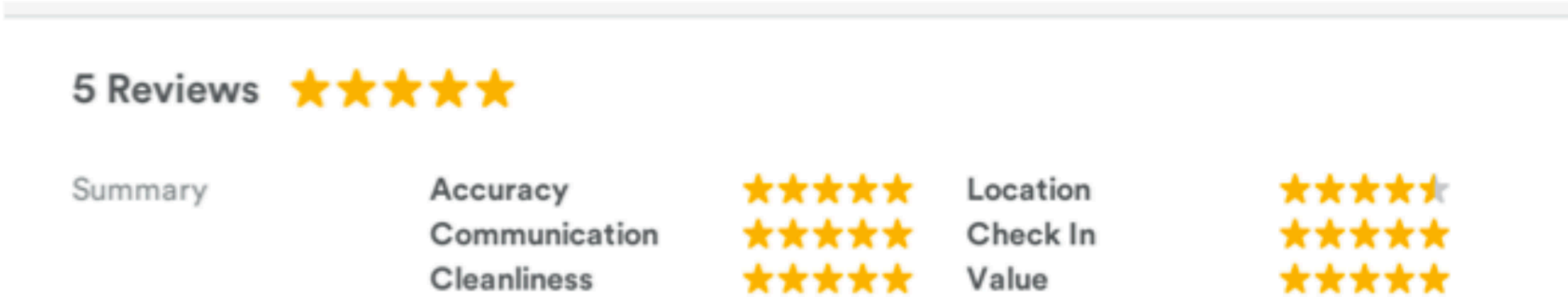


# Business Experimentation and Causal Methods

Prof. Fradkin  
Topic: Experimenting in Practice




# This Time


1. How to design an experiment.
  - Problem to be solved?
  - Types of treatment arms.
  - Number of treatment arms.
  - Unit of randomization.
  
2. How to analyze an experiment after.
  - Collect appropriate data.
  - Check for proper randomization.
  - Obtaining estimates.
  - Thinking about effect sizes.


# Example based on my research.

## MARKETING SCIENCE

[JOURNAL HOME](#) [ARTICLES IN ADVANCE](#) [CURRENT ISSUE](#) [ARCHIVES](#) [ABOUT](#) [SUBMIT](#) [SUBSCRIBE](#)

 Open Access | 

 View PDF


 Tools |  Share

### Go to Section

Abstract

[Home](#) > [Marketing Science](#) > [Vol. 40, No. 6](#) >

## Reciprocity and Unveiling in Two-Sided Reputation Systems: Evidence from an Experiment on Airbnb

Andrey Fradkin , Elena Grewal, David Holtz 

Published Online: 8 Oct 2021 | <https://doi.org/10.1287/mksc.2021.1311>

### Abstract

Reputation systems are used by nearly every digital marketplace, but designs vary and the effects of these designs are not well understood. We use a large-scale experiment on Airbnb to study the causal effects of one particular design choice—the timing with which feedback by one user about another is revealed on the platform. Feedback was hidden until both parties submitted a review in the treatment group and was revealed immediately after submission in the control group. The treatment stimulated more reviewing in total. This is due to users’ curiosity about what their counterparty wrote and/or the desire to have feedback visible to other users. We also show that the treatment reduced retaliation and reciprocation in feedback and led to lower ratings as a result. The effects of the policy on feedback did not translate into reduced adverse selection on the platform.

[< Previous](#) [Back to Top](#) [Next >](#)

# What problem are we trying to solve?

---

Problem (circa 2014):

Guests and hosts sometimes left retaliatory reviews.

This eroded trust in the system and the fear of retaliation may have prevented honest reviews.

Some academic literature on this topic.



# Proposed solution: Simultaneous Reveal

---

- Basic Idea: If the guest can't see the host's review (and vice versa) then they can't retaliate.
- Eventually reviews must be revealed. After all, they must be made public.
- Reveal reviews when both sides have submitted the review.
- But what if only one review is submitted?

Review period is 30 days. After 30 days, no more reviews can be submitted and reviews are revealed.





# Think about unintended consequences

---

- Airbnb wanted reviews to be shown early since that would help the users.
- Under 30 day review period, that may take a while.
- Solution: change review period to 14 days.



# Problem: We're changing two things at once.

---

- Now the proposal makes reviews simultaneous reveal and lowers the review period to 14 days.
- If we run that as an experiment, we won't know what had the effect, simultaneous reveal or 14 day review period.



# Solution: Two treatment arms

---

- Control Group: 30 day review period
- Treatment Arm 1: 14 day review period
- Treatment Arm 2: 14 day review period + simultaneous reveal.
- Compare Arm 1 and 2 to get the effect of JUST simultaneous reveal.





# Think about what determines launch decision.

---

- Under what condition will the treatment be given to everyone?
- Very problem specific.
  - Advertising: positive ROI
  - This experiment:
    - Review rates don't fall too much.
    - No consumer backlash.

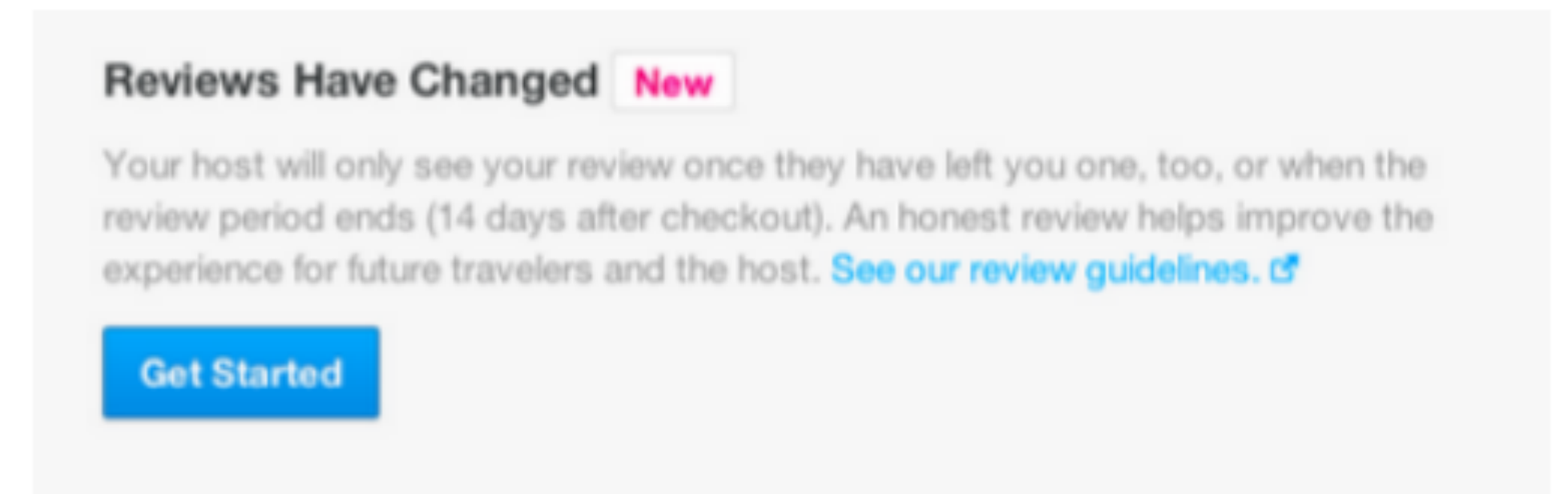


# Implementation details are part of the treatment.

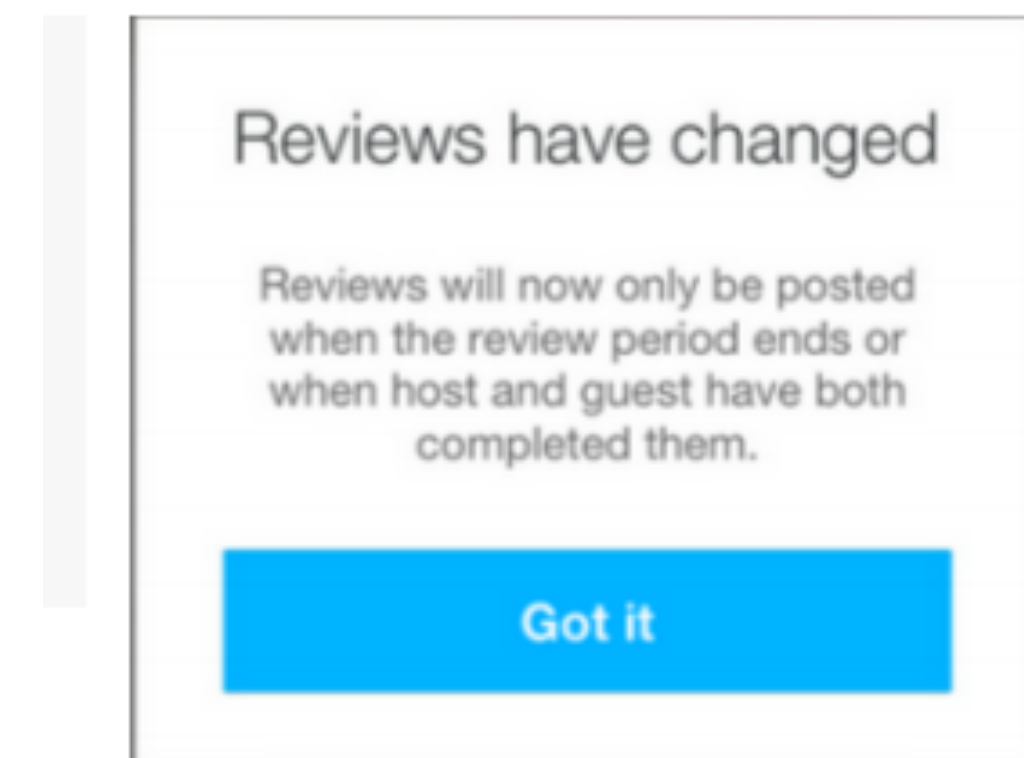
---

- How should Airbnb inform people of the new policy?
- Need design for:
  - Mobile App
  - Desktop
  - Mobile browser?
  - Different Languages

## (a) Desktop



## (b) Mobile



# Implementation Mistake:

---

## (a) Treatment

Hi [REDACTED],

Please leave a review for your recent guest, [REDACTED]. It's quick, easy, and your guest can only see it if they leave you a review too.

Reviews become public when both host and guest have submitted them or 14 days after the checkout date of the stay.

[Leave a review](#)

By leaving prompt, thorough, honest reviews, you help uphold one of the nine Airbnb Hospitality Standards for successful hosting.

Thank you for your part in building our worldwide community!  
The Airbnb Team

## (b) Control

Dear [REDACTED],

We hope you enjoyed hosting [REDACTED]. Please contribute back to the Airbnb community by [leaving a review and rating](#) for [REDACTED]. [REDACTED] help future hosts make an educated decision.

You have **14 days** to submit a public review for [REDACTED].

If you do not want to leave a review, you can also tell us about your experience by [leaving private feedback only for Airbnb](#).

Thanks,  
The Airbnb Team



# Implementation Mistake:

## Different ways to link to review form! Unintended difference.

(a) Treatment

Hi [REDACTED],

Please leave a review for your recent guest, [REDACTED]. It's quick, easy, and your guest can only see it if they leave you a review too.

Reviews become public when both host and guest have submitted them or 14 days after the checkout date of the stay.

[Leave a review](#)

By leaving prompt, thorough, honest reviews, you help uphold one of the nine Airbnb Hospitality Standards for successful hosting.

Thank you for your part in building our worldwide community!  
The Airbnb Team

(b) Control

Dear [REDACTED],

We hope you enjoyed hosting [REDACTED]. Please contribute back to the Airbnb community by [leaving a review and rating](#) for [REDACTED]. [REDACTED] help future hosts make an educated decision.

You have **14 days** to submit a public review for [REDACTED].

If you do not want to leave a review, you can also tell us about your experience by [leaving private feedback only for Airbnb](#).

Thanks,  
The Airbnb Team



# Do we have enough power?

---

- Step 1: What outcome are we interested in measuring?

For simultaneous reveal: ratings behavior

- Step 2: What effect sizes do we want to detect?

From Airbnb's perspective, sizable drops in reviewing and ratings. (Downside risk)

- Step 3: Do a power calculation to figure out what sample size we need.

This will tell us what share of users to randomize and potentially how long to run the experiment. The longer we run it, the more users.



# What is the unit of randomization?

---

- Guest?
  - Most power since more guests than hosts.
  - Same listing may get some guests with treatment and others in control.  
Inconsistent experience.
- Listing?
- Host?
- City of stay?



# What is the unit of randomization?

---

- Guest?
- Listing?
  - Guests tend to use Airbnb less frequently so we're not as worried about inconsistent experience for guests.
  - But, same host has multiple listings.
  - Potential for spillovers across listings for same host.
- Host?
- City of stay?



# What is the unit of randomization?

---

- Guest?
- Listing?
- **Host?**
  - **How the experiment was randomized.**
  - Consistent experience for every host.
  - Captures within-host spillovers if review for one listing changes demand for other listings of same host.
- City of stay?





# What is the unit of randomization?

---

- Guest?
- Listing?
- Host?
- City of stay?
  - Advantage: captures spillovers across listings in same city.
  - Disadvantage: fewer cities than users, so little statistical power.



# Types of Randomization: Blocking

---

- Blocking:  
Randomizing in a way that balances the treatment group on certain characteristics. This is a way to increase statistical power.

Example: If we are studying the effect of an education related treatment on grades, we may want to make sure that the treatment and control groups have the same proportion of A and B students based on last year's GPA. Blocking would accomplish this.

- The reason we choose past grades, is that they predict future grades.



# Types of Randomization: Clustering

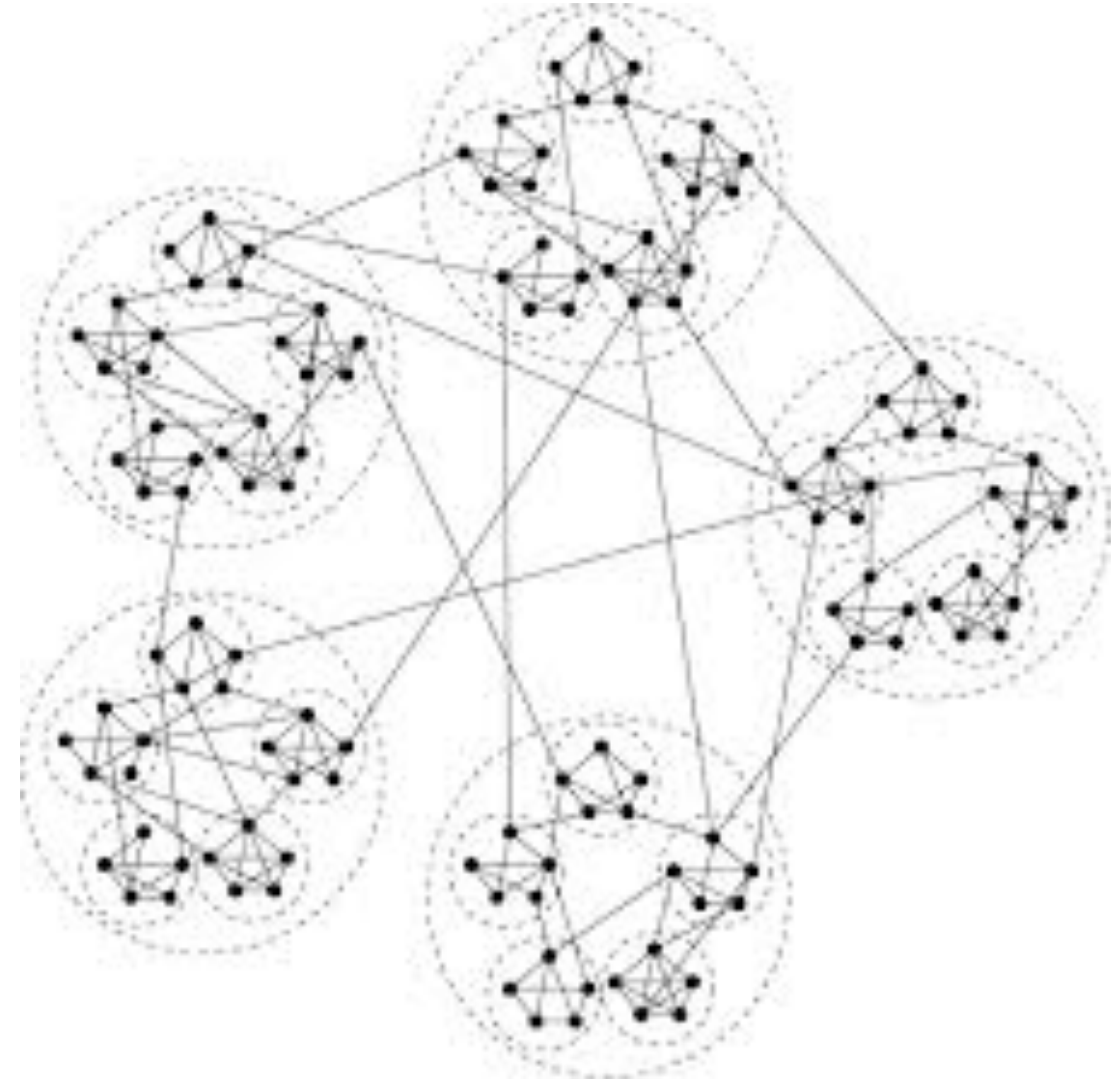
---

- Clustering:  
Assign the treatment to groups or clusters rather than individuals.

Example 1: Assign an educational treatment by classroom rather than student. Clustering at a classroom level.

Example 2: Assign simultaneous reveal treatment by host rather than listing. Clustering at a host level.

- This is way to ‘capture spillovers’ in our experiment. It measures the effect of treating everyone in the cluster.
- Measure the outcome at the level of the cluster rather than the level of individual.



**Warning about clustering!**

**Number of observations is the number of clusters, not  
the number of individuals in those clusters!**



# Make sure to collect the right data.

---

- Think of what outcomes you want to measure. Make a way to collect them and link them to treatment assignment.
- Examples:
  - Create a log in the database when a user clicks on something or opens an email.
  - Embedding a survey in the experiment.
  - Collecting information (cookies, etc...) that allows you to link the experimental data with other data.



# Coming up next:

---

February 6, 2024	Case	Assignment 2 Due	Case, intro to final project, assignment solutions	Finish Assignment 2
February 8, 2024			Regression + Analyzing Experiment Lecture, in-class regression exercise. Create regression tables.	
February 13, 2024			Heterogeneity, p-hacking, ethics, Guest Lecture (Madhav Kumar)	

# This Time

1. How to design an experiment.
  - Problem to be solved?
  - Types of treatment arms.
  - Number of treatment arms.
  - Unit of randomization.
  
2. **How to analyze an experiment after.**
  - **Collect appropriate data.**
  - **Check for proper randomization.**
  - **Obtaining estimates.**
  - **Thinking about effect sizes.**