

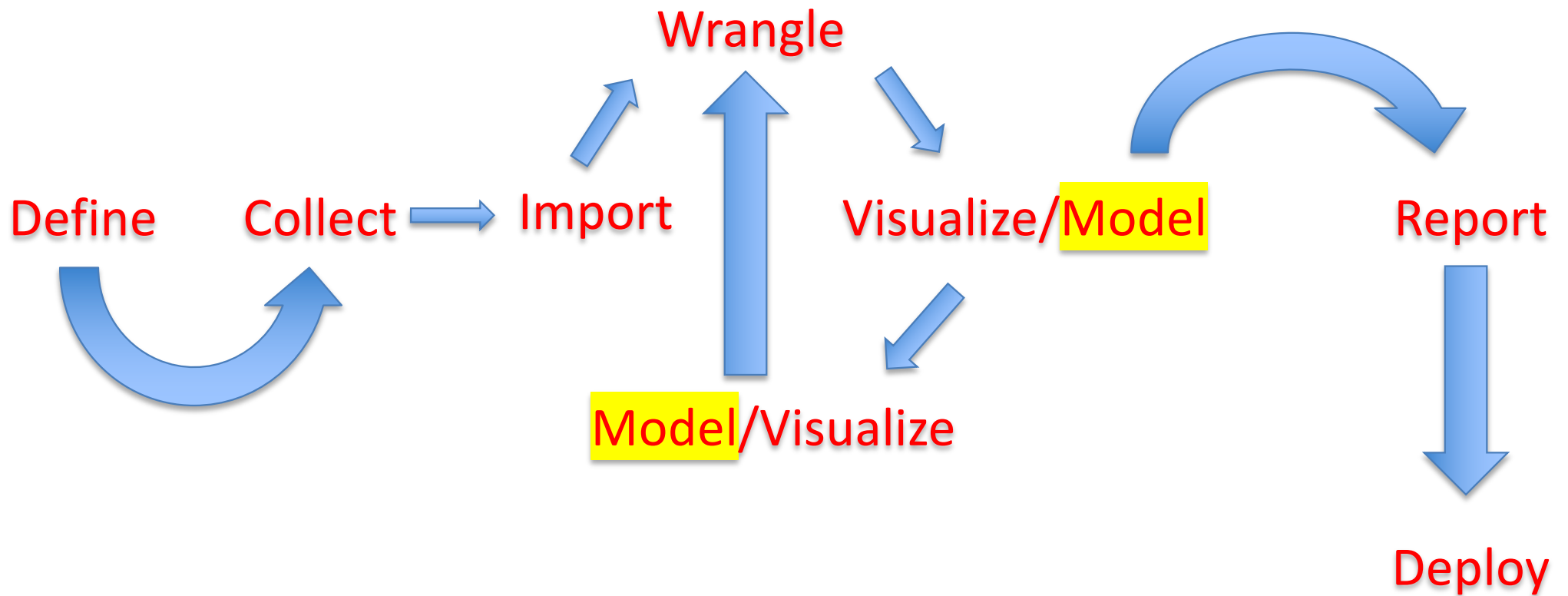
Financial Analytics

Lesson 04 | Randomized Controlled Trials (RCTs) and Matching
for Causal Inference

Prof. SUNGJONG ROH, PhD

AY 2024-2025 April-Term | QF627 Extras | 2025-05-06

Workflow of Data-Driven Decision Making





How to make a confident claim of X leads to Y
(Understanding randomized experiment, a.k.a., A/B RCTs)

Elements of Causal Effects



Three Necessary (NOT sufficient) Conditions of Causality

Three Necessary (NOT sufficient) Conditions of Causality

1. Correlation

Three Necessary (NOT sufficient) Conditions of Causality

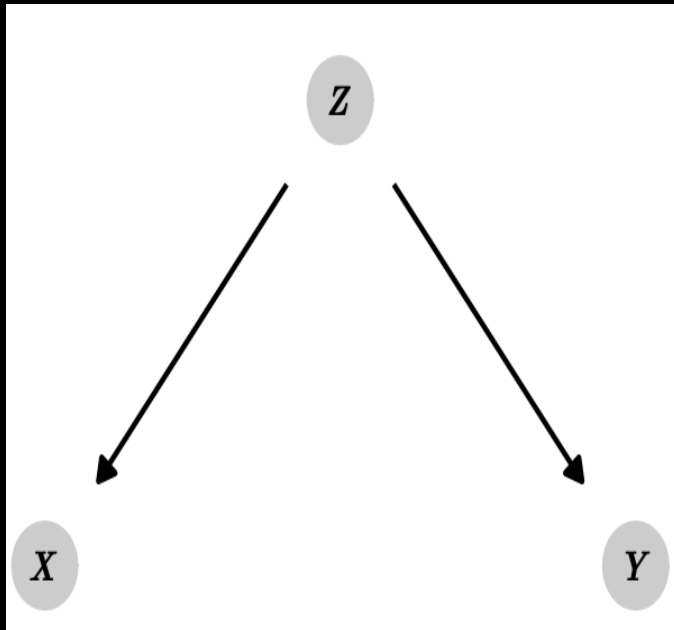
1. Correlation

2. Time Lag (Temporal Precedence)

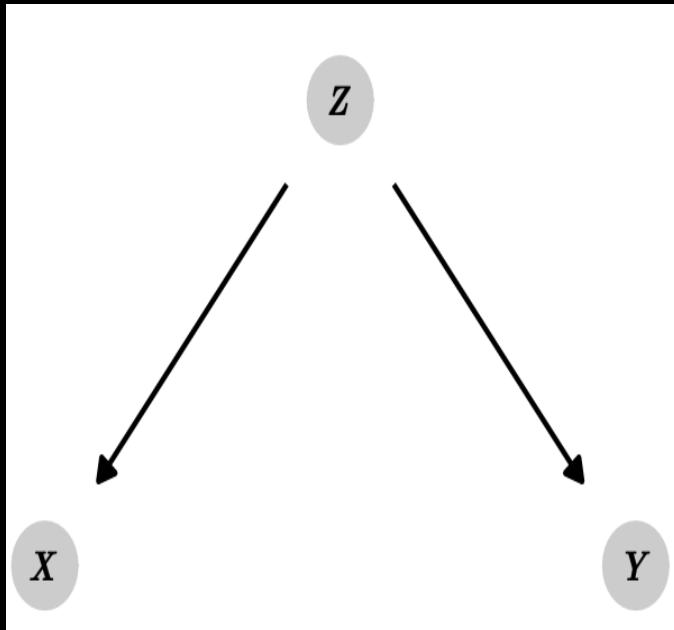
Three Necessary (NOT sufficient) Conditions of Causality

1. Correlation
2. Time Lag (Temporal Precedence)
3. Non-Spuriousness
(Addressing **Confounders**)

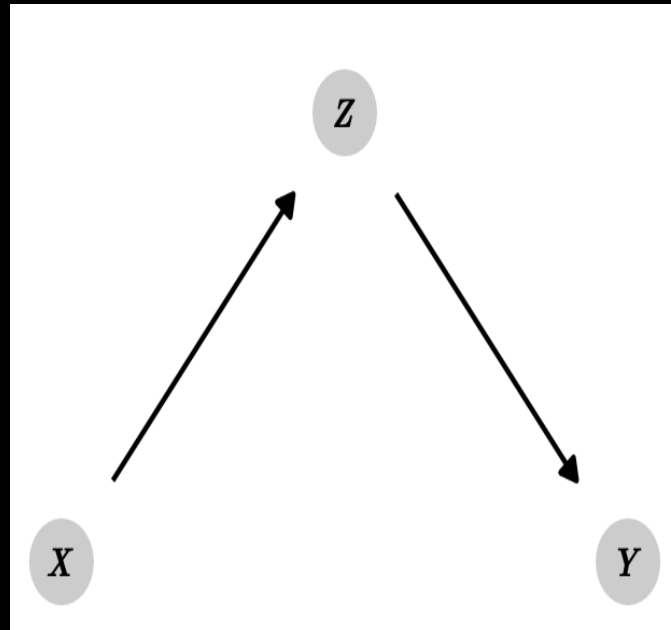
Confounder



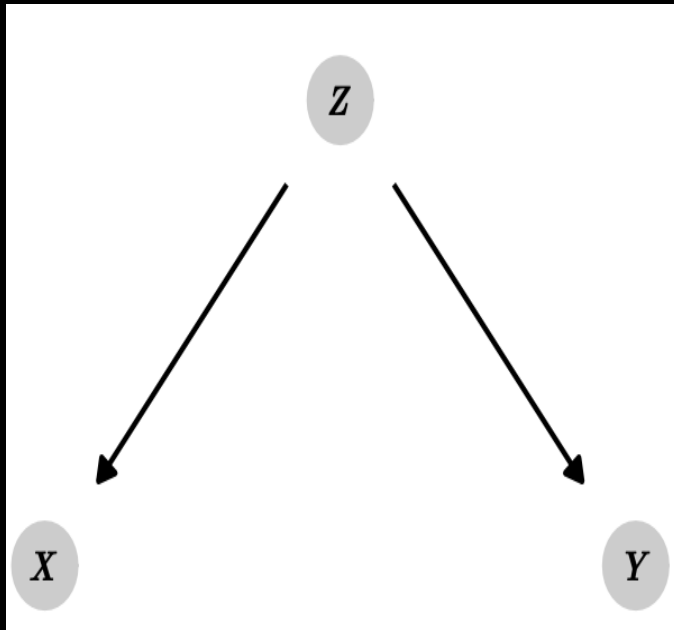
Confounder



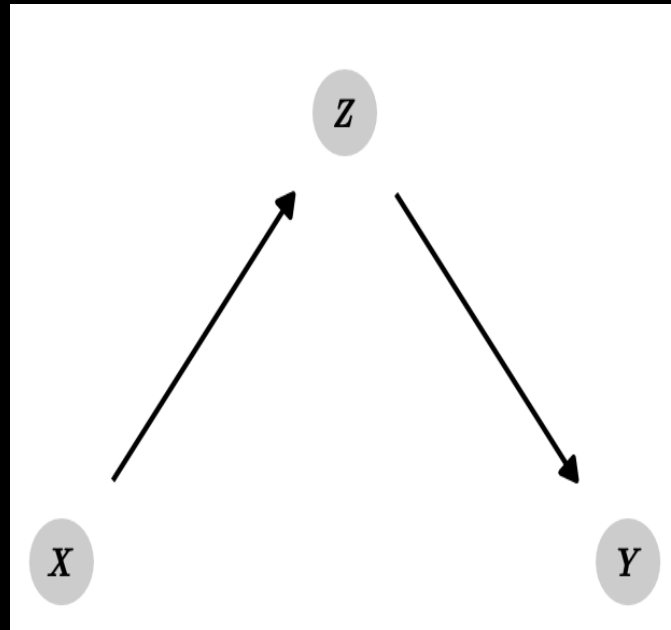
Mediator



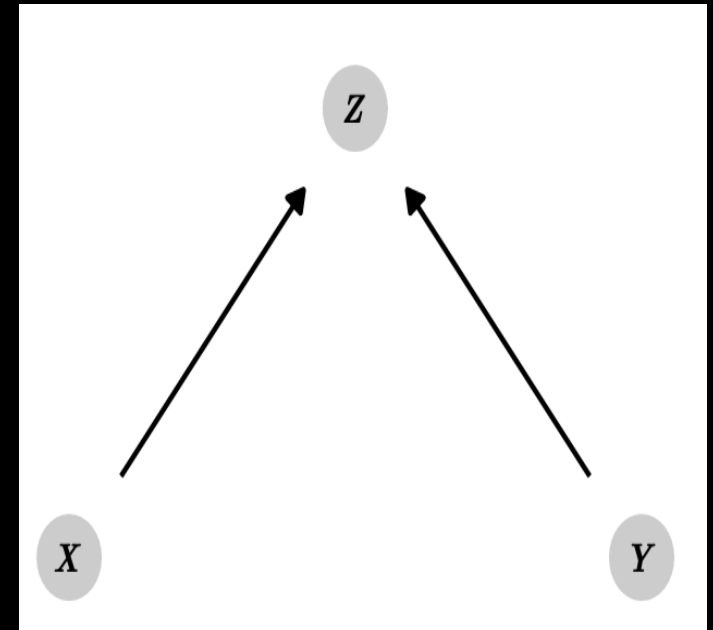
Confounder



Mediator



Collider



Data Collection Methods (Study Design)

Observational Studies

Collect data in a way that does not directly interfere with how the data arise (“observe”)

Only establish an association

Experimental Studies

Randomly assign subjects to treatments

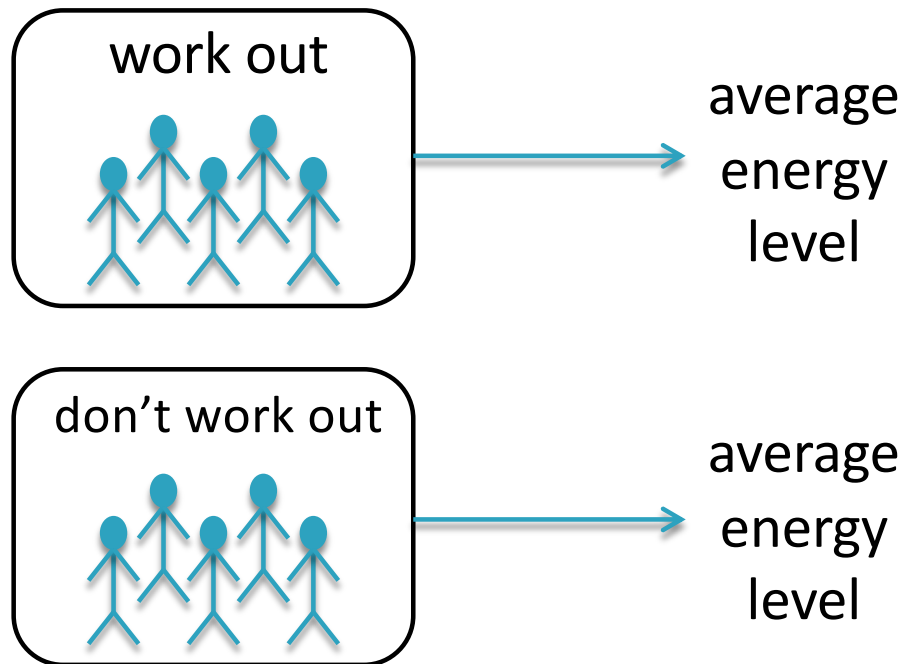
Establish causal connections

Q: Evaluate the relationship between regularly working out and energy level

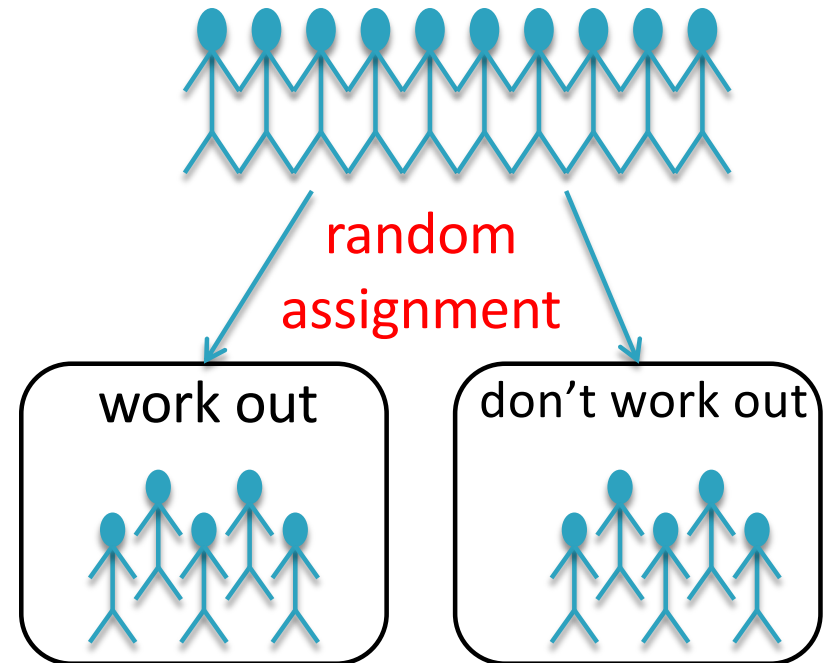
Random Assignment

Q: Evaluate the relationship between regularly working out and energy level

Observational Studies



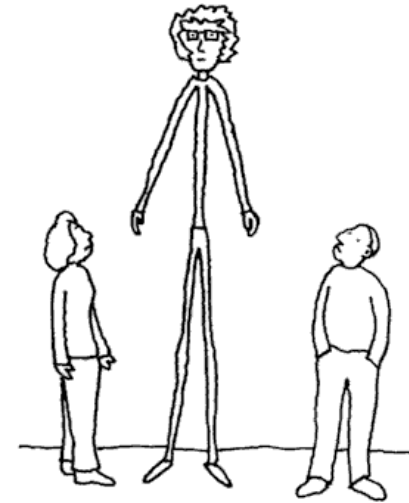
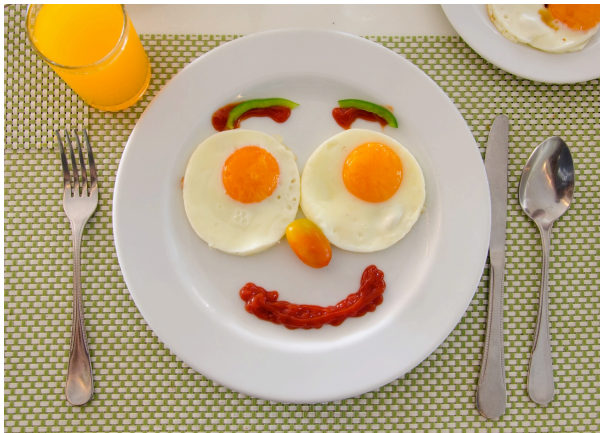
Experimental Studies



Let's do some exercise.

3 Plausible Explanations

Eating breakfast causes girls to be slimmer

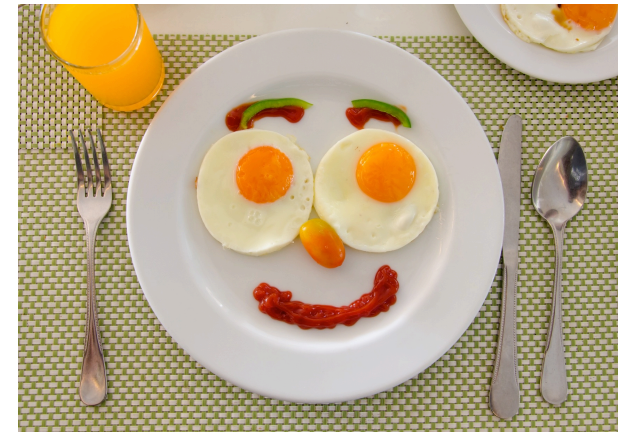
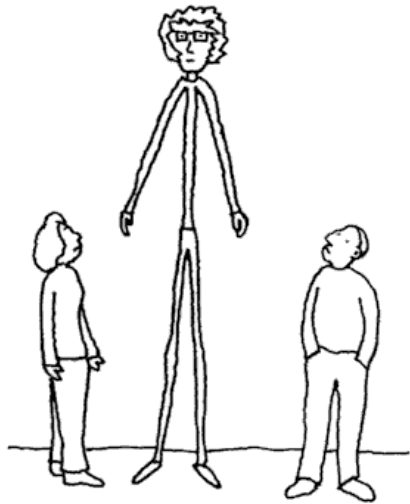


<http://pas-wordpress-media.s3.amazonaws.com/wp-content/uploads/2014/03/Culture-Eats-Strategy-For-Breakfast.jpg>

http://2.bp.blogspot.com/-b_7DdJVRo_A/UsUGJfKRfBI/AAAAAAAAACWg/vzKksyjLvTI/s1600/me-tall.gif

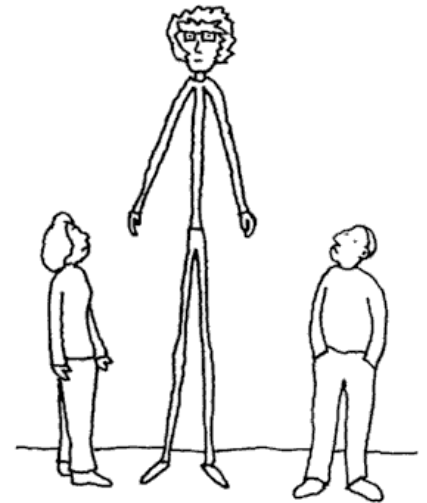
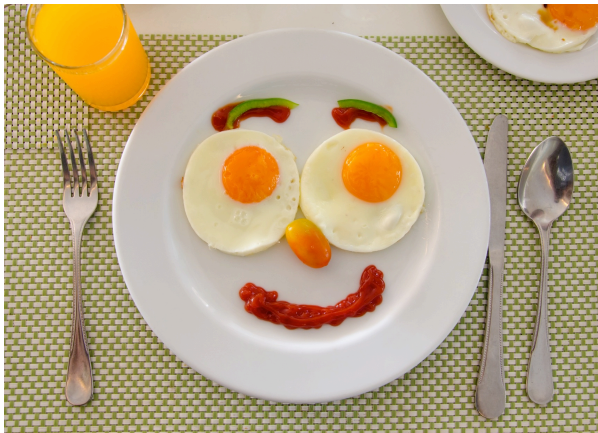
3 Plausible Explanations

Being slim causes girls to eat breakfast



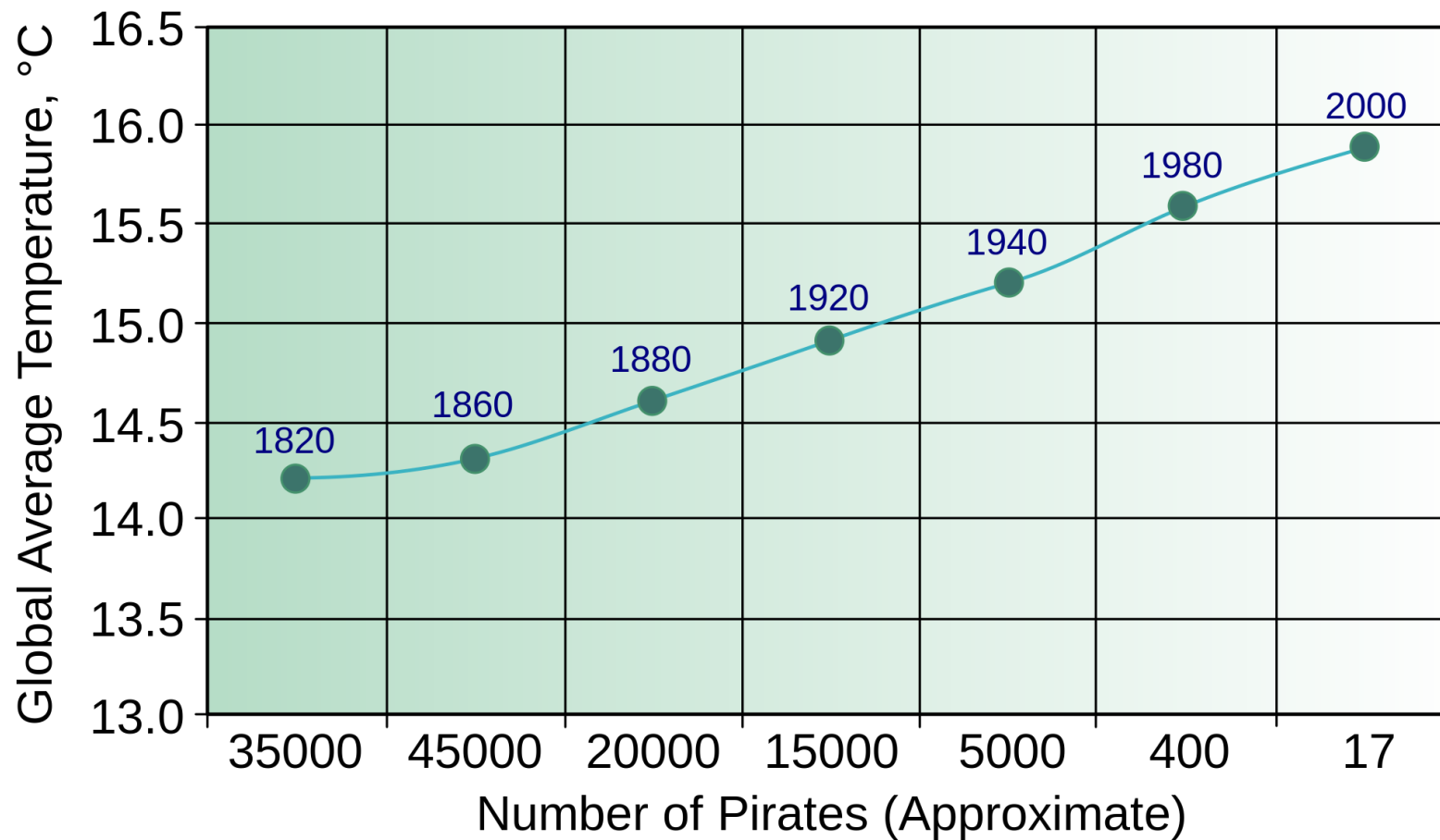
Extraneous variables that affect both the independent and the dependent variable, and that make it seem like there is a relationship between them

Confounding Variables



Correlation Does NOT Imply Causation

Global Average Temperature vs. Number of Pirates

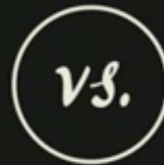


Some executives still mistakenly believe that all they need to do is establish correlation, and causality can be inferred.

Oh Dear... That's Wrong T.T

Let's do some further exercise.

Web Browser & Job Performance





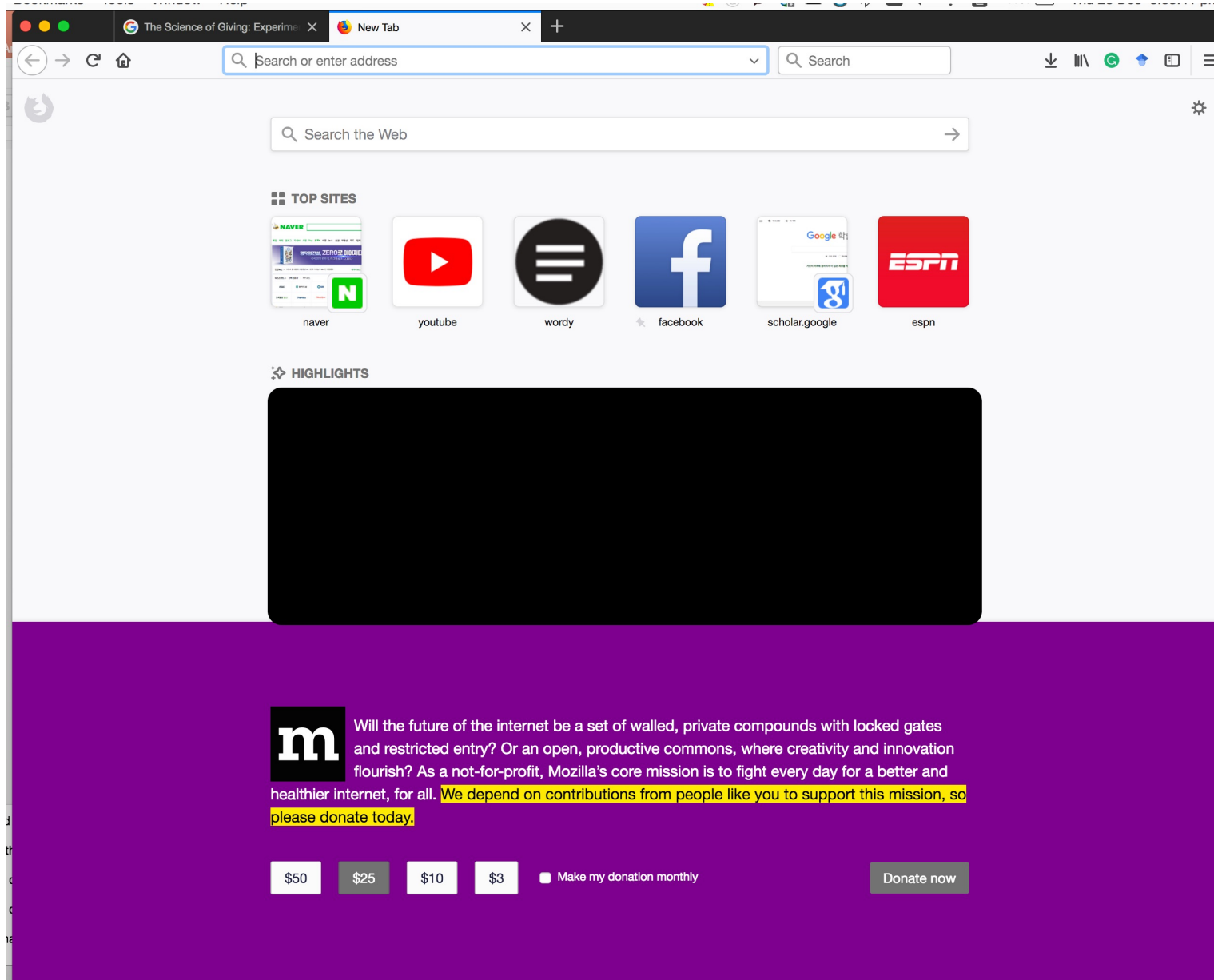
There is good evidence that
Firefox and Chrome users significantly
outperform Internet Explorer and Safari users.



They also stay in their jobs 15 percent longer.



Why? It's about how you got the browser.




BookmarksToolsWindowHelp


R Korea - KRUG(Korean R User)New Tab


Search or enter addressSearch


Search the Web


TOP SITES


naver

youtube


scholar.google

facebook

wordy

espn

HIGHLIGHTS



Big corporations want to restrict entry. Fake news and filter bubbles are making it harder for us to find our way. Online bullies are silencing inspired voices. And our desire to explore is hampered by threats to our safety and privacy. It's time to join Mozilla and do our part as digital citizens. Please donate today to support programs that keep the internet healthy, free and open for us all.

\$50

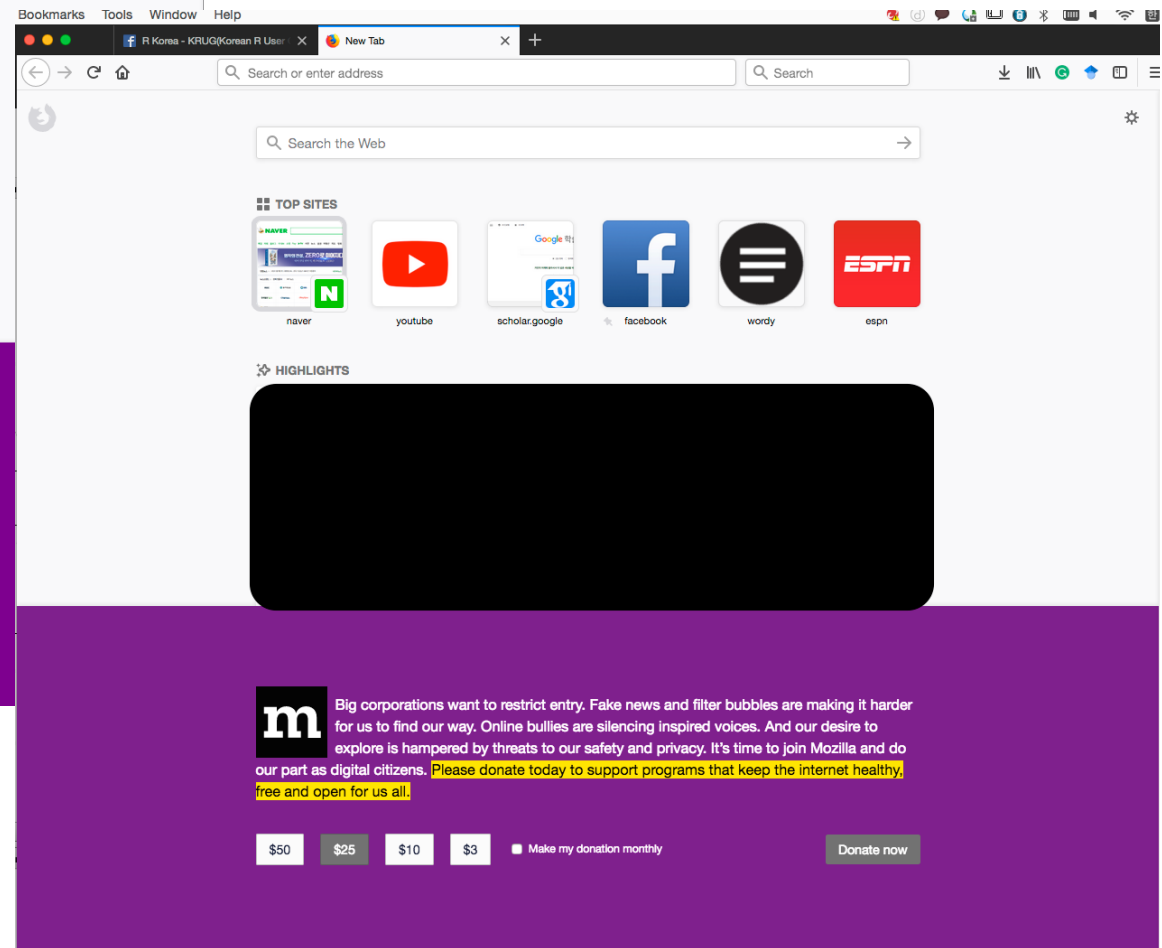
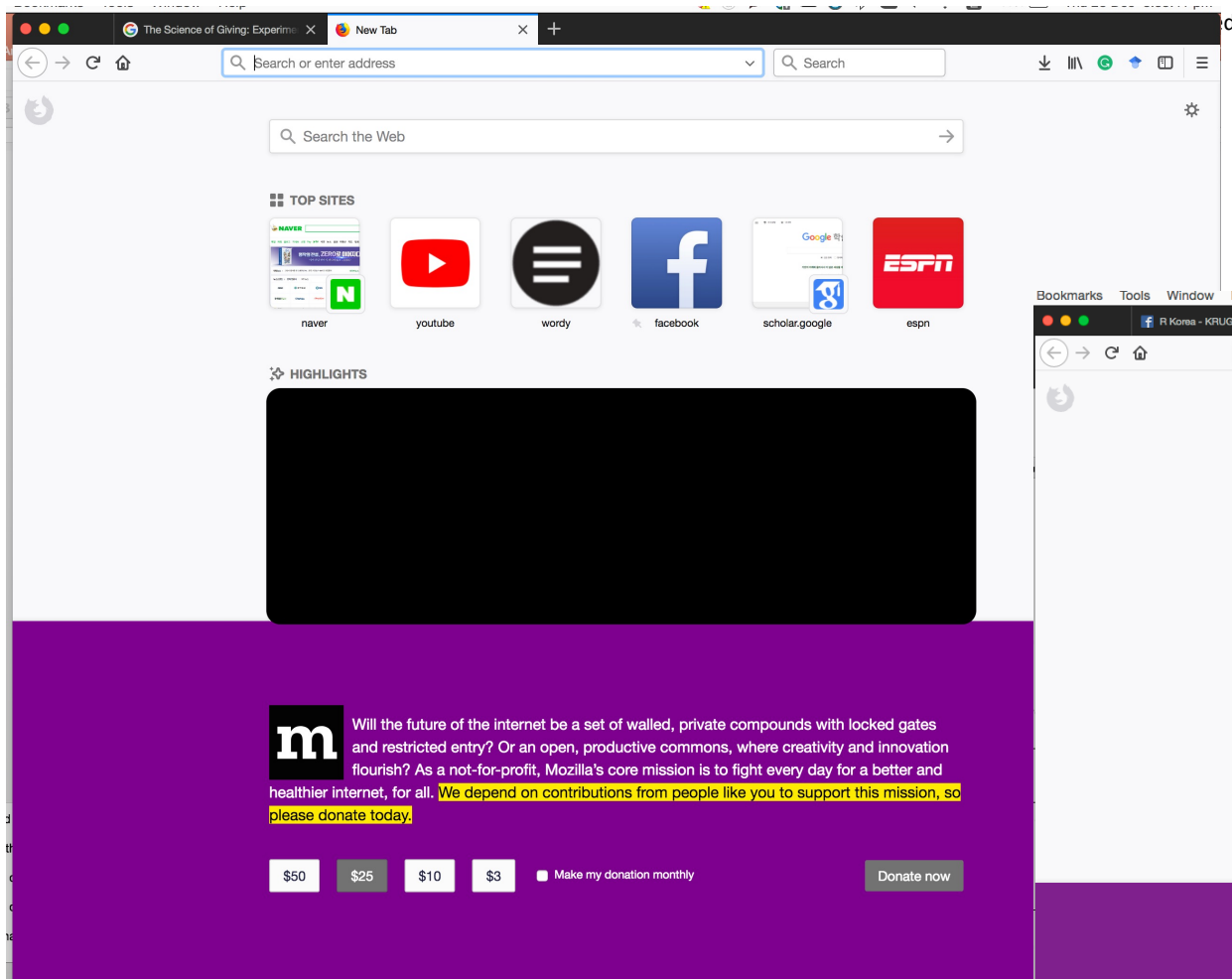
\$25

\$10

\$3

☐ Make my donation monthly

Donate now



“Experiment with Everything” Approach

Today, Microsoft and several other leading companies (including Amazon, Booking.com, Facebook, and Google) each conduct more than 10,000 online controlled experiments annually, with many tests engaging millions of users.

A

B

“Experiment with Everything” Approach

Start-ups and companies without digital roots, such as Walmart, Hertz, and *Singapore Airlines*, also run them regularly, though on a smaller scale.

A

B

“Experiment with Everything” Approach

These organizations have discovered that an “experiment with everything” approach has surprisingly large payoffs.

A

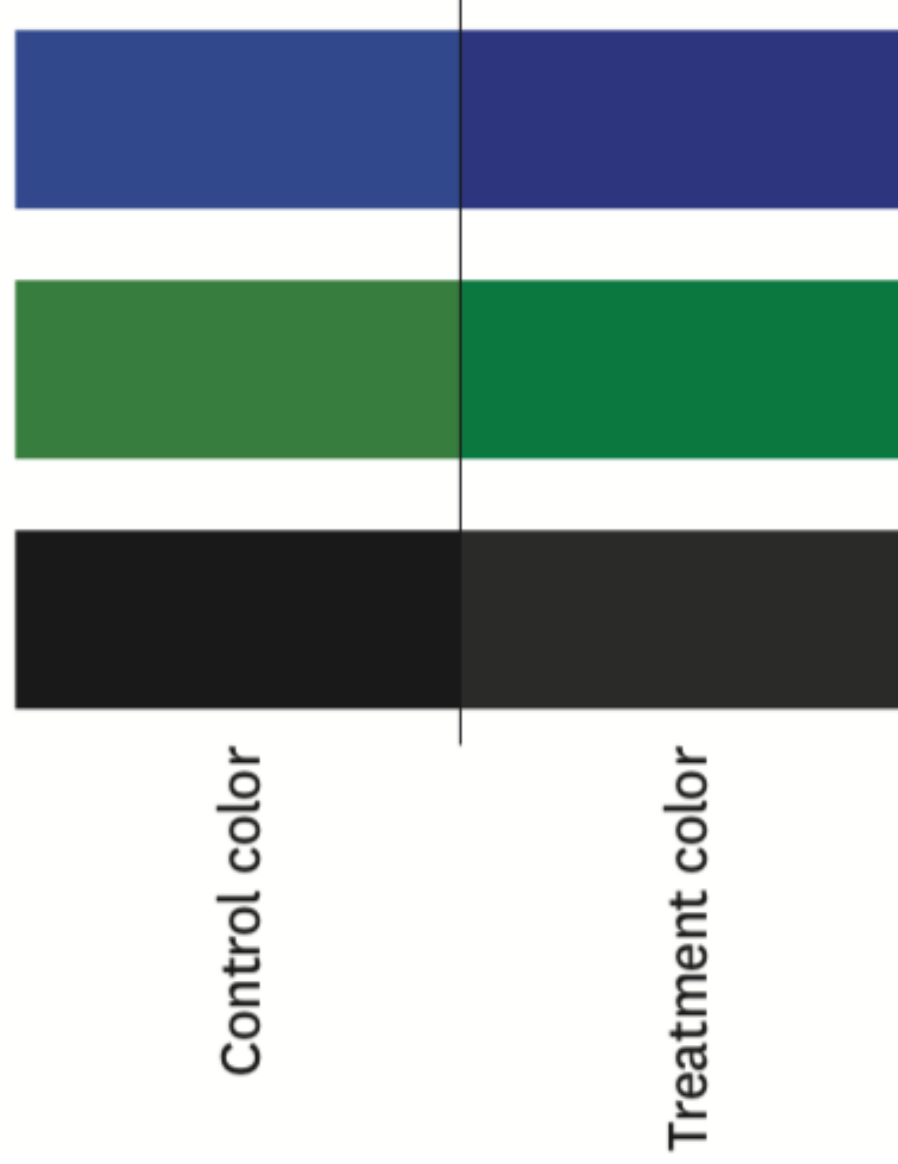
B



Case Study:

Bing's small changes with a HUGE impact

A

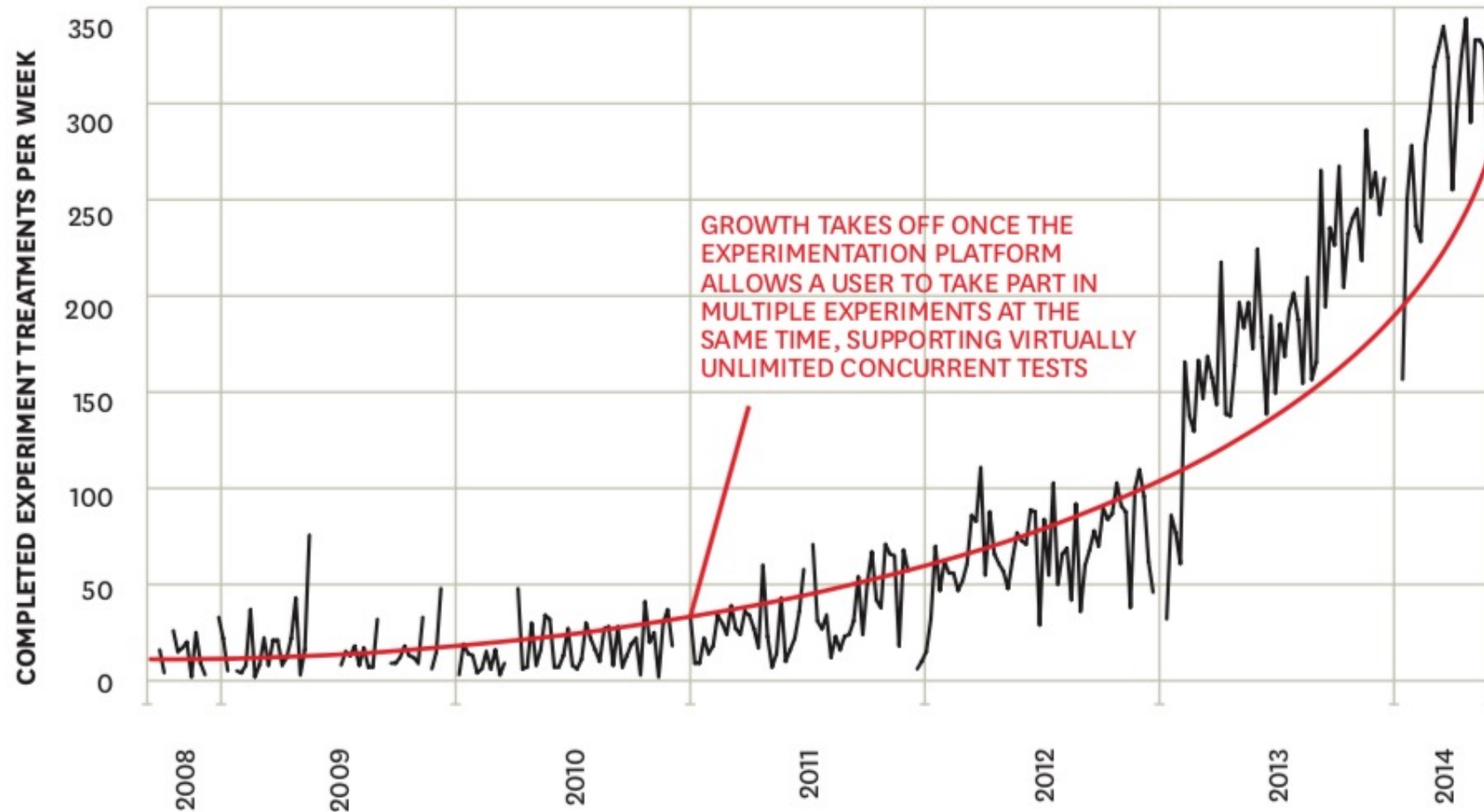


B



- ❑ Bing's experiments showed that slightly darker blues and greens in titles and a slightly lighter black in captions improved the users' experience.
- ❑ When rolled out to all users, the color changes boosted revenue by more than \$10 million annually.

The Growth of Experimentation (A/B Testing) at Bing



A Case from *booking.com*

(Thomke, 2020)

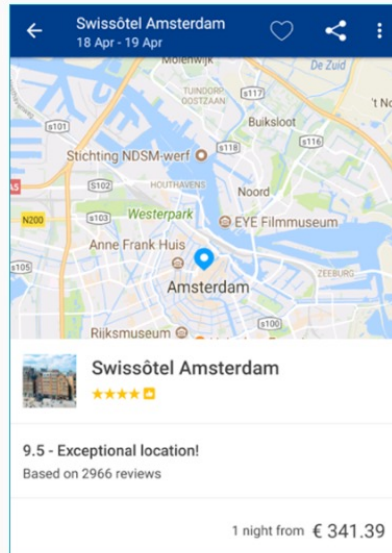
Scenario #1

Hypothesis

Highlighting a neighborhood's walkability helps users make better decisions about property location.

A: The Control

Shows the site's current practice



A Case from *booking.com*

(Thomke, 2020)

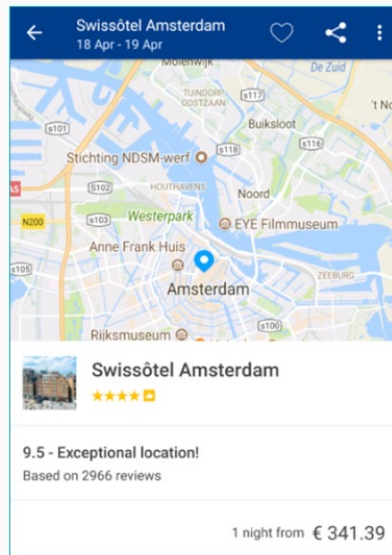
Scenario #1

Hypothesis

Highlighting a neighborhood's walkability helps users make better decisions about property location.

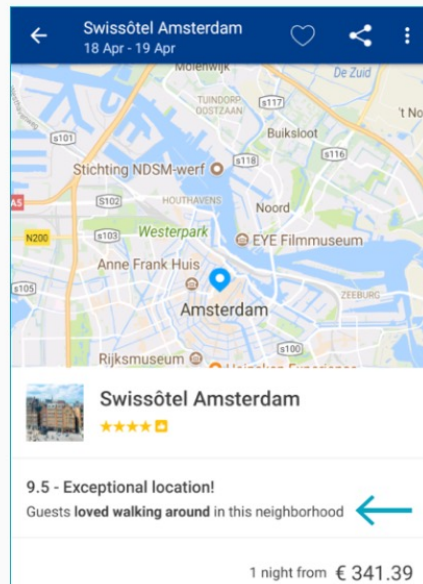
A: The Control

Shows the site's current practice



B: The Treatment

Adds walkability information



A Case from *booking.com*

(Thomke, 2020)

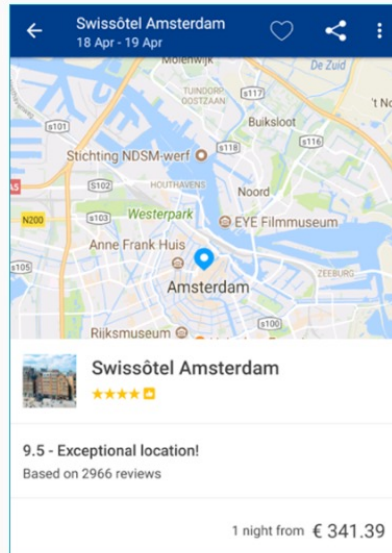
Scenario #1

Hypothesis

Highlighting a neighborhood's walkability helps users make better decisions about property location.

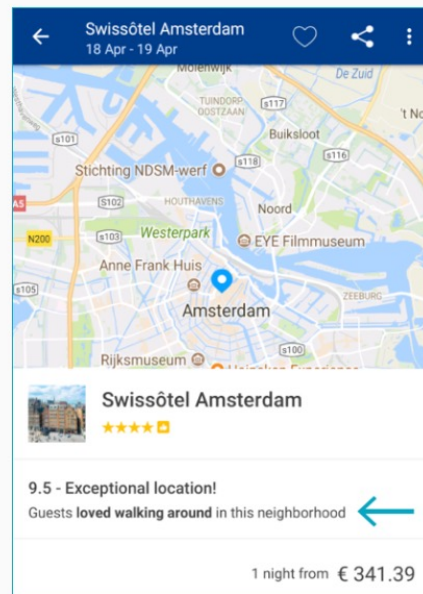
A: The Control

Shows the site's current practice



B: The Treatment

Adds walkability information



The Result

- The treatment had no significant impact on the key metric.
- The current practice is kept in place.

A Case from *booking.com*

(Thomke, 2020)

Scenario #2

Hypothesis

Displaying the checkout date when users select the age of children in their party improves their experience.

A: The Control

Shows the site's current practice

Rooms 1	Adults 2	Children 2
Ages of children at check-out		
4	7	

A Case from *booking.com*

(Thomke, 2020)

Scenario #2

Hypothesis

Displaying the checkout date when users select the age of children in their party improves their experience.

A: The Control

Shows the site's current practice

Rooms 1	Adults 2	Children 2
Ages of children at check-out		
4	7	

B: The Treatment

Adds the checkout date above children's ages

Rooms 1	Adults 2	Children 2
Children's ages on Jul 23, 2016		
4	7	

A Case from *booking.com*

(Thomke, 2020)

The Result

- The treatment had a significant positive impact on the key metric, and the change is implemented.

Scenario #2

Hypothesis

Displaying the checkout date when users select the age of children in their party improves their experience.

A: The Control

Shows the site's current practice

Rooms 1	Adults 2	Children 2
Ages of children at check-out		
4	7	

B: The Treatment

Adds the checkout date above children's ages

Rooms 1	Adults 2	Children 2
Children's ages on Jul 23, 2016		
4	7	

Interventions, A/B Testing, and Causal Inference

- ❑ When creating corporate strategies, many companies make decisions—on everything from new product features, to look and feel, to marketing campaigns— using subjective opinions rather than hard data.

Interventions, A/B Testing, and Causal Inference

- ❑ When creating corporate strategies, many companies make decisions—on everything from new product features, to look and feel, to marketing campaigns— using subjective opinions rather than hard data.
- ❑ Companies should conduct controlled experiments (A/B tests) to evaluate their ideas.

Interventions, A/B Testing, and Causal Inference

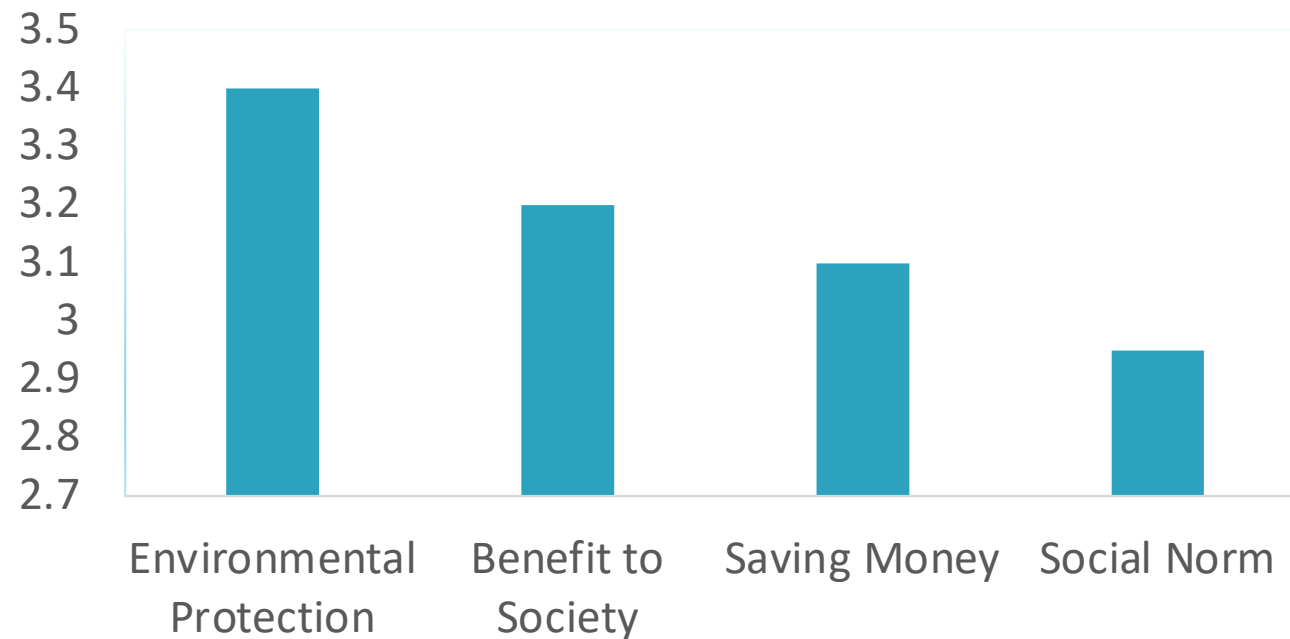
- ❑ When creating corporate strategies, many companies make decisions—on everything from new product features, to look and feel, to marketing campaigns— using subjective opinions rather than hard data.
- ❑ Companies should conduct controlled experiments (A/B tests) to evaluate their ideas.
- ❑ Potential improvements should be rigorously tested, because large investments can fail to deliver, and some tiny changes can be surprisingly detrimental while others have big payoffs.

Interventions, A/B Testing, and Causal Inference

Executives should understand how to properly design and execute A/B tests and other controlled experiments, ensure their integrity, interpret their results, and avoid pitfalls.

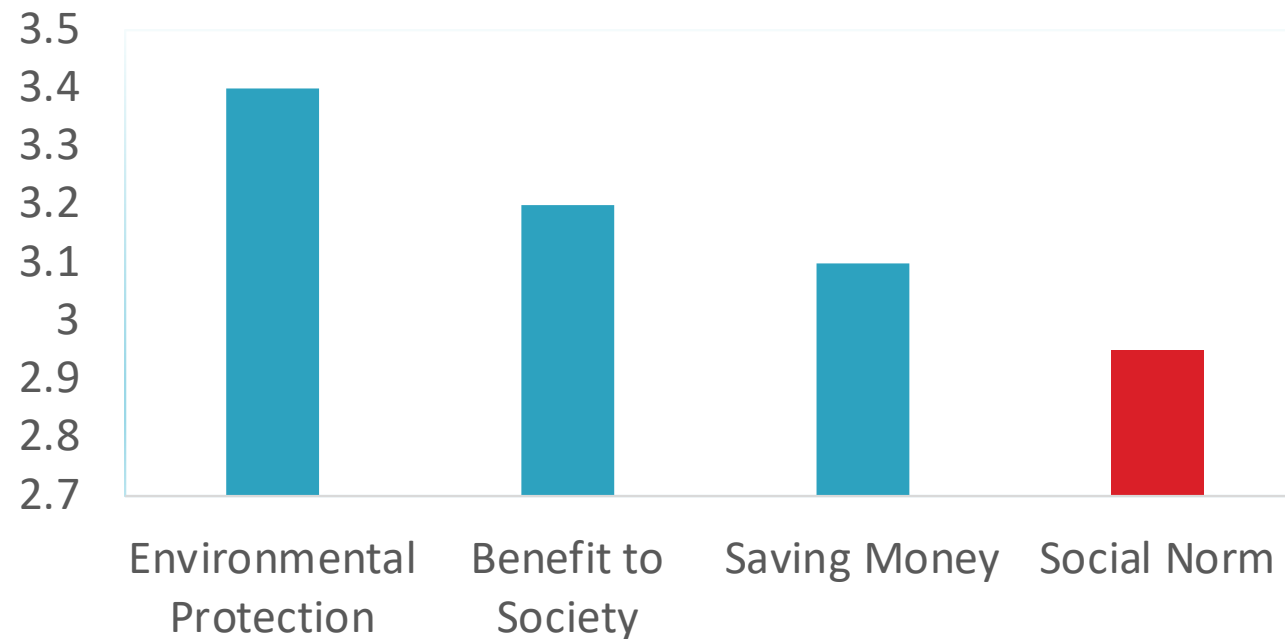
California Energy Savings Survey (Goldstein et al., 2010)

Reported Beliefs Regarding the Influence of Each Motive



California Energy Savings Survey (Goldstein et al., 2010)

Reported Beliefs Regarding the Influence of Each Motive



Does This Perception Match with Reality?

48

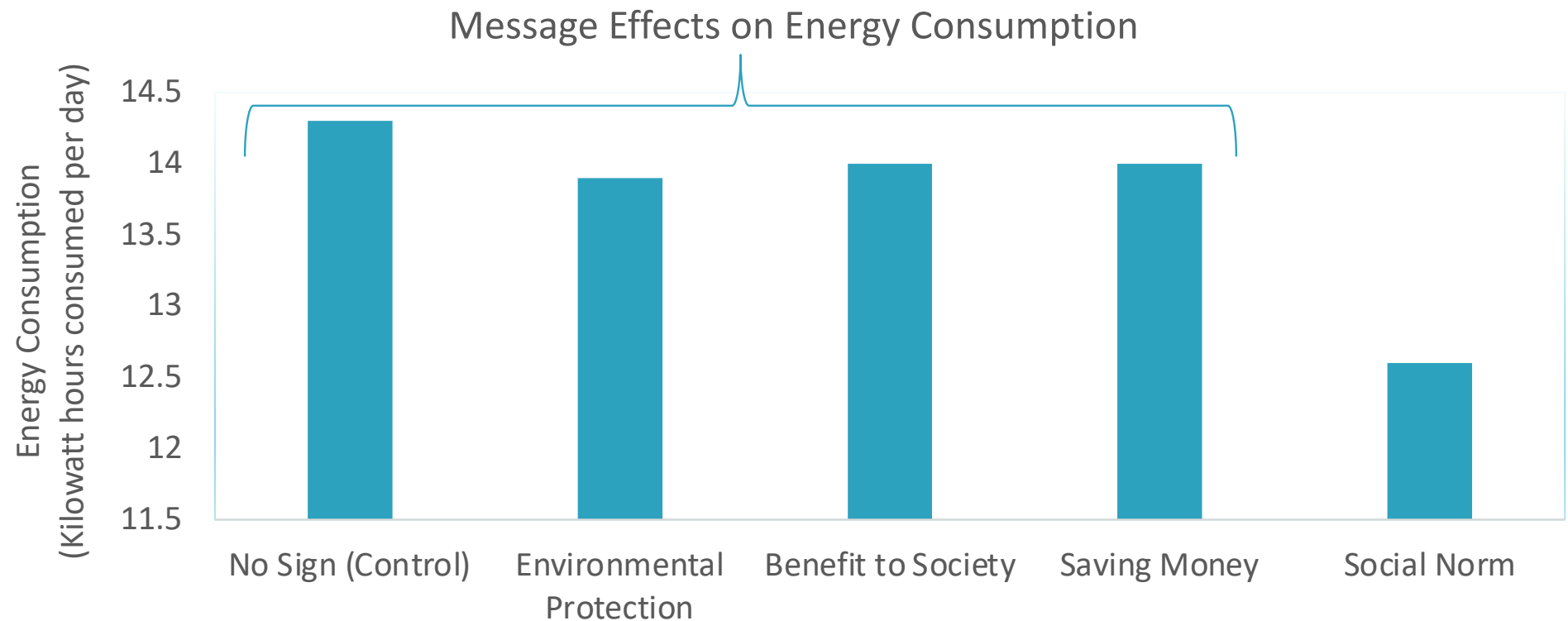
- Researchers then conducted an experiment in a California neighborhood
- They placed signs on residents' door encouraging them to conserve energy

Does This Perception Match with Reality?

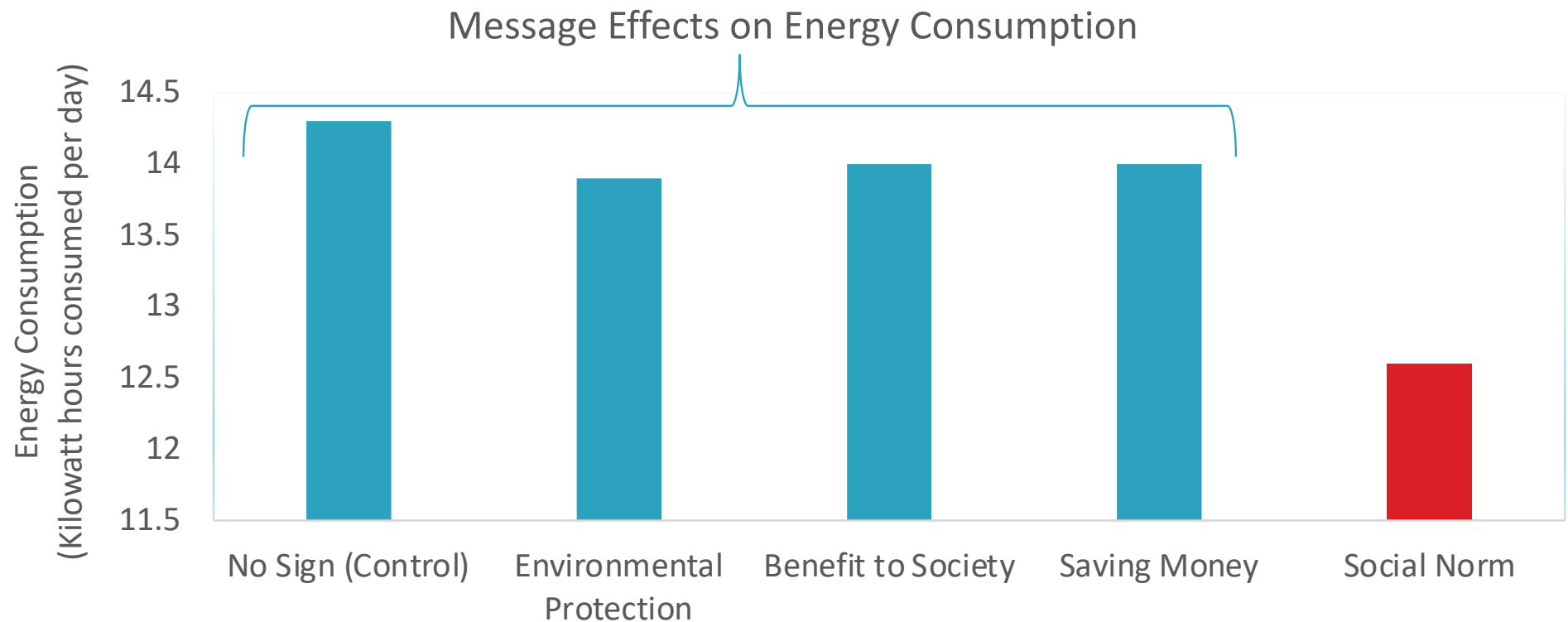
49

- Researchers then conducted an experiment in a California neighborhood
- They placed signs on residents' door encouraging them to conserve energy
- Researchers varied the reasons for conserving on the signs to be in accordance with the various motivating factors

Results of Field Experiment (Goldstein et al., 2010)



Results of Field Experiment (Goldstein et al., 2010)







Would it matter what color that the competitors wear as they compete?

Winning Color?

54

Researchers looked into tone of different games on one-on-one sports: Boxing, Taekwondo, and Wrestling.

They wanted to see if there would be a systematic advantage depending on what color they are wearing.

Winning Color?

55

Researchers looked into tone of different games on one-on-one sports: Boxing, Taekwondo, and Wrestling.

They wanted to see if there would be a systematic advantage depending on what color they are wearing.

This is an interesting and important question, because the IOC has a rule to toss a coin to randomly assign different color uniforms to competitors, to eliminate any systematic advantage.

Winning Color?

56

Researchers looked into tone of different games on one-on-one sports: Boxing, Taekwondo, and Wrestling.

They wanted to see if there would be a systematic advantage depending on what color they are wearing.

This is an interesting and important question, because the IOC has a rule to toss a coin to randomly assign different color uniforms to competitors, to eliminate any systematic advantage.

That seems fair, but it turns out that the color of uniform one wears, creates systematic advantages and disadvantages for competitors.



Winning Color?

58

Yes, this is small sample.

Winning Color?

59

Yes, this is small sample.

So researchers tried to analyze entire matches on different games on different Olympic games and found a systematic advantage in terms of winning a game about 10% advantage: 55% and 45%.

Winning Color?

60

Yes, this is small sample.

So researchers tried to analyze entire matches on different games on different Olympic games and found a systematic advantage in terms of winning a game about 10% advantage: 55% and 45%.

Players have different odds of winning games to begin with, so researchers match such competitiveness of players and the results are even more perplexing: now in those even matches: 62% vs. 38%.

Winning Color?

61

Yes, this is small sample.

So researchers tried to analyze entire matches on different games on different Olympic games and found a systematic advantage in terms of winning a game about 10% advantage: 55% and 45%.

Players have different odds of winning games to begin with, so researchers match such competitiveness of players and the results are even more perplexing: now in those even matches: 62% vs. 38%.

You have a two thirds greater chance of winning changes when you wear red uniforms.

What's the Mechanism?

62

Wearing red you feel more dominant,
and when you see someone wearing
red they look more dominant.

To figure this out, researchers showed
referees a close match and asked them
to judge who won.

They also asked how many points the
referees would give.

What's the Mechanism?

63

Wearing red you feel more dominant, and when you see someone wearing red they look more dominant.

To figure this out, researchers showed referees a close match and asked them to judge who won.

They also asked how many points the referees would give.

On average, referees give competitors wearing red 8 more points.



What's the Mechanism?

64

Yes, again, this could be the case because the competitor performed better.

What's the Mechanism?

65

Yes, again, this could be the case because the competitor performed better.

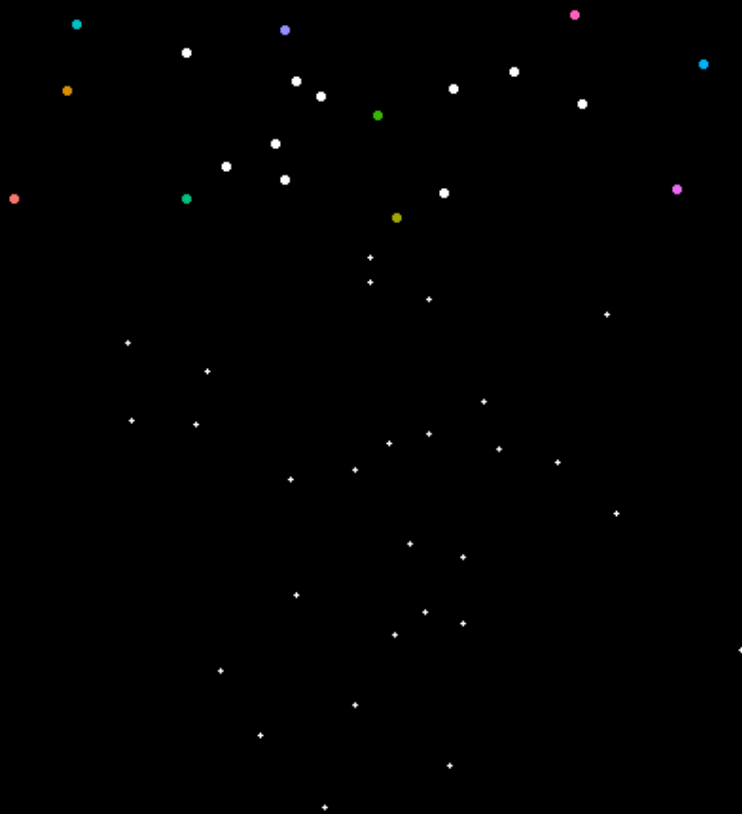
To rule out such an alternative account, the researchers digitally changed the color of uniform and showed the same game to different groups of referees.

What's the Mechanism?

66



Again, the competitor wearing red received 8 more points on average from referees, although these same competitors received 8 points less in the previous study when they were shown to referees wearing blue.



Analytically Yours,
Prof. Roh