

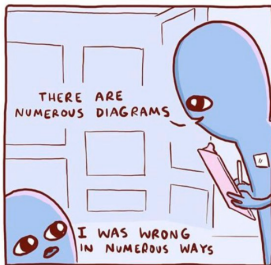
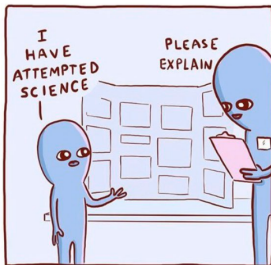
Research Club Talk
Maison de L'Inde

Research Biases in Social Sciences:

An overview of common cases and how to avoid them

by Thiago Scarelli
PhD candidate, Paris School of Economics

6 March 2020, 21:00
Salle APJ Abdul Kalam



NATHANWPYLE

What are we going to talk about?

- ▶ Biases that affect research outcomes.

What are we going to talk about?

- ▶ Biases that affect research outcomes.

What are we NOT going to talk about:

- ▶ Behavioral or cognitive biases per se.

What are we going to talk about?

- ▶ Biases that affect research outcomes.

What are we NOT going to talk about:

- ▶ Behavioral or cognitive biases per se.

This is a methodological discussion. Who should care about it?

A general concept

$$\mathbb{E}(S) = \theta + \text{bias}$$

A general concept

$$\mathbb{E}(S) = \theta + \text{bias}$$

Bias or statistical noise?

A. When we fail to disentangle confounding factors

Fundamental condition to claim that X is directly related to Y :
there is no other hidden factor driving X and Y at the same time.

If that condition does not hold, we have an **omitted variable bias**.

A. When we fail to disentangle confounding factors

Fundamental condition to claim that X is directly related to Y:
there is no other hidden factor driving X and Y at the same time.

If that condition does not hold, we have an **omitted variable bias**.

- Does birth order affect the odds of Down syndrome?

A. When we fail to disentangle confounding factors

Fundamental condition to claim that X is directly related to Y:
there is no other hidden factor driving X and Y at the same time.

If that condition does not hold, we have an **omitted variable bias**.

- ▶ Does birth order affect the odds of Down syndrome?
- ▶ Does impatient kids have worse grades?

A. When we fail to disentangle confounding factors

Fundamental condition to claim that X is directly related to Y:
there is no other hidden factor driving X and Y at the same time.

If that condition does not hold, we have an **omitted variable bias**.

- ▶ Does birth order affect the odds of Down syndrome?
- ▶ Does impatient kids have worse grades?
- ▶ Are alternative teaching methods better than the traditional one?

A. How to avoid omitted variable bias?

1. **Multivariate statistics:** Explicitly account for the omitted variables.

A. How to avoid omitted variable bias?

- I. Multivariate statistics:** Explicitly account for the omitted variables.
- II. Instrumental variable:** Break the link with omitted variables by exploring factors that induce changes in X without affecting the outcome Y .

A. How to avoid omitted variable bias?

- I. **Multivariate statistics:** Explicitly account for the omitted variables.
- II. **Instrumental variable:** Break the link with omitted variables by exploring factors that induce changes in X without affecting the outcome Y.
- III. **Experiment:** Break the link with omitted variables by **randomization**.
 1. **Natural experiments:** Explore a situation where X was applied as an unintended consequence of an external event, and then observe Y.
 2. **Field or lab experiments:** Artificially apply X and then observe Y.

B. When randomization does not hold

Fundamental condition to claim that the experiment is eliciting the true association between X and Y is that all other factors were randomized.

What can go wrong?

B. When randomization does not hold

Fundamental condition to claim that the experiment is eliciting the true association between X and Y is that all other factors were randomized.

What can go wrong? Some individuals are more likely to be observed.

I. Selection bias, sampling bias

B. When randomization does not hold

Fundamental condition to claim that the experiment is eliciting the true association between X and Y is that all other factors were randomized.

What can go wrong? Some individuals are more likely to be observed.

I. Selection bias, sampling bias

II. Self-selection, volunteering, non-response, refusal

B. When randomization does not hold

Fundamental condition to claim that the experiment is eliciting the true association between X and Y is that all other factors were randomized.

What can go wrong? Some individuals are more likely to be observed.

I. Selection bias, sampling bias

II. Self-selection, volunteering, non-response, refusal

III. Loss to follow-up, transfer bias, migration bias, attrition bias

B. When randomization does not hold

Fundamental condition to claim that the experiment is eliciting the true association between X and Y is that all other factors were randomized.

What can go wrong? Some individuals are more likely to be observed.

I. Selection bias, sampling bias

II. Self-selection, volunteering, non-response, refusal

III. Loss to follow-up, transfer bias, migration bias, attrition bias

- ▶ Does training improve job market outcomes for the unemployed?

C. When expectations drive the findings

Even if randomization holds *ex ante*, being part of an experiment in itself may affect the behavior of the different individuals involved in it.

C. When expectations drive the findings

Even if randomization holds *ex ante*, being part of an experiment in itself may affect the behavior of the different individuals involved in it.

Biases can emerge from the change in the behavior of the **subjects**:

1. Hawthorne effect, observer effect

C. When expectations drive the findings

Even if randomization holds *ex ante*, being part of an experiment in itself may affect the behavior of the different individuals involved in it.

Biases can emerge from the change in the behavior of the **subjects**:

I. Hawthorne effect, observer effect

II. Placebo effect

C. When expectations drive the findings

Even if randomization holds *ex ante*, being part of an experiment in itself may affect the behavior of the different individuals involved in it.

Biases can emerge from the change in the behavior of the **subjects**:

I. Hawthorne effect, observer effect

II. Placebo effect

III. John Henry effect

C. When expectations drive the findings

Even if randomization holds *ex ante*, being part of an experiment in itself may affect the behavior of the different individuals involved in it.

Biases can emerge from the behavior of **researchers** or **collaborators**:

IV. Pigmalion Effect

C. When expectations drive the findings

Even if randomization holds *ex ante*, being part of an experiment in itself may affect the behavior of the different individuals involved in it.

Biases can emerge from the behavior of **researchers** or **collaborators**:

IV. Pigmalion Effect

V. Reporting bias, measurement bias, detection bias

C. How to avoid bias from participants behaviour?

I. Blinding

Subjects don't know their status.

C. How to avoid bias from participants behaviour?

I. Blinding

Subjects don't know their status.

II. Double blinding

Subjects *and* researchers don't know subject's status.

C. How to avoid bias from participants behaviour?

I. Blinding

Subjects don't know their status.

II. Double blinding

Subjects *and* researchers don't know subject's status.

III. Triple blinding

Subjects, researchers *and* collaborators don't know subject's status.

C. When the literature does not tell the whole story

When assessing evidence and exploring alternative causes, we might be biased by the fact that some results are more likely to be seen than others.

I. Language bias

C. When the literature does not tell the whole story

When assessing evidence and exploring alternative causes, we might be biased by the fact that some results are more likely to be seen than others.

I. Language bias

II. Publication bias, citation bias, outcome bias, file-drawer effect

Summary

Failures in the methodological assumptions can create systematic deviations in the statistical measurement of interest.

In a simplified (non-exhaustive) scheme, biases can be due to:

1. omitted factors;
2. randomization failures;
3. expectations of the agents involved in the study;
4. literature availability.

Takeaways

Key message to research readers:

- ▶ Be critical and vigilant when assessing the available evidence.

Takeaways

Key message to research readers:

- ▶ Be critical and vigilant when assessing the available evidence.

Key message to researchers:

- ▶ *ex ante*, try to anticipate those effects whenever possible;
- ▶ *ex post*, be transparent about the likely direction of any bias that the research design could not mitigate.

Takeaways

Key message to research readers:

- ▶ Be critical and vigilant when assessing the available evidence.

Key message to researchers:

- ▶ *ex ante*, try to anticipate those effects whenever possible;
- ▶ *ex post*, be transparent about the likely direction of any bias that the research design could not mitigate.

Revision is not a bug the scientific method — is its defining feature.