# Advances in the Design and Analysis of Experiments

University of Illinois at Urbana-Champaign

Sample syllabus

## Instructor

Gustavo Diaz
diazdia2@illinois.edu
420 David Kinley Hall MC-713
1407 W Gregory Drive
Urbana, IL 61801

## Course Description

This course covers recent advances in the design and analysis of experiments in Political Science (and beyond). We focus on two areas. First, we discuss the merit of doing experiments: What can we learn from experiments? Does our research question require an experiment? If so, what kind of experiment? How can we build cumulative knowledge from experiments?

Second, we discuss the nuts and bolts of experiments: How can we know how an experiment should look like before knowing the results? How should we measure our outcome variables? How to assign units to treatment and control groups? How to test hypothesis or calculate treatment effects given our research design? How to distinguish meaningful effects from random chance? What of all this should we include in a pre-analysis plan?

As usual, the answer to these questions is a big "it depends." A good experiment looks different depending on the discipline, field, and sub-field. But in this course we will collectively develop our own standards to navigate these questions in a productive way.

**Note:** The current version of this syllabus aims for a 16-week graduate seminar that follows after (at least) introductory courses to research design and quantitative data analysis. I could make it more general and fill the space of an introduction to research design, or turn it into more specific modules on field experiments, survey experiments, or an experiment design workshop. This may fit better in schools with a trimester system.

## Course Goals

As we transition from consumers to creators of knowledge, we learn about cutting-edge innovations in experiments in isolation. As a result, what we learn is path-dependent on the subject areas we follow. The main purpose of this course is to cast a wider net and create a space to collectively

reflect on our own work through the lenses of experiments, in the hopes of increasing the quality of our contributions.

By the end of this course, you will be able to engage with future developments on experiments in the social sciences, incorporate them to your own work, and to make unique and meaningful contributions both empirically and methodologically.

## Expectations

I assume you have taken introductory graduate courses on research design and data analysis in Political Science or an equivalent in statistics, social sciences, or medical sciences. You know the basics of experiments and can analyze data from experiments using statistical software. Maybe you have already collaborated on or conducted your own experiment. I also assume you want to learn more about experiments to make unique contributions in your career path of choice.

I expect familiarity, yet not expertise, with experiments. A few examples may help you calibrate this expectation.

- You know that experiments involve random assignment to conditions
- You use, or plan to use, experiments in your work and you want to have all your bases covered
- You would like to bring students to the laboratory, but you worry about students not being "real people"
- You like to read about experiments and are excited to participate in a conversation about them with like-minded people
- You have strong feelings about experiments that you want to articulate into productive critiques

I expect you to participate actively, productively, and respectfully in our meetings. Some of the material addresses complicated concepts or uses math extensively. I do not expect you to understand every single equation for this course, but I do expect you to read carefully enough to make a judgment about how the material relates to your work. That means you will get every detail if you choose to revisit the material after taking this course.

## Requirements

### Attendance

This course meets regularly per school policy. I will not take attendance, but repeated absences without justification (defined by school policy) will be considered a reason to fail the course. I will revisit this policy based on roster size and school regulations.

### Participation

X% of your grade will be based on participation. Participation involves actively engaging in discussion during our meetings according to the course expectations, as well as attending to office hours, and other forms of participation that suit your learning style. We will discuss ways to adjust the participation policy to suit everyone's goals and interests in our first meeting.

### Leading discussion

X% of your grade will be based on your role as discussion leader. The discussion leader's job is to briefly summarize the material and identify critical discussion topics in the form of questions, reflections, or critiques. In general, your goal as a discussion leader is to ensure a productive discussion. Based on enrollment I will determine a minimum of readings or weeks in which you will serve as discussion leader. If feasible, I will ask you to work in pairs in preparation to lead discussion. That may alleviate the stress of being put in the spot. You can choose to split the reading, or to collectively come up with talking points.

### Pre-analysis plan

Your final assignment will be to write an extended pre-analysis plan. Ideally, the topic will be an empirical contribution using experiments, or a methodological contribution to the literature on experiments. I am open to accommodate for what makes the most sense for your career path as long as you keep me updated.

The extended pre-analysis is similar to the pre-analysis plans that most political scientists write. Common templates include the EGAP (http://egap.org/content/how-to-register) and the OSF (https://osf.io/registries) registries. However, for the purposes of this course, I also expect you to include a section discussing the theoretical motivation for your pre-analysis plan. This is mostly so you can get better feedback. More details to come soon.

X% of your final grade will be based on a brief project proposal (around 5 pages) due on Week 7. We will meet on Week 6 to discuss project proposal ideas.

You will also be assigned to provide written feedback on a project proposal, due on week 11. Your feedback should read like a reviewer or discussant report, summarizing the project, emphasizing its importance and contributions, as well as providing realistic and productive to improve. X% of your grade will be based on your written feedback report.

X% of your final grade will be based on the final pre-analysis plan draft at the end of the term.

## Schedule

### Week 1: Introduction/Organization

Druckman, James N. and Arthur Lupia. 2012. "Experimenting with Politics." *Science* 335(6073): 1177-1179

Fujii, Lee Ann. 2012. "Research Ethics 101: Dilemmas and Responsibilities." *PS: Political Science and Politics* 45(4): 717-723

### Week 2: Why Experiments? Back to Basics

Fisher, R.A. 1935. *The Design of Experiments*. London: Oliver and Boyd

Campbell, Donald T. 1957. "Factors relevant to the validity of experiments in social settings." *Psychological Bulletin* 54(4): 297-312

Mook, Douglas G. 1983. "In Defense of External Invalidity". *American Psychologist* 38(4): 379-387

Smith, Herbert L. 1990. "Specification Problems in Experimental and Nonexperimental Social Research." *Sociological Methodology* 20: 59-91

### Week 3: Why Experiments? Modern Debates

Gerber, Alan, Kevin Arceneaux, Cheryl Boudreau, Conor Dowling, Sunshine Hillygus, Thomas Palfrey, Daniel R. Biggers, and David J. Hendry. 2014. "Reporting Guidelines for Experimental Research: A Report from the Experimental Research Section Standards Committee." *Journal of Experimental Political Science* 1(1): 81-98

Mutz, Diana and Robin Pemantle. 2015. "Standards for Experimental Research: Encouraging a Better Understanding of Experimental Methods." *Journal of Experimental Political Science* 2(2): 192-215

Giacomini, Mita. 2009. "Theory-Based Medicine and the Role of Evidence: Why the Emperor Needs New Clothes, Again" *Perspectives in Biology and Medicine* 52(2): 234-251

Deaton, Angus and Nancy Cartwright. 2018. "Understanding and misunderstanding randomized controlled trials." *Social Science & Medicine* 210: 2-21

Green, Donald P. and Alan S. Gerber. 2002. "The Downstream Benefits of Experimentation." *Political Analysis* 10(4): 394-402

Druckman, James N., and Cindy D. Kam. 2011. "Students as Experimental Participants." In *Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 41–57. Cambridge: Cambridge University Press

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2(2): 109-138

### Week 4: Learning from Experiments

Open Science Collaboration. 2015. "Estimating the reproducibility of psychological science." *Science* 349(6251): aac4716

Benjamin, Daniel J., James O. Berger, Magnus Johannesson, et al. 2018. "Redefine statistical significance." *Nature Human Behavior* 2: 6-10

McShane, Blakeley B., David Gal, Andrew Gelman, Christian Robert, and Jennifer L. Tackett. 2019. "Abandon Statistical Significance." *The American Statistician* 73(sup1): 235-245

Stroebe, Wolfgang. 2019. "What Can We Learn from Many Labs Replications?" *Basic and Applied Social Psychology* 41(2): 91-103

Dunning, Thad. 2016. "Transparency, Replication, and Cumulative Learning: What Experiments Alone Cannot Achieve." *Annual Review of Political Science* 19: S1-S23

Duch, Raymond, Denise Laroze, Thomas Robinson, and Pablo Beramendi. 2020. "Multi-Modes for Detecting Experimental Measurement Error." Political Analysis 28(2): 263–283

Bowers, Jake and Paul Testa. 2019. "Better Government, Better Science: The Promises and Challenges Facing the Evidence-Informed Policy Movement." *Annual Review of Political Science* 22: 521-542

### Week 5: Planning Experiments

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

Laitin, David D. 2013. "Fisheries Management." *Political Analysis* 21(1): 42-47

Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Lawrence F. Katz, Benjaming A. Olken, and Anja Sautmann. 2020. "In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plans for RCTs in Economics." NBER Working paper No. 26993

Blair, Graeme, Jasper Cooper, Alexander Coppock, and Macartan Humphreys. 2019. "Declaring and Diagnosing Research Designs." *American Political Science Review* 113(3): 838-859

Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do About It." *American Journal of Political Science* 62(3): 760-775

Slough, Tara. 2019. "On Theory and Identification: When and Why We Need Theory for Causal Identification." Working paper

### Week 6: Discuss Ideas for Pre-Analysis Plan

Chen, Noule and Christopher Grady. "10 Things to Know About Pre-Analysis Plans." EGAP Methods Guides

### Week 7: Randomization (Proposal draft due)

Kasy, Maximilian. 2016. "Why Experimenters May Not Always Want to Randomize, and What They Could Do Instead." *Political Analysis* 24(3): 324-338

Wu, Jason and Peng Ding. 2020. "Randomization Tests for Weak Null Hypotheses in Randomized Experiments." *Journal of the American Statistical Association*

Li, Xinran, Peng Ding, and Donald B. Rubin. 2020. "Rerandomization in $2^K$ factorial experiments." *Annals of Statistics* 48(1): 43-63

Miratrix, Luke, Jasjeet S. Sekhon, and Bin Yu. 2013. "Adjusting treatment effect estimates by post-stratification in randomized experiments." *Journal of the Royal Statistical Society. Series B* 75(2): 369-396

Moore, Ryan T. 2012. "Multivariate Continuous Blocking to Improve Political Science Experiments." *Political Analysis* 20(4): 460-479

Moore, Ryan T and Sally A. Moore. 2013. "Blocking for Sequential Political Experiments." *Political Analysis* 21(4): 507-523

Green, Donald P. and Lynn Vavreck. 2008. "Analysis of Cluster-Randomized Experiments: A Comparison of Alternative Estimation Approaches." *Political Analysis* 16(2): 138-152 [Perhaps replace with something addressing when and why to use cluster randomization]

### Week 8: Field Experiments I (Methods)

Esterling, Kevin M., Michael A. Neblo, and David M.J. Lazer. 2011. "Estimating Treatment Effects in the Presence of Noncompliance and Nonresponse: The Generalized Endogenous Treatment Model." *Political Analysis* 19(2): 205-226

Coppock, Alexander, Alan S. Gerber, Donald P. Green, and Holger L. Kern. 2017. "Combining Double Sampling and Bounds to Address Nonignorable Missing Outcomes in Randomized Experiments." *Political Analysis* 25(2): 188-206

Marbach, Moritz and Dominik Hangartner. 2020. "Profiling Compliers and Noncompliers for Instrumental-Variable Analysis.' *Political Analysis* 28(3): 435-444

Egami, Naoki and Erin Hartman. 2019. "Covariate Selection for Generalizing Experimental Results: Application to a Large-Scale Development Program in Uganda." Working paper. https://arxiv.org/abs/1909.02669

Aronow, Peter M., Dean Eckles, Cyrus Samii, and Stephanie Zonszein. 2020. "Spillover Effects in Experimental Data." Working paper. https://arxiv.org/abs/2001.05444

### Week 9: Field Experiments II (Applications)

Broockman, David E., Joshua L. Kalla, and Jasjeet S. Sekhon. 2017. "The Design of Field Experiments With Survey Outcomes: A Framework for Selecting More Efficient, Robust, and Ethical Designs." *Political Analysis* 25(4): 435-464

Slough, Tara. 2019. "The Ethics of Electoral Experimentation: Design-Based Recommendations." Working paper. http://taraslough.com/assets/pdf/eee.pdf

Butler, Daniel, and Jonathan Homola. 2017."An Empirical Justification for the Use of Racially Distinctive Names to Signal Race in Experiments" *Political Analysis* 25(1): 122-130

Coppock, Alexander. 2019. "Avoiding Post-Treatment Bias in Audit Experiments" *Journal of Experimental Political Science* 6(1): 1-4

Grossman, Guy, and Kristin Michelitch. 2018. "Information Dissemination, Competitive Pressure, and Politician Performance Between Elections: A Field Experiment in Uganda." *American Political Science Review* 112(2): 280-301

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, et al. 2015. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries" Science 348(6236): 1260799

Rossiter, Erin. 2020. "The Consequences of Interparty Conversation on Outparty Affect and Stereotypes." Working paper

### Week 10: Placeholder for mid-semester break

### Week 11: Survey Experiments for Causal Inference

Gaines, Bryan, James H. Kuklinski, and Paul J. Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15(1): 1-20

Brutger, Ryan, Joshua D. Kertzer, Jonathan Renson, Dustin Tingley, and Chagai M. Weiss. 2020. "Abstraction and Detail in Experimental Design." Working paper [Very early draft, check back later

for a stable version]

Mummolo, Jonathan and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113(2): 517-529

Miratrix, Luke, Jasjeet S. Sekhon, Alexander G. Theodoridis, and Luis F. Campos. 2018. "Worth Weighting? How to Think About and Use Weights in Survey Experiments." *Political Analysis* 26(3): 275-291

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating vignette and conjoint experiments against real-world behavior." *Proceedings of the National Academy of Sciences* 112(8): 2395-2400

De la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. 2020. "Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution." *Political Analysis*

Abramson, Scott, Korhan Koçak, and Asya Magazinik. 2019. "What Do We Learn About Voter Preferences from Conjoint Experiments?" Working paper

### Week 12: Survey Experiments to Measure Sensitive Attitudes

Schwarz, Norbert. 1999. "Self-Reports: How the Questions Shape the Answers." *American Psychologist* 54(2): 93-105

Schuldt, Jonathon P., Sara H. Konrath, and Norbert Schwarz. 2011. "'Global warming' or 'climate change'? Whether the planet is warming depends on question wording." *Public Opinion Quarterly* 75(1): 115-124

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "When to Worry About Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments." *American Political Science Review*

Rosenfeld, Bryn, Kosuke Imai, and Jacob N. Shapiro. 2016. "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions." *American Journal of Political Science* 60(3): 783-802

Chou, Winston, Kosuke Imai, and Bryn Rosenfeld. 2017. "Sensitive Questions with Auxiliary Information." *Sociological Methods & Research* 49(2): 418-454

Blair, Graeme, Kosuke Imai, and Jason Lyall. 2014. "Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan." *American Journal of Political Science* 58(4): 1043-1063

Alvarez, R. Michael, Lonna Rae Atkeson, Ines Levin, Yimeng Li. "Paying Attention to Inattentive Survey Respondents." *Political Analysis* 27(2): 145-162

### Week 13: Laboratory Experiments in Behavorial Economics

Del Ponte, Alessandro, Reuben Kline, and John Ryan. 2020. "Behavioral Analysis in the Study of Politics: The Conflict Laboratory." In *Oxford Encyclopedia of Political Decision Making*

Belot, Michele, Raymond Duch, and Luis Miller. 2015. "A comprehensive comparison of students and non-students in classic experimental games." *Journal of Economic Behavior & Organization* 113(1): 26-33

Krupnikov, Yanna and John Barry Ryan. 2017. "Choice vs. Action: Candidate Ambiguity and Voter Decision Making." *Quarterly Journal of Political Science* 12(4): 479-505

Andrews, Talbot M., Andrew W. Delton, and Reuben Kline. 2018. "High-risk high-reward investments to mitigate climate change." *Nature Climate Change* 8: 890-894

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision." *American Political Science Review* 101(4): 709-725

### Week 14: Laboratory Experiments in Psychology

Bortolotti, Lisa, and Matteo Mameli. 2006. "Deception in psychology: moral costs and benefits of unsought self-knowledge." *Accountability in Research* 13(3): 259-275

Westfall, Jacob, David A. Kenny, and Charles M. 2014. "Statistical Power and Optimal Design in Experiments in Which Samples of Participants Respond to Samples of Stimuli." *Journal of Experimental Psychology: General* 143(5): 2020-2045

Brysbaert, Marc. 2019. "How Many Participants Do We Have to Include in Properly Powered Experiments? A Tutorial of Power Analysis with Reference Tables." *Journal of Cognition* 2(1): 1-38

Lamprianou, Iasonas, and Antonis A. Ellinas. 2019. "Emotion, Sophistication and Political Behavior: Evidence from a Laboratory Experiment." *Political Psychology* 40(4): 859-876

Bakker, Bert N., Gijs Schumacher, Claire Gothreau, and Kevin Arceneaux. 2020. "Conservatives and liberals have similar physiological responses to threats." *Nature Human Behavior* 4: 613-621

### Week 15: Heterogeneous Treatment Effects and Causal Mechanisms

Ding, Peng, Avi Feller, and Luke Miratrix. 2019. "Decomposing Treatment Effect Variation." *Journal of the American Statistical Association* 114(525): 304-317

Imai, Kosuke and Aaron Strauss. 2011. "Estimation of Heterogeneous Treatment Effects from Randomized Experiments, with Application to the Optimal Planning of the Get-Out-the-Vote Campaign." *Political Analysis* 19(1): 1-19

Gaines, Brian J. and James H. Kuklinski. 2011. "Experimental Estimation of Heterogeneous Treatment Effects Related to Self-Selection." *American Journal of Political Science* 55(3): 724-736

Imai, Kosuke, Dustin Tingley, and Teppei Yamamoto. 2013. "Experimental designs for identifying causal mechanisms." *Journal of the Royal Statistical Society. Series A* 176(1): 5-51

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. "Analyzing Causal Mechanisms in Survey Experiments." *Political Analysis* 26(4): 357-378

### Week 16: Experiments and Computational Social Science

Efron, Bradley. 2007. "Size, power and false discovery rates." *Annals of Statistics* 1351-1377

Imai, Kosuke and Marc Ratkovic. 2013. "Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation." *Annals of Applied Statistics* 7(1): 443-470

Shiraito, Yuki. 2016. "Uncovering Heterogeneous Treatment Effects." Working paper.

Imai, Kosuke and Michael Lingzhi Li. 2020. "Experimental Evaluation of Individualized Treatment Rules." Working paper

Offer-Westort, Molly, Alexander Coppock, and Donald P. Green. 2020. "Adaptive Experimental Design: Prospects and Applications to Political Science." Working paper

**Placeholder for Pre-analysis Plan Deadline**