STA 235H - Potential Outcomes

Fall 2021

McCombs School of Business, UT Austin

Let's do a short exercise

Take three pieces of papers: One of each color

Don't look!

The problem

- Imagine **everyone** here has a headache
- Everyone takes an aspirin to mitigate the headache
- How many students **still have a headache** (look at your **orange** paper)

Wow, did the aspirin work?

How? Potential Outcomes Framework

What? Causal Estimands

Why? Causal Questions and Study Design

The "How": Potential outcomes framework

Geoffrey Supran
@GeoffreySupran

"The @GretaThunberg Effect" is now an empirically demonstrated, peer-reviewed phenomenon:

"We find that those who are more familiar with Greta Thunberg have higher intentions of taking collective actions to reduce global warming."

Open access: onlinelibrary.wiley.com/doi/epdf/10.11...

Received: 5 December 2020 | Accepted: 18 December 2020
DOI: 10.1111/jasp.12737

ORIGINAL ARTICLE

Journal of Applied Social Psychology WILEY

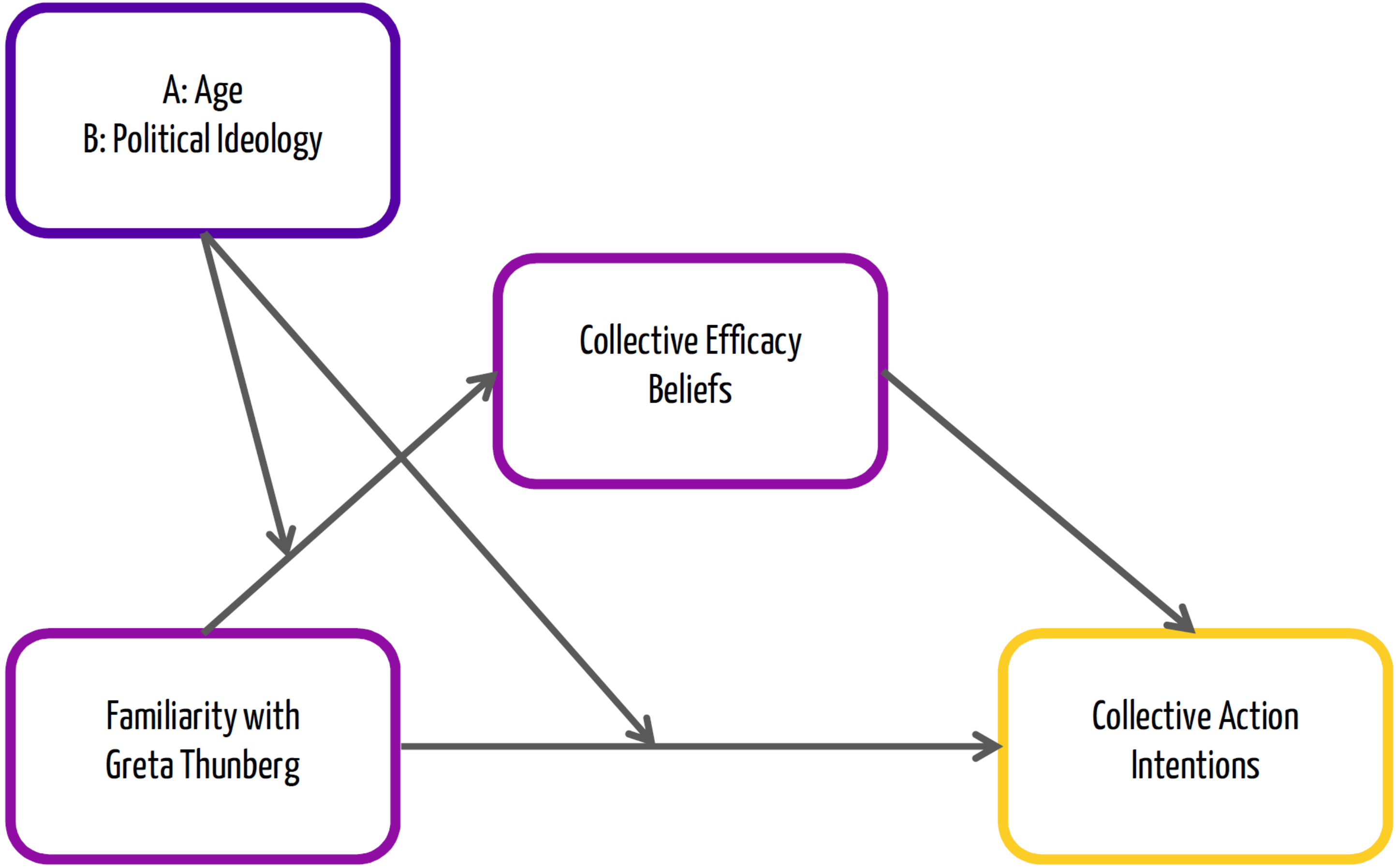
The Greta Thunberg Effect: Familiarity with Greta Thunberg predicts intentions to engage in climate activism in the United States

Anandita Sabherwal[1] | Matthew T. Ballew[2] | Sander van der Linden[1] | Abel Gustafson[3] | Matthew H. Goldberg[4] | Edward W. Maibach[5] | John E. Kotcher[5] | Janet K. Swim[4] | Seth A. Rosenthal[4] | Anthony Leiserowitz[4]

[1]Department of Psychology, University of Cambridge, Cambridge, UK
[2]Crean College of Health and Behavioral Sciences, Chapman University, Orange, CA, USA
[3]Department of Communication, University of Cincinnati, Cincinnati, OH, USA
[4]Yale Program on Climate Change Communication, Yale University, New Haven, CT, USA
[5]Mason Center for Climate Change Communication, George Mason University,

Abstract
Despite Greta Thunberg's popularity, research has yet to investigate her impact on the public's willingness to take collective action on climate change. Using cross-sectional data from a nationally representative survey of U.S. adults ($N = 1{,}303$), we investigate the "Greta Thunberg Effect," or whether exposure to Greta Thunberg predicts collective efficacy and intentions to engage in collective action. We find that those who are more familiar with Greta Thunberg have higher intentions of taking collective actions to reduce global warming and that stronger collective efficacy beliefs mediate this relationship. This association between familiarity with Greta Thunberg, collective ef-

the Year by some, and asked to "work on her anger management issues" by others (Alter et al., 2019; McCarthy, 2019). The present study, to date, is one of the first to present empirical evidence supporting the "Greta Thunberg Effect," and to offer a potential explanation of why a young leader could be a powerful influence on collective action. We find that familiarity with Greta Thunberg is


A: Age
B: Political Ideology
Collective Efficacy Beliefs
Familiarity with Greta Thunberg
Collective Action Intentions

What do you think are the biggest issues here?

Khoa Vu
@KhoaVuUmn

"The Greta effect" Effect: Your misuse of causal language is never too wrong to make famous people retweet your study.

> **Hillary Clinton** @HillaryClinton · Jan 28
> Data proving @GretaThunberg right—"you are never too small to make a difference." twitter.com/GeoffreySupran...

11:22 AM · Jan 29, 2021 · Twitter Web App

What about other topics?

Before we start...

Be clear about your language

Before we start...

Be clear about your language

Be clear about your data

Before we start...

Be clear about your language

Be clear about your data

Be clear about your assumptions

What is Causal Inference?

Inferring the effect of one thing on another thing

What is Causal Inference?

Inferring the effect of one thing on another thing

- "My headache went away because I took an aspirin".

What is Causal Inference?

Inferring the effect of one thing on another thing

- "My headache went away because I took an aspirin".
- "The new marketing campaign increased our sales by 20%"

What is Causal Inference?

Inferring the effect of one thing on another thing

- "My headache went away because I took an aspirin".
- "The new marketing campaign increased our sales by 20%"
- "Providing students support when filling out FAFSA forms improves college access and completion."

A world of potential (outcomes)

- Under a binary treatment or intervention, there are **two potential worlds**:

A world of potential (outcomes)

- Under a binary treatment or intervention, there are two potential worlds:

- World 1: You take the pill
- World 2: You don't take the pill

A world of potential (outcomes)

- A potential outcome is the outcome under each of these scenarios or "worlds".
    - *There will be one for each path!*

A world of potential (outcomes)

- A **potential outcome** is the outcome under each of these scenarios or "worlds".
    - *There will be one for each path!*
- A priori, each of these scenarios has a *potential outcome*
- A posteriori, I can only observe **at most one of the potential outcomes**

A world of potential (outcomes)

- A **potential outcome** is the outcome under each of these scenarios or "worlds".
  - *There will be one for each path!*
- A priori, each of these scenarios has a *potential outcome*
- A posteriori, I can only observe **at most one of the potential outcomes**

Fundamental Problem of Causal Inference

What are the potential outcomes for our previous example?

Potential Outcomes Examples

- "My headache went away because I took an aspirin".

Outcome if I take an aspirin/ Outcome if I don't take an aspirin

- "The new marketing campaign increased our sales by 20%"

Sales with a marketing campaign/ Sales without a marketing campaign

- "Providing students support when filling out FAFSA forms improves college access and completion."

College access and completion with support/ College access and completion without support

What are some causal questions you might be interested in?

Potential Outcomes Framework

Let's introduce some notation:

- Let Y_i be the observed outcome for unit i (e.g. whether I have a headache or not in an hour).
- Let Z_i be the treatment or intervention (e.g. taking a pill).

Then,

$$Y_i|(Z_i = 1) \triangleq Y_i(1)$$

where $Y_i(1)$ is the potential outcome under treatment.

In the same fashion,

$$Y_i|(Z_i = 0) \triangleq Y_i(0)$$

where $Y_i(0)$ is the potential outcome under control.

Potential Outcomes Framework

This means that we can write the observed outcome as a function of the *potential outcomes*:

$$\rightarrow Y_i = Z_i \cdot Y_i(1) + (1 - Z_i) \cdot Y_i(0)$$

- This definition will be useful because we can see this as a **missing data problem**.

Causal Effects

Individual Causal Effect

$$ICE_i = Y_i(1) - Y_i(0)$$

Causal Effects

Individual Causal Effect

$$ICE_i = Y_i(1) - Y_i(0)$$

Can we ever observe individual causal effects?

Causal Effects

Individual Causal Effect

$$ICE_i = Y_i(1) - Y_i(0)$$

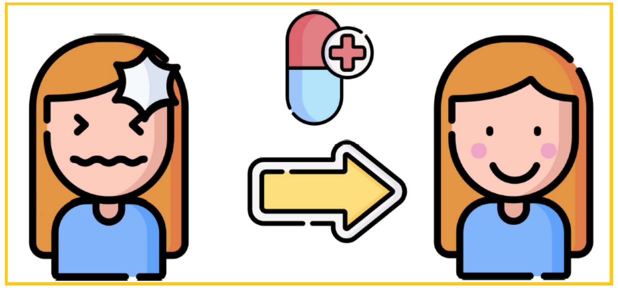
Can we ever observe individual causal effects?

No!*

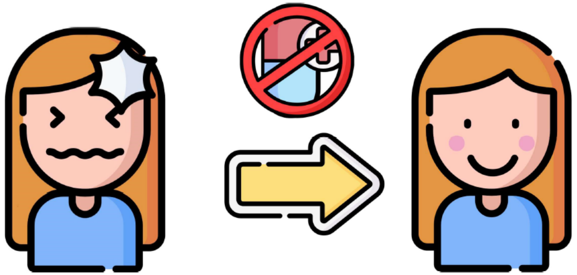
Only one realization

Only one realization

Z=1



Z=0



The "What": Causal estimands, estimates, and estimators

Estimands vs Estimates vs Estimators

Estimand

A quantity we want to estimate

Estimator

A rule for calculating an estimate based on data

Estimate

The result of an estimation

Estimands vs Estimates vs Estimators

Estimand

A quantity we want to estimate

E.g.: Population mean

μ

Estimator

A rule for calculating an estimate based on data

E.g.: Sample mean

$$\frac{1}{n}\sum_i Y_i$$

Estimate

The result of an estimation

E.g.: Result of the sample mean for a given sample S

$\hat{\mu}$

Estimands vs Estimates vs Estimators

Source: Deng, 2022

Estimands vs Estimates vs Estimators

- Some important **estimands** that we need to keep in mind:

Average Treatment Effect (ATE)

Average Treatment Effect on the Treated (ATT)

Conditional Average Treatment Effect (CATE)

Estimands vs Estimates vs Estimators

- Some important estimands that we need to keep in mind:

$$ATE = E[Y(1) - Y(0)]$$

$$ATT = E[Y(1) - Y(0)|Z = 1]$$

$$CATE = E[Y(1) - Y(0)|X]$$

Getting around the fundamental problem of causal inference

- Let's go back to our original example: **Does a pill help reduce headaches?**

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1	?	1	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	0	0	?	?

Getting around the fundamental problem of causal inference

- We have a **missing data problem**

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1	?	1	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	0	0	?	?

Getting around the fundamental problem of causal inference

- Compare those who **took the pill** to the ones **did not take it.**

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1	?	1	?
2	**1**	**1**	**1**	**?**	**?**
3	**1**	**0**	**0**	**?**	**?**
4	0	0	?	0	?
5	0	1	?	1	?
6	**1**	**0**	**0**	**?**	**?**

Getting around the fundamental problem of causal inference

- Compare those who **took the pill** to the ones **did not take it.**

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1	?	1	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	0	0	?	?

Getting around the fundamental problem of causal inference

- Compare those who took the pill to the ones did not take it.

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1	?	1	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	0	0	?	?

$$\hat{\tau} = \frac{1}{3}\left(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i\right) = -0.333$$

Getting around the fundamental problem of causal inference

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i) = -0.333$$

- What is the **estimand**?

Getting around the fundamental problem of causal inference

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i) = -0.333$$

- What is the **estimand**?

Average Treatment Effect

Getting around the fundamental problem of causal inference

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i) = -0.333$$

- What is the estimand?

Average Treatment Effect

- What is the estimator?

Getting around the fundamental problem of causal inference

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i) = -0.333$$

- What is the **estimand**?

Average Treatment Effect

- What is the **estimator**?

Difference in sample means

Getting around the fundamental problem of causal inference

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i) = -0.333$$

- What is the estimand?

Average Treatment Effect

- What is the estimator?

Difference in sample means

- What is the estimate and *how do we interpret it*?

Getting around the fundamental problem of causal inference

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i) = -0.333$$

- What is the estimand?

Average Treatment Effect

- What is the estimator?

Difference in sample means

- What is the estimate and *how do we interpret it*?

33.3 percentage point decrease in probability of having a headache

What could be the problem with comparing the sample means?

The "Why": Causal questions and study designs

Under what assumptions is our estimate causal?

We are using:

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i)$$

to estimate:

$$\tau = E[Y_i(1) - Y_i(0)]$$

Under what assumptions is our estimate causal?

We are using:

$$\hat{\tau} = \frac{1}{3}(\sum_{i \in Z=1} Y_i - \sum_{i \in Z=0} Y_i)$$

to estimate:

$$\tau = E[Y_i(1) - Y_i(0)]$$

Let's do some math

Under what assumptions is our estimate causal?

$$\tau = E[Y_i(1) - Y_i(0)]$$

$$= E[Y_i(1)] - E[Y_i(0)]$$

Key assumption:

Ignorability

- Ignorability means that the potential outcomes $Y(0)$ and $Y(1)$ are independent of the treatment, e.g. $(Y(0), Y(1)) \perp\!\!\!\perp Z$.

Under what assumptions is our estimate causal?

$$\tau = E[Y_i(1) - Y_i(0)]$$

$$= E[Y_i(1)] - E[Y_i(0)]$$

Key assumption:

Ignorability

- Ignorability means that the potential outcomes $Y(0)$ and $Y(1)$ are independent of the treatment, e.g. $(Y(0), Y(1)) \perp\!\!\!\perp Z$.
    - *Remember that if* $A \perp\!\!\!\perp B \;\rightarrow\; E[A|B] = E[A]$

Under what assumptions is our estimate causal?

$$\tau = E[Y_i(1) - Y_i(0)]$$

$$= E[Y_i(1)] - E[Y_i(0)]$$

Key assumption:

Ignorability

- Ignorability means that the potential outcomes $Y(0)$ and $Y(1)$ are independent of the treatment, e.g. $(Y(0), Y(1)) \perp\!\!\!\perp Z$.

$$\tau = E[Y_i(1)] - E[Y_i(0)] = E[Y_i(1)|Z=1] - E[Y_i(0)|Z=0]$$

Under what assumptions is our estimate causal?

$$\tau = E[Y_i(1) - Y_i(0)]$$

$$= E[Y_i(1)] - E[Y_i(0)]$$

Key assumption:

Ignorability

- Ignorability means that the potential outcomes $Y(0)$ and $Y(1)$ are independent of the treatment, e.g. $(Y(0), Y(1)) \perp\!\!\!\perp Z$.

$$\tau = E[Y_i(1)] - E[Y_i(0)] = \underbrace{E[Y_i(1)|Z=1]}_{\text{Obs. Outcome for T}} - \overbrace{E[Y_i(0)|Z=0]}^{\text{Obs. Outcome for C}}$$

Under what assumptions is our estimate causal?

$$\tau = E[Y_i(1) - Y_i(0)]$$

$$= E[Y_i(1)] - E[Y_i(0)]$$

Key assumption:

Ignorability

- Ignorability means that the potential outcomes $Y(0)$ and $Y(1)$ are independent of the treatment, e.g. $(Y(0), Y(1)) \perp\!\!\!\perp Z$.

$$\tau = E[Y_i(1)] - E[Y_i(0)] = E[Y_i(1)|Z=1] - E[Y_i(0)|Z=0] =$$

$$= E[Y_i|Z=1] - E[Y_i|Z=0]$$

Ignorability Assumption

We can just "ignore" the missing data problem:

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1		1	
2	1	1	1		
3	1	0	0		
4	0	0		0	
5	0	1		1	
6	1	0	0		

Ignorability Assumption

We can just "ignore" the missing data problem:

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1		1	
2	1	1	1		
3	1	0	0		
4	0	0		0	
5	0	1		1	
6	1	0	0		

Ignorability Assumption

We can just "ignore" the missing data problem:

i	Z	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	1		1	
2	1	1	1		
3	1	0	0		
4	0	0		0	
5	0	1		1	
6	1	0	0		
			1/3	2/3	

Let's do a little exercise

Look at your **green** piece of paper and go to the following website

Would you go to a physician/urgent care?

Now let's assume I randomly allocate whether you go or not go to the hospital

Do the previous results hold?

Randomization is an awesome tool for causal inference

Main takeaway points

Causal Inference is hard

Main takeaway points

Causal Inference is hard

- Think about the causal problem

Main takeaway points

Causal Inference is hard

- Think about the causal problem
- Always look at your data

Main takeaway points

Causal Inference is hard

- Think about the causal problem
- Always look at your data
- Check validity of assumptions (*Is ignorability plausible? Am I controlling for the right covariates?*)

Main takeaway points

Causal Inference is hard

- Think about the **causal problem**
- Always **look at your data**
- Check **validity** of assumptions (*Is ignorability plausible? Am I controlling for the right covariates?*)
- Most of this chapter will be spent on looking for **exogeneous variation** to make the ignorability assumption happen.

Next week

- Randomized Controlled Trials:
  - Pros and Cons
  - Concept of validity
  - A/B Testing

References

- Angrist, J. & S. Pischke. (2015). "Mastering Metrics". *Chapter 1*.
- Cunningham, S. (2021). "Causal Inference: The Mixtape". *Chapter 4: Potential Outcomes Causal Model*.
- Neil, B. (2020). "Introduction to Causal Inference". *Fall 2020 Course*