Design and analysis of experiments

HU | Spring 2026 (Winter term 25-26)

Instructor	**Teaching assistant**
Macartan Humphreys	Lennard Naumann
macartan.humphreys@wzb.eu	lennard.naumann@wzb.eu
macartan.github.io	E306

Overview: The course addresses the design and analysis of experiments. It is a hands-on course in which theoretical results are introduced through lecture, demonstrated in practice, and worked on in groups. In-class discussions focus on major works, features of designs, and experimental strtegies. Topics include randomization schemes, analysis strategies, design evaluation, and workflows.

1 Class resources

- All resources: https://macartan.github.io/ci
- This syllabus: https://macartan.github.io/ci/syllabus_experiments_2026.pdf
- Slides: https://macartan.github.io/ci/slides_experiments_2026.html
- Student survey
- Git repo https://github.com/macartan/ci: This repo is for both my causal inference class and this experiments class.

2 Times and locations

Meetings will be held at the WZB

Date	Time	Location	Topic
9 Jan 2026	9–13	WZB A 305	Getting started
16 Jan 2026	9–13	WZB A 305	Doing experiments
30 Jan 2026	9–13	WZB A 305	Causality and causal inquiries
6 Feb 2026	9–13	WZB A 305	Answer strategies
20 Feb 2026	9–13	WZB A 305	Data strategies and design evaluation
27 Feb 2026	9–13	WZB B 001	Topics: downstream experimentation; spillover experiments; patient preference trials; survey experiments

The sessions will be structured roughly as follows.

- 9:00 - 11:00: lecture or discussions of readings
- 11:00 - 11:15: break
- 11:15 - 13:00: discussions, team presentations and in-class exercises

Instructor office hours: Fridays, 13:00 - 15:00, in person

Teaching assistant office hours: Wednesdays, 13:00 - 14:00 on Zoom (excluding 25 Feb)

3 Expectations

3.1 Required (6 ECTS)

- Do the readings and take part in discussion (30%)
- Take part in field experiment I (30%)
- Design an experiment as a candidate for survey experiment II (40%)

3.2 Optional (+4 ECTS)

- Prepare a research design or short paper, perhaps building on existing work. Typically this contains:
  - a problem statement
  - a description of an experimental strategy to address the problem
  - empirical (ideally) or simulation based results (ok)
  - a discussion of implications
  - A passing paper will illustrate identify a causal effect of substantive interest credibly; a good paper will innovate on method or topic.
- Deadline: August 31, 2026

4 Sessions and readings

The course combines lectures, discussion, and activities.

All readings are linked from this document.

I will draw material especially from Blair, Coppock, and Humphreys (2023): Research Design in the Social Sciences

Recommended but not open access: Gerber and Green (2012) with supplementary material available here: https://isps.yale.edu/FEDAI

Readings for each session are given below.

4.1 Intro (9 January)

Aims:

- Introduce the class
- Practice running experiments by hand
- Introduce and discuss field experiment I
- To prepare for next session: proposed adaptation of field experiment 1

Course outline, tools

4.2 Experimentation (16 January)

Aims:

- Review lifecycle of an experiment (discussion)
- Discuss ethical challenges to experimentation
- Discuss a set of strong experiments
- Develop group field experiment 1, decide on strategy, and start ethics application

To do before class: 1. the readings; 2. develop specific proposals in small groups for field experiment 1 (4 slides), completing individually or collectively a copy of the EGAP research design form

Required reading:

on ethics:

- Johnson: Protecting the Community: Lessons from the Montana Flyer Project
- APSA guidelines

some cases:

- Example 1: Gerber, Green, and Larimer (2008)
- Example 2: Lowe (2021)

- The Hijab penalty and the Appendix

Optional pieces on experimentation:

- Humphreys (2015) on ethics in experiments
- Design lifecycle Blair, Coppock, and Humphreys (2023) p 319 - 358
- DD, Ch 2 introduces ideas at a high level
- Alvarez, Key, and Núñez (2018)
- Humphreys, De la Sierra, and Van der Windt (2013) on fishing and registration

Optional substantive pieces:

- Kurzban, Tooby, and Cosmides (2001)
- Haas et al. (2025)
- Mousa (2020)
- Dietrich and Sands (2023)

4.3 Causality (30 January)

Aims:

- Understand the potential outcomes framework
- Understand identification challenges and the role of randomization
- Understand varieties of estimands

To do before class: 1. the readings; 2. collectively complete a pre-analysis plan and ethics application for Experiment 1

Required:

- Holland is a beautiful classic reading on the fundamental problem of causal inference
- DD, Ch 7

Optional:

- II Ch 2, goes over both potential outcomes and DAGs
- Keele (2015) gives a good high level overview of many key ideas in this course
- Imai, King, and Stuart (2008)
- Hernán and Robins (2024) (Section 3.1, 7.2, 8.4, 10.1)
- Hernán and Robins (2024) Ch 6
- II Ch 4
- Lundberg, Johnson, and Stewart (2021)

4.4 Answer strategies (6 February)

Aims:

- Understand alternative strategies to estimating treatment effects
- Understand design based strategies for estimating standard errors and p-values
- Understand confidence intervals
- Understand the role of covariates in experimental design

To do before class: 1. Collective presentation on the implementation and findings from Experiment 1

Optional:

- Freedman (2008) helps make connections between design based inference and regression
- Bind and Rubin (2020) on RI
- Lin (2012) relieves some worries you might have after reading Freedman (2008)
- Chang, Middleton, and Aronow (2024) follows up on Lin (2012).

4.5 Experimental Design and Evaluation (20 February)

Aims:

- Understand alternative assignment strategies
- Understand statistical power
- Understand how to assess alternative diagnosands

To do before class: 1. The readings 2. Share initial design for for survey based experiment.

Required:

- DD, Ch 7
- DD, Ch 8 on data strategies
- DD, Ch 9 on answer strategies
- DD, Ch 13 gives more practical guidance on getting started

4.6 Topics (27 February)

To do before class: 1. The readings 2. Round-robin report on design performance. Teams receive a class-mate's design, declare a model that produces possible data from the design (without consulting the authors design declaration, if any) and applies the original authors answer strategy, reporting performance.

Required reading:

4.6.1 survey experiments

- Hainmueller, Hopkins, and Yamamoto (2014)

4.6.2 mediation experiments

- Imai, Keele, and Tingley (2010) on mediation

4.6.3 Lab experiments

- Roth (1986; Palfrey 2009)

4.6.4 Experiments on networks

- Baird et al. (2014) [Optional: Sinclair, McConnell, and Green (2012); Aronow and Samii (2017)]

4.6.5 downstream experimentation

- Coppock and Green (2016)

Not required:

4.6.6 patient preference trials

- Benedictis-Kessner et al. (2019)

5 Prerequisites and tools

You should already have background in statistics up to the point of feeling comfortable with regression.

In addition you should know some `R`. Really, the more you can invest on getting on top of `R` before the class the better.

Before the first class please make sure your `R` is up-to-date and that you are working in `R studio`. Then make sure you have the following packages installed.

```
pacman::p_load(
  rstan,
  dagitty,
  DeclareDesign,
  CausalQueries,
  tidyverse,
  knitr
  )
```

Resources:

- Resources for learning `R`: http://www.r-bloggers.com/how-to-learn-r-2/
- Hadley Wickham's R for Data Science
- Grant McDermott on data science

5.1 Writing with `qmd`

Please plan to do drafting in `Quarto`. This is a simple markup language that lets you integrate writing and coding. This document is written in `Quarto` and the slides will be also.

The key thing is that you can insert code chunks like this.

```
# Define a random number
x <- rnorm(1)
```

Code like this is run as the document compiles, and results can be accessed as needed, like this: we just sampled the random number x = -1.5230672.

- See Wickham, Çetinkaya-Rundel, and Grolemund (2023) on Workflow and Quarto

I recommend using `Rstudio` as an editor. More information here: https://quarto.org/docs/get-started/hello/rstudio.html

5.2 AI

AI can be a great aid to help understand this material and also to develop code. However it should only be used as an aid: it makes many many mistakes; you remain responsible for your code and more importantly for your comprehension.

References

Alvarez, R Michael, Ellen M Key, and Lucas Núñez. 2018. "Research Replication: Practical Considerations." *PS: Political Science & Politics* 51 (2): 422–26.

Aronow, P M, and Cyrus Samii. 2017. "Estimating Average Causal Effects Under General Interference, with Application to a Social Network Experiment." *Annals of Applied Statistics* 11 (4): 1912–47.

Baird, Sarah, J Aislinn Bohren, Craig McIntosh, and Berk Ozler. 2014. "Designing Experiments to Measure Spillover Effects." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2505070.

Benedictis-Kessner, Justin de, Matthew A Baum, Adam J Berinsky, and Teppei Yamamoto. 2019. "Persuading the Enemy: Estimating the Persuasive Effects of Partisan Media with the Preference-Incorporating Choice and Assignment Design." *American Political Science Review* 113 (4): 902–16.

Bind, M-AC, and DB Rubin. 2020. "When Possible, Report a Fisher-Exact p Value and Display Its Underlying Null Randomization Distribution." *Proceedings of the National Academy of Sciences* 117 (32): 19151–58.

Blair, Graeme, Alexander Coppock, and Macartan Humphreys. 2023. *Research Design in the Social Sciences: Declaration, Diagnosis, and Redesign*. Princeton University Press.

Chang, Haoge, Joel A Middleton, and PM Aronow. 2024. "Exact Bias Correction for Linear Adjustment of Randomized Controlled Trials." *Econometrica* 92 (5): 1503–19.

Coppock, Alexander, and Donald P Green. 2016. "Is Voting Habit Forming? New Evidence from Experiments and Regression Discontinuities." *American Journal of Political Science* 60 (4): 1044–62.

Dietrich, Bryce J., and Melissa L. Sands. 2023. "Seeing Racial Avoidance on New York City Streets." *Nature Human Behaviour* 7 (8): 1275–81.

Freedman, David A. 2008. "On Regression Adjustments to Experimental Data." *Advances in Applied Mathematics* 40 (2): 180–93.

Gerber, Alan S, and Donald P Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. Norton.

Gerber, Alan S, Donald P Green, and Christopher W Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102 (1): 33–48.

Haas, Violeta I, Tim Wappenhans, Ferdinand Geißler, Felix Hartmann, Daniel Bischof, Johannes Giesecke, Macartan Humphreys, Heike Klüver, and Lukas F Stoetzer. 2025. "Beyond Persuasion: Protest's Direct Behavioral Impact on Bystanders." *Working Paper*.

Hainmueller, Jens, Daniel J Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22 (1): 1–30.

Hernán, Miguel A, and James M Robins. 2024. *What If?* Boca Raton: Chapman & Hall/CRC.

Humphreys, Macartan. 2015. "Reflections on the Ethics of Social Experimentation." *Journal of Globalization and Development* 6 (1): 87–112.

Humphreys, Macartan, Raul Sanchez De la Sierra, and Peter Van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21 (1): 1–20.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15 (4): 309–34.

Imai, Kosuke, Gary King, and Elizabeth A Stuart. 2008. "Misunderstandings Between Experimentalists and Observationalists about Causal Inference." *Journal of the Royal Statistical Society Series A: Statistics in Society* 171 (2): 481–502.

Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology." *Political Analysis* 23 (3): 313–35.

Kurzban, Robert, John Tooby, and Leda Cosmides. 2001. "Can Race Be Erased? Coalitional Computation and Social Categorization." *Proceedings of the National Academy of Sciences* 98 (26): 15387–92.

Lin, Winston. 2012. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *arXiv Preprint arXiv:1208.2301*.

Lowe, Matt. 2021. "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *American Economic Review* 111 (6): 1807–44.

Lundberg, Ian, Rebecca Johnson, and Brandon M Stewart. 2021. "What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory." *American Sociological Review* 86 (3): 532–65.

Mousa, Salma. 2020. "Building Social Cohesion Between Christians and Muslims Through Soccer in Post-ISIS Iraq." *Science* 369 (6505): 866–70.

Palfrey, Thomas R. 2009. "Laboratory Experiments in Political Economy." *Annual Review of Political Science* 12: 379–88.

Roth, Alvin E. 1986. "Laboratory Experimentation in Economics." *Economics & Philosophy* 2 (2): 245–73.

Sinclair, Betsy, Margaret McConnell, and Donald P Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments." *American Journal of Political Science* 56 (4): 1055–69.

Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2023. *R for Data Science*. O'Reilly Media, Inc.