfaction (**SAT**)?



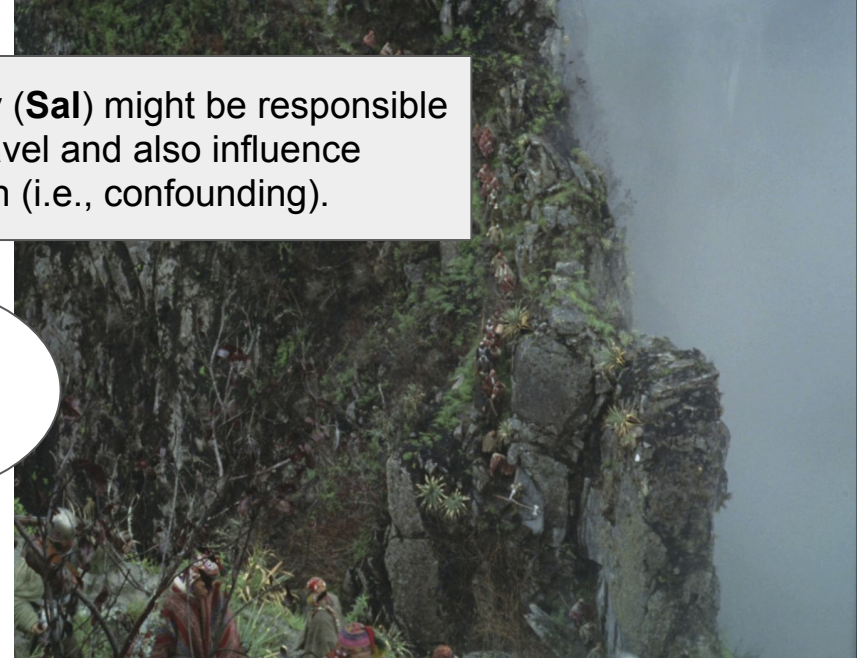
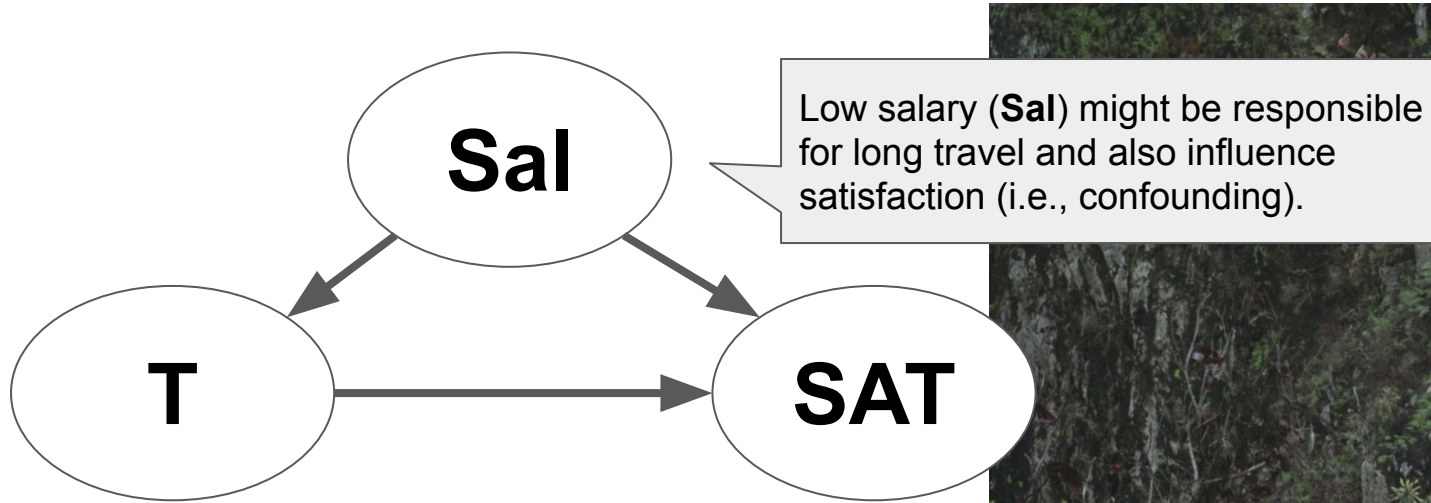
# Does the time it takes to reach your workplace (**T**) affect your job satisfaction (**SAT**)?



It is very **complicated to modify** the time it takes to reach work; must be **observational**.



# Does the time it takes to reach your workplace (**T**) affect your job satisfaction (**SAT**)?





# What type of pitches are useful for fast bowlers in cricket?



# What type of pitches are useful for fast bowlers in cricket **in terms of goals**?

I adjusted this a little so that we can **quantify “useful”**  
(Hope this works. I don’t know cricket very well).



# Does the type of pitches influence goals for fast bowlers?

More changes to better separate X and Y.

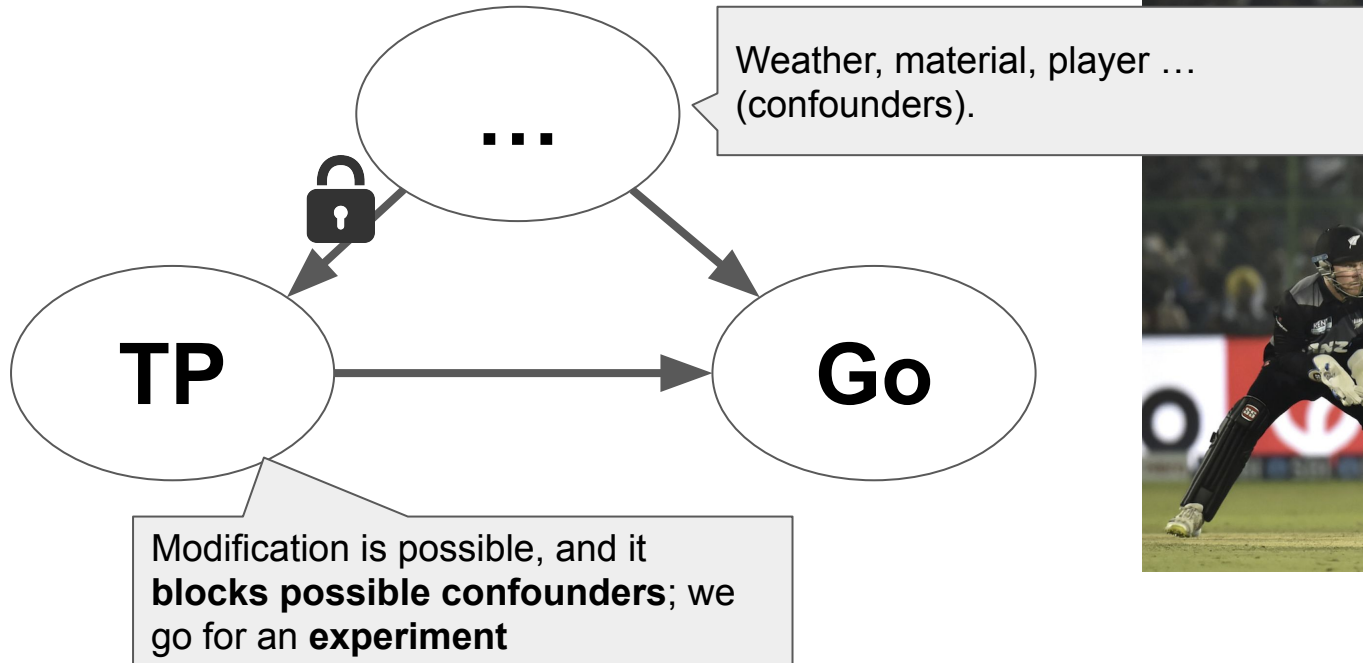




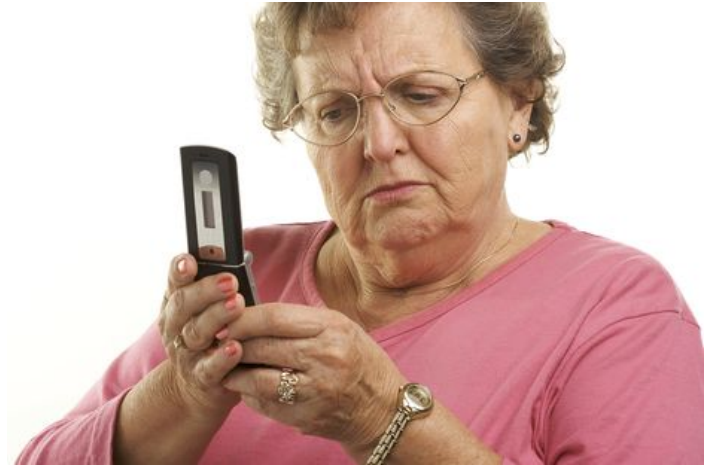
# Does the type of pitches (**TP**) influence goals (**Go**) for fast bowlers?



# Does the type of pitches (X) influence **goals** (Y) for fast bowlers?



Do users between an age of 18 - 30 post more frequently on Twitter than people of an age  $> 30$ ?



# Does the age (18 - 30 or $> 30$ ) influence the frequency of postings on Twitter?

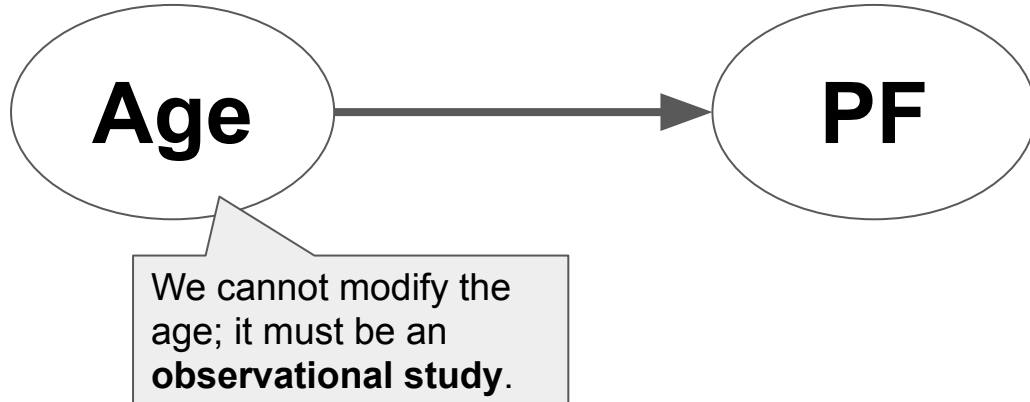
Reformulate



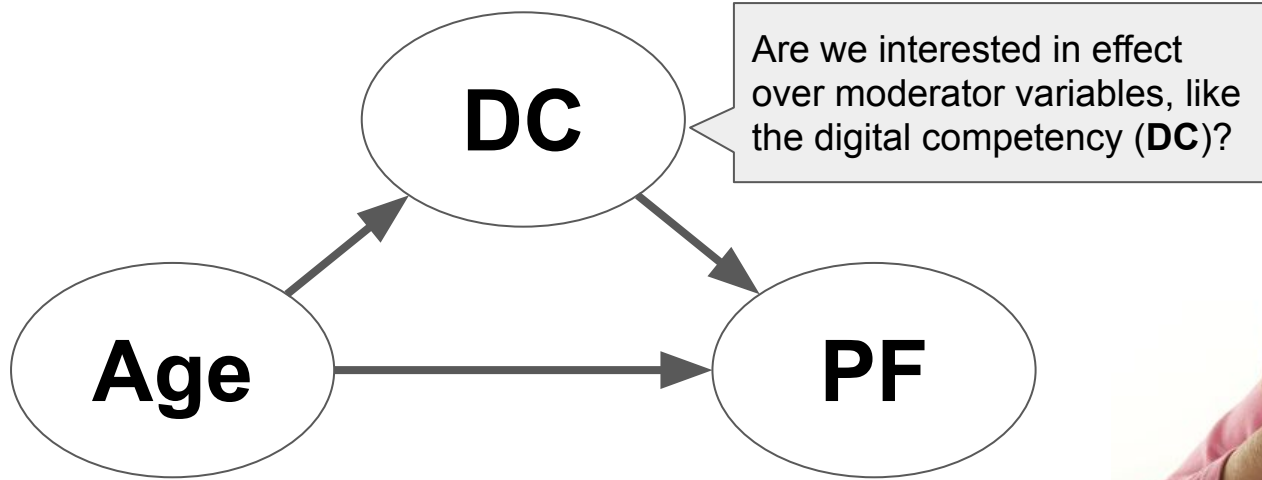
Does the age (18 - 30 or > 30, **Age**) influence the frequency of postings on Twitter (**PF**)?



# Does the age (18 - 30 or > 30, **Age**) influence the frequency of postings on Twitter (**PF**)?



# Does the age (18 - 30 or > 30, **Age**) influence the frequency of postings on Twitter (**PF**)?



Would cancelling all hygiene rules with respect to corona cause a higher mortality rate in a certain country?





# Do Corona hygiene rule influence mortality rate in a country?

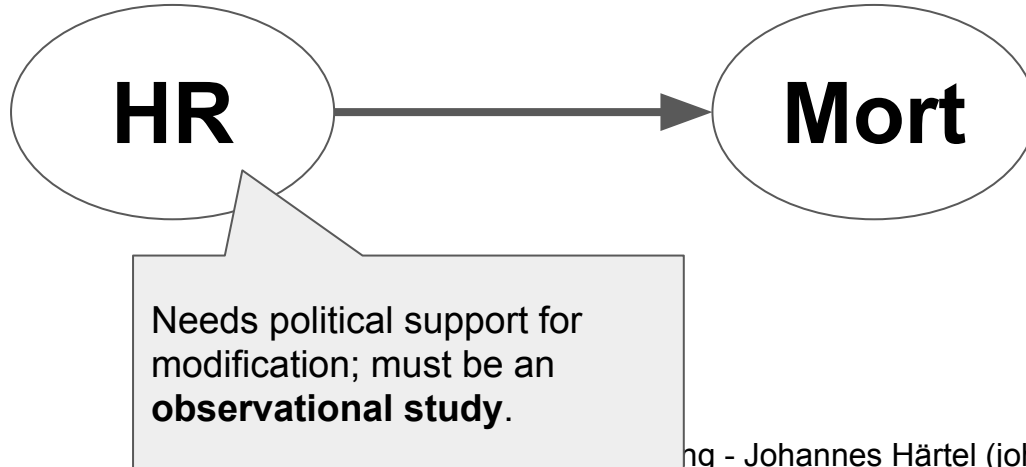
Reformulated



# Do Corona hygiene rule (**HR**) influence mortality rate in a country (**Mort**)?



# Do Corona hygiene rule (**HR**) influence mortality rate in a country (**Mort**)?



# Are female citizens less likely to hold a political office in the European Union?



# Does the Sex influence the assignment of political positions in the European Union?

Reformulate



# Does the Sex (**Sex**) influence the assignment of political positions in the European Union (**Pos**)?



# Does the Sex (**Sex**) influence the assignment of political positions in the European Union (**Pos**)?



We can not modify the sex; it must be an **observational study**.



# Does the Sex (**Sex**) influence the assignment of political positions in the European Union (**Pos**)?

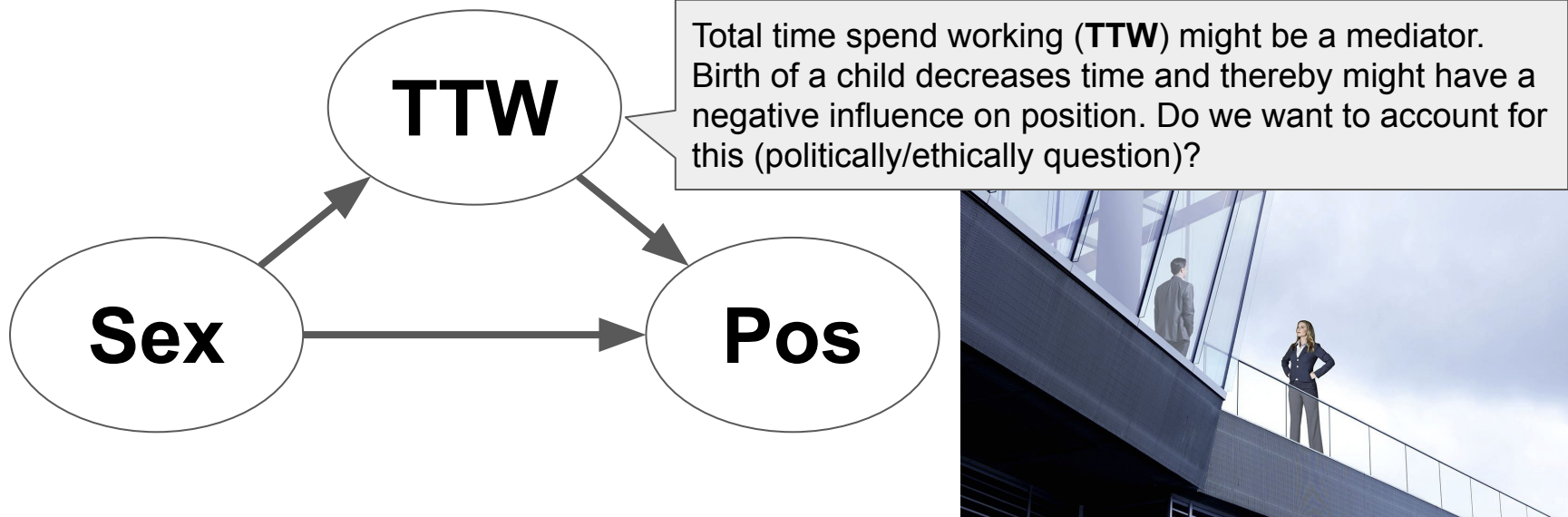


Depending on the context, there are some **experimental** ways how to “modify” the sex. In a job application, we can test for different responses with sex assigned randomly. However, does **not** cover the problem in its **original environment**.





# Does the Sex (**Sex**) influence the assignment of political positions in the European Union (**Pos**)?



# Does Covid-19 effect the global number of flight passengers?



# Does Covid-19 (**Cov**) effect the global number of flight passengers (**FIP**)?



# Does Covid-19 (**Cov**) effect the global number of flight passengers (**FIP**)?



We can not travel into the past and prevent Covid from breaking out (modify it).

# Does Covid-19 (**Cov**) effect the global number of flight passengers (**FIP**)?



However, we can also **not compare two worlds**; one with and one without covid. In this case, we may compare previous years with covid years (a trick employing auto-correlation). This is (longitudinal) **observational** study.

# Does Covid-19 (**Cov**) effect the global number of flight passengers (**FIP**)?

