

2024 | CS-E4730 | Computational social science
10 | Experiments & interventions at scale | Arttu
Malkamäki

Project deadline 27 May 2024.

Project submission not mandatory.

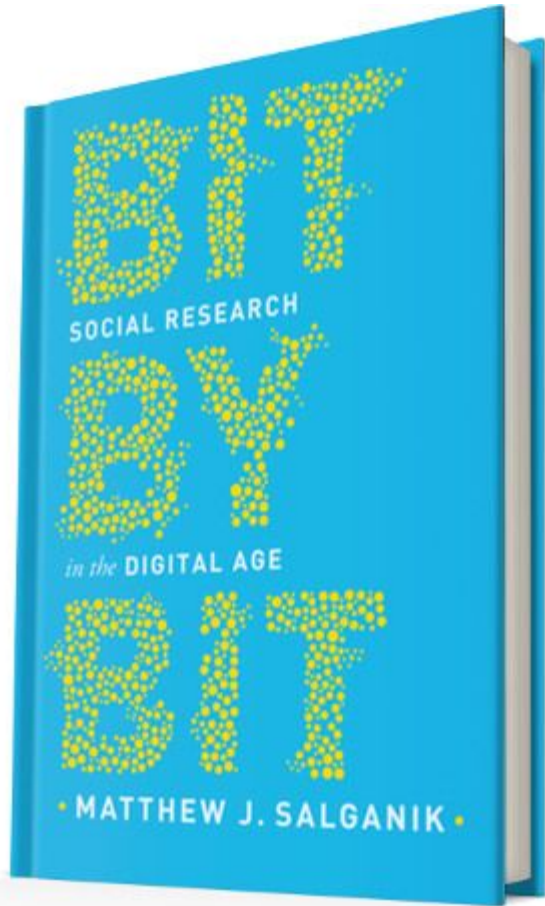
Note clarification of B4 on A+.

Week	Lecture	Exercise dl	Extended dl	Topic
01	Feb-26	Mar-01	Mar-13	Introduction to CSS
02	Mar-04	Mar-08	Mar-20	Artificial societies & agent based models
03	Mar-11	Mar-15	Mar-27	Data & digital traces
04	Mar-18	Mar-22	Apr-03	Counting things & analysing text
05	Mar-25	Apr-05	Apr-17	Social networks: structure
06	Apr-08	*	—	Introduction to CSS-project
—	—	—	—	Exams
07	Apr-22	Apr-26	May-08	Ethics, privacy, legal
—	—	—	—	Wappu
08	May-06	May-10*	May 22	Agent-based models & emergence
09	May-13	May-19**	May 29	Social networks: dynamics
10	May 20	May 24**	Jun 05	Experiments & interventions at scale

* Project deadline May 27

** Bonus round

*** Only lecture questions



The lecture of today draws from the chapter *Running Experiments* in **Bit-By-Bit** by Matthew J. Salganik.

1. Defining experiments
2. Improving experiments
3. Doing digital experiments

bitbybitbook.com/en/1st-ed/running-experiment

Defining experiments

Improving experiments

Doing digital experiments

Defining experiments



Types?
Experiments?
Designs?
Objectives?
Scales?

Defining experiments

Perturb and observe experiment [P&O]

- Setup system modification, feedback observation
- Focus system behaviour
- Methodology iterative, adaptive
- Common field natural sciences, engineering
- Bias control less formal, empirical
- Measurement observational, practical results

Randomised controlled experiment [RCE]

- Setup randomisation, controlled environment
- Focus treatment efficacy
- Methodology random assignment, control group comparison
- Common field **social sciences**, medicine
- Bias control formal, statistical rigour
- Measurement statistical significance, reliability

Defining experiments

Modus operandi of randomised controlled experiments in the social sciences

Goal >> Hypothesis >> Participants >> Random sample A >> Treatment >> Outcome >> Compare
>> Random sample B >> ~~Treatment~~ >> Outcome

0. Determine the goal of the experiment [optimisation, understanding, both]
1. Formulate a meaningful hypothesis
2. Recruit participants
3. Assign participants randomly into *treatment* [A] and *control* [B] groups
4. Deliver treatment
5. Compare outcomes

Defining experiments

Dimensions of randomised controlled experiments in the social sciences



Defining experiments

Correll, Benard, Paik [2007]. *Getting a Job: Is There a Motherhood Penalty?* American Journal of Sociology 112[5]: 1297-1338.

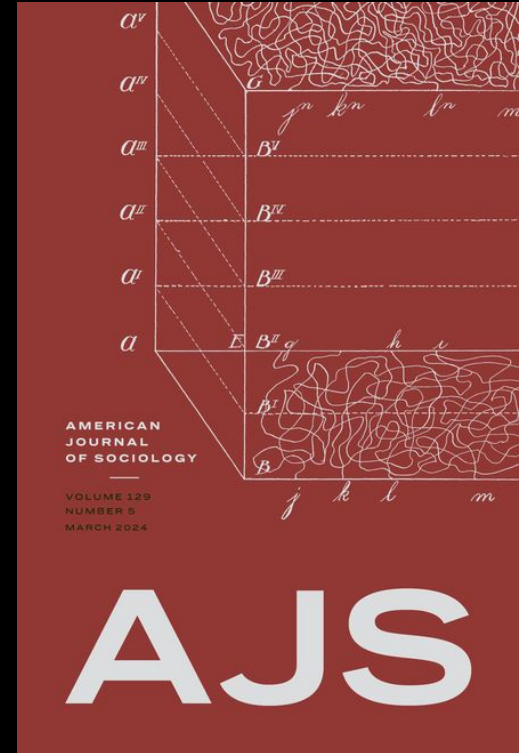
How come mothers earn less than childless women in same jobs and with similar skills?

Lab-like

- **Participants:** college undergraduates
- **Task:** review resumes of job applicants [treatment signalling motherhood], rate on competence, warmth, and commitment to work; recommend whether to hire or not and if yes, what starting salary.
- **Outcome:** mothers [treatment] less likely to get hired, and if they get hired, they start with lower salary.
- **Analysis:** outcome due to lower ratings in competence and commitment.

Field-like

- **Participants:** real firms
- **Task:** response to hundreds of job advertisements with fake applications [treatment signalling motherhood].
- **Outcome:** mothers [treatment] less likely to receive an invitation to a job interview.
- **Analysis:** falls short in explaining behaviour; lacks information on the decision-making process by the participants.



Defining experiments

Correll, Benard, Paik [2007]. *Getting a Job: Is There a Motherhood Penalty?* American Journal of Sociology 112[5]: 1297-1338.

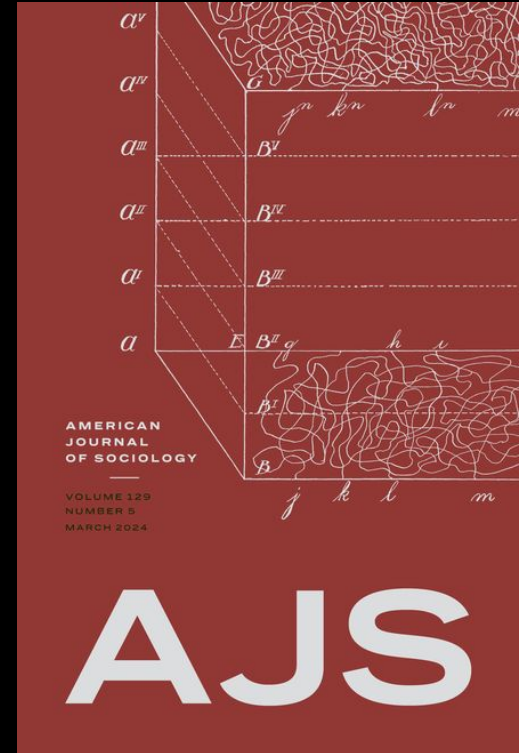
How come mothers earn less than childless women in same jobs and with similar skills?

Advantages of lab over field

- Near total control over the environment; resumes read in lab-like setting, whereas some resumes were perhaps never read in the field.
- Easier access to information on the mechanisms underpinning participant behaviour.

Disadvantages of lab over field

- Participants perhaps knew the goal of the study, leading to bias in behaviour.
- Small differences stand out in a lab, perhaps leading to overestimation of the effect.
- Participants do not represent the target population.



Defining experiments

Dimensions of randomised controlled experiments in the social sciences

DIGITAL

- Way more participants
- Low marginal cost of adding participants **once infrastructure is set up**
- If measurement *always-on*, pre-treatment data could facilitate participant selection
- If measurement *always-on*, experiment could be run for much longer
- Ethical considerations are real

ANALOG

- Correll et al. 2007.
-

Defining experiments

Restivo and van de Rijdt [2012]. *Experimental Study of Informal Rewards in Peer Production*. PLOS ONE 7[3]: e34358.

How come individuals contribute to public goods?

DIGITAL FIELD EXPERIMENT

- **Participants:** active Wikipedia editors who had not yet been awarded a ★.
- **Task:** give some editors ★s [treatment] and examine effect on editing behaviour.
- **Outcome:** treatment group edited less, but control group edited even lesser.
- **Analysis:** editing behaviour on Wikipedia is bursty; ★s allowed for sustaining productivity when compared to what would have happened without them; without control group, the analysis would have gotten it all wrong.



Improving experiments

Improving experiments

VALIDITY

- **Do the results of an experiment support a more general conclusion?**
 - statistical conclusion validity: was the statistical analysis appropriate
 - internal validity: was the experiment run appropriately [A/B stayed separate, treatment delivery succeeded]
 - construct validity: do data correspond to the theoretical construct [e.g., bias, sexism, cohesion] of interest
 - external validity: whether the results apply to other situations, sets of participants, or ways of measurement
- **Digital experiments perhaps improve internal and external validity, but struggle with construct validity**
 - treatment delivery and outcome measurement easier
 - experiment replication cheaper
 - digital traces do not directly translate into meaningful theoretical constructs

Improving experiments

HETEROGENEITY OF TREATMENT EFFECTS

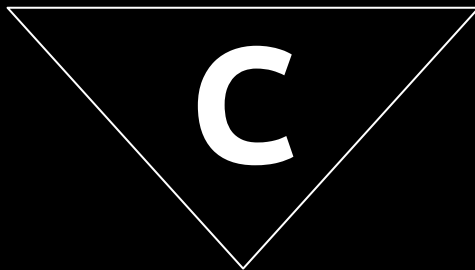
- **Treatments affect different participants differently**
 - Schulz et al. 2007, electricity consumption for one group went down, up for another; zero average effect
 - Larger sets of participants increase heterogeneity
- **Splitting participants either *a priori* or *in situ* could remedy heterogeneity of treatment effects**
 - Pre-treatment information [advantage of digital experiments] provide cues to participant attributes
 - Treatment group could also be split up during experiment; enough splits isolate the source of heterogeneity
 - Machine learning-based predictive approaches could automate detection of heterogeneous treatment effects
 - Helps to understand how and for whom the treatment works, and how to improve the treatment

Improving experiments

FOCUS ON MECHANISMS

- Digital experiments somewhat easier to specifically design to test for mechanisms.
- Understanding mechanisms is core to all contemporary research, but often difficult.

More citrus



Less scurvy

Doing digital experiments

Doing digital experiments

COST

- How much money, time, and effort to run the experiment

CONTROL

- How much control over the social environment

REALISM

- Does the experiment involve naturally occurring treatments, environments, or participants

ETHICS

- Does the experiment involve ethical considerations

Doing digital experiments

	COST	CONTROL	REALISM	ETHICS
Partner with governments, firms, or non-profit organisations	low	medium	high	harder
<hr/>				
Use existing digital systems [Doleac and Stein 2013; Restivo and van de Rijt 2012]	low	low	high	harder
Build a digital experiment [Centola 2010; Brown et al. 2022]	medium	high	medium	easier
Build a digital product [a game, an application]	high	high	high	easier

Doing digital experiments

Centola [2010]. *The Spread of Behavior in an Online Social Network Experiment*. *Science* 329[5996]: 1194-1197.

How does social network structure influence spread of behaviour?

DIGITAL FIELD EXPERIMENT

- **Participants:** online health community [recruited via online ads].
- **Task:** observe the spread of same behaviour in populations embedded in different network structures [otherwise indistinguishable]; assign health buddies to participants upon arrival to induce either random [control] or community structures [treatment]; inject a behaviour into each network [a chance to register to a health information portal]; notify participants if their buddies register.
- **Outcome:** behaviour spread faster and farther in a community structure [diffusion measured by the fraction of the total network adopting the behaviour; speed of diffusion evaluated by comparing the time required for the behaviour to spread to the greatest fraction reached].



Doing digital experiments

Brown et al. [2022]. *Echo Chambers, Rabbit Holes, and Algorithmic Bias: How YouTube Recommends Content to Real Users.*

The Solution: Remove User Choice!

- Recruit YouTube users to install **browser extension**
- Start on randomly assigned **seed video**
- Follow **traversal rule** to select recommended video



YouTube Recommendation Downloader

Offered by: camaplugin



www.youtube.com/watch?v=o_nUNx8UDWI

GREAT POWER, GREAT RESPONSIBILITY

PNAS PNAS PNAS



Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer^{a,1}, Jamie E. Guillory^{b,2}, and Jeffrey T. Hancock^{b,c}

^aCore Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and Departments of ^bCommunication and ^cInformation Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness. Emotional contagion is well established in laboratory experiments, with people transferring positive and negative emotions to others. Data from a large real-world social network, collected over a 20-y period suggests that longer-lasting moods (e.g., depression, happiness) can be transferred through networks [Fowler JH, Christakis NA (2008) *BMJ* 337:a2338], although the results are controversial. In an experiment with people who use Facebook, we test whether emotional contagion occurs outside of in-person interaction between individuals by reducing the amount of emotional content in the News Feed. When positive

demonstrated that (i) emotional contagion occurs via text-based computer-mediated communication (7); (ii) contagion of psychological and physiological qualities has been suggested based on correlational data for social networks generally (7, 8); and

(iii) people emotional experiences is no experim in the absen On Face later seen b (8). Becau

Editorial Expression of Concern and Correction

PSYCHOLOGICAL AND COGNITIVE SCIENCES

PNAS is publishing an Editorial Expression of Concern regarding the following article: “Experimental evidence of massive-scale emotional contagion through social networks,” by Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock, which appeared in issue 24, June 17, 2014, of *Proc Natl Acad Sci USA* (111:8788–8790; first published June 2, 2014; 10.1073/pnas.1320040111). This paper represents an important and emerging area of social science research that needs to be approached with sensitivity and with vigilance regarding personal privacy issues.

Questions have been raised about the principles of informed consent and opportunity to opt out in connection with the research in this paper. The authors noted in their paper, “[The

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Correction for “Experimental evidence of massive-scale emotional contagion through social networks,” by Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock, which appeared in issue 24, June 17, 2014, of *Proc Natl Acad Sci USA* (111:8788–8790; first published June 2, 2014; 10.1073/pnas.1320040111).

The authors note that, “At the time of the study, the middle author, Jamie E. Guillory, was a graduate student at Cornell University under the tutelage of senior author Jeffrey T. Hancock, also of Cornell University (Guillory is now a postdoctoral fellow at Center for Tobacco Control Research and Education, University of California, San Francisco, CA 94143).” The author and affiliation lines have been updated to reflect the above changes

Doing digital experiments

GREAT POWER, GREAT RESPONSIBILITY

- **Replace**
 - If possible, replace experiments with less intrusive approaches
 - Natural experiments leverage a naturally occurring situation **instead** of imposing conditions that make people behave in a certain way [but offer limited control over the environment].
 - Coviello et al. 2014: people tend to write more negatively when it is raining; then, does the emotion spread].
- **Refine**
 - Minimise the harm that the treatment has on participants; not easy.
 - Emotional contagion study blocked people from seeing potentially important posts; an alternative design could have simply boosted positive and negative words.
- **Reduce**
 - Minimise the number of participants as much as possible; always uncertainty about the consequences
 - Emotional contagion study used difference-in-means estimator [measure difference between groups after treatment]; an alternative design could have used difference-in-differences estimator [measure difference between groups before and after treatment] that produces more valid results and often requires less participants.

Finale

Finale

Defining experiments: what are experiments, how digital differs from [lab or field] analog experiments

Improving experiments: validity, heterogeneity of treatment effects, mechanism

Doing digital experiments: four ways to run experiments

Attention to ethics: replace, refine, and reduce