PS 522 Research Design

Spring 2026
Tuesdays 9:30 - 11:50 a.m.
404 David Kinley Hall
4.0 Credits

Jake Bowers	Matthew S. Winters
jwbowers@illinois.edu	mwinters@illinois.edu
432 David Kinley Hall	324 Coble Hall
Office Hours by Appointment	Office Hours: Wednesday 2:30 - 4:30 p.m.
Sign up at https://calendly.com/jakebowers	Sign up at https://calendly.com/mswinters1

A Brief Recent History of PS522

For many years, Jim Kuklinski taught this course alone. As his retirement neared, the department figured no one could truly replace him and that it would be a good idea for the course to be co-taught. For the last several years, some combination of two of Bowers, Gaines, Winters, and Wong have taught this course.

Course Description[1]

We have designed this course to be deliberately very broad. Part of the design stage of research involves identifying tractable, interesting questions and developing theoretically interesting explanations that answer those questions. Researchers must also plan collection and analysis of data as a part of research design. Although this course does not teach about statistical methods in depth, we will discuss different statistical methods, aiming for conceptual understanding of those methods and their value, rather than a mathematical or computational understanding. We focus on the kinds of comparisons and measurements we aim to make, rather than on the precise calculations that we use to execute those comparisons or measurements.

Researchers can also design more or less accurate and/or efficient ways of describing research findings. We expect to discuss aspects of what makes for good (scientific) writing from time to time.

The content and organization of this course have changed over the years, not only with changes of instructor, but also because political scientists have, over time, rethought the foundations of empirical research. To quote Jim Kuklinski:

> Until about the end of the 1950s, political scientists would commonly undertake an in-depth field study of a phenomenon in a particular country or region of a country. The emphasis was detailed understanding of a specific context. A particularly powerful critique of this work was its limited capacity to reach conclusions that apply across units.

[1] Make sure you have the current version of the syllabus. We will probably change it throughout the term.

> This critique, combined with a growing availability of quantitative data and statistical methods to analyze them, led to the emergence of the statistical analysis of relatively large data sets. For more than half a century, political scientists would rely heavily on regression analysis and related methods to make general causal claims. In retrospect, it is now clear that the regression approach emphasized the "general" more than the "causal." Currently, more and more scholars are questioning the utility of this relative emphasis: "What good is generalization if the estimated causal effect is wrong?" (Of course, one can easily reverse the question: "What good is showing cause and effect if it cannot be generalized beyond the specific cases?")

The discipline now pays much more attention to the difficulties associated with inferring cause and effect from patterns of observed association than was common when the instructors were PhD students. Since randomized experiments offer one clear approach to causal inference, these days researchers from across subfields of political science use more experiments and more diverse experiment-inspired research designs than before. Given the increasingly careful thought around causal inference across the social sciences and since the understanding of causal inference in general (via randomized experiments or not) requires an understanding of the logic of randomization, experiments occupy a privileged position in the course.

The academic community, commentators and pundits, casual followers of politics, and practitioners of politics will often define what are and are not adequate explanations of political phenomena. Our goal this term is not so much to give you a set of best practices that represent a consensus in the field but rather to identify some of the main issues in research design so that you can approach your research (and the research done by others that you consume) thoughtfully.

Course Goals

We intend for this course to increase your understanding of:;

- current disciplinary standards and expectations;
- research question formulation;
- the connection between question formulation, concept formulation, theory formulation, and research design; and
- the core challenges in research design and analysis.

Course Requirements

This is not a lecture course. The instructors expect to be involved in discussion, steering it to some degree, but we also expect each week to play out largely as discussion among students. To that end, students must do the reading in advance. We have tried to select serious, important, and helpful readings and to limit the number of readings so that each week everyone can read every assigned article. The "Additional Readings" are pieces that we think are worth knowing about, but we do not expect you to have read them for a given class meeting.

<u>Class Participation 30%</u>

Quality matters much more than quantity, but if you rarely speak, neither you nor your colleagues benefit. In-class discussion should be lively but also civil.

We may begin classes with ungraded quizzes on the readings.

Weekly Reading Exercises 30%
You will be required to apply the core concepts engaged by the readings to your own work and/or your own substantive interests each week. We expect that these will be (about) one single-spaced page each. These exercises should help make the abstract discussions in the readings of measurement, theory, causality, inference, and explanation more concrete in your mind and thus help you take one step towards improving your own thinking about research design.

These essays will help the instructors guide in-class discussion and should be posted to the Canvas website by 11:59 p.m. CT on the Sunday before class.

These exercises will be graded on a satisfactory / unsatisfactory basis.

Final Test 40%
We expect the material covered in PS522 to inform your research decisions and to help you critique research and provide research guidance to students for years into the future. In order to give you the opportunity to synthesize material from across the semester and to revisit topics from throughout the semester, we will administer a cumulative final test on the last day of class. We currently plan to allow you to bring in six (6) pages of notes for the test. The test will consist of a series of short-answer questions dealing with critical ideas discussed over the course of the semester.

General Policies

Emergency Response
The University of Illinois has a set of guidelines and policies relating to emergencies. Please review these here: https://police.illinois.edu/em/run-hide-fight/

Students with Disabilities
The instructors will attempt to make appropriate accommodations for students with disabilities. As described in the University of Illinois Student Code, these accommodations can be coordinated through the Division of Disability Resources and Educational Services (DRES). The DRES Student Services Office is reachable at 217-333-4603 or disability@illinois.edu. Students must notify the instructors and provide proper documentation during the first week of class.

Academic Integrity
Any act of academic dishonesty or misconduct will be penalized. Please refer to Art. 1, Part 4 of the Student Code at the University of Illinois.

AI and Fair Credit Policy
This class is about how we work—and part of working well is understanding which tools help us think versus which tools substitute for our thinking. In this class, we ask you to list the tools you use for each piece of work you turn in: AI/LLMs (including how you used them), text editors, software packages, citation managers, and so on. For an example, see the acknowledgements footer at jakebowers.org. For another example, see this conversation Jake had with the Claude AI to develop this policy. If you use the same tools across assignments, you can write "Same tools as before" and note any new additions.

For AI specifically: the weekly exercises exist to help you push through the discomfort of connecting abstract ideas to your own research. That discomfort is where learning happens. If you

use AI in ways that bypass the struggle—rather than support it—you will produce acceptable-looking text and miss the point. We trust you to make that judgment.

Why the limited-note exam matters: The course concludes with a limited-note exam that you must pass to receive a passing grade. This exam exists for your benefit: preparing for it pushes you to synthesize the core concepts that thread through the course, and our feedback tells you whether you've misunderstood something important before it shows up in your dissertation prospectus. The exam allows us to adopt a trust-based AI policy for everything else—you have every incentive to actually learn the material, because you'll need to demonstrate that learning without assistance at the end.

Student Conduct

Students are expected to behave in accordance with the penal and civil statutes of all applicable local, state, and federal governments, with the rules and regulations of the Board of Regents, and with university regulations and administrative rules. For more information about the student code and handbook, see the CITL course policies page.

Inclusivity

In line with the Department of Political Science's commitment to create a community of care and inclusivity ("A Commitment to Equity, Diversity, and Inclusion"), it is our conviction that our shared learning experience is greatly enriched when students from diverse backgrounds and perspectives find a positive and safe environment. Respect for different viewpoints must be a chief principle during this class, and we expect all students to maintain and nurture this environment.

Land Acknowledgement

As a land-grant institution, the University of Illinois at Urbana-Champaign has a responsibility to acknowledge the historical context in which it exists. We are currently on the lands of the Peoria, Kaskaskia, Peankashaw, Wea, Miami, Mascoutin, Odawa, Sauk, Mesquaki, Kickapoo, Potawatomi, Ojibwe, and Chickasaw Nations. It is necessary for us to acknowledge these Native Nations and for us to work with them as we move forward as an institution with Native peoples at the core of our efforts.

Student Wellness Resources

The University of Illinois strives to promote student success through the support of student psychological and emotional well-being. Please take advantage of the resources listed on the Student Affairs website.

Sexual Misconduct Policy and Reporting

The University of Illinois is committed to combating sexual misconduct. Faculty and staff members are required to report any instances of sexual misconduct to the university's Title IX and Disability Office. In turn, an individual with the Title IX and Disability Office will provide information about rights and options, including accommodations, support services, the campus disciplinary process, and law enforcement options.

A list of the designated university employees who, as counselors, confidential advisors, and medical professionals, do not have this reporting responsibility and can maintain confidentiality, can be found in the Confidential Resources section. Other information about resources and reporting is available at wecare.illinois.edu.

COVID-19 Classroom Policy
Following University policy, all students are required to engage in appropriate behavior to protect the health and safety of the community. Students are also required to follow the campus COVID-19 protocols, as they evolve.

Weekly Schedule

1. January 20: Course Introduction and the State of Knowledge Accumulation in the Social Sciences

Week aim: You have joined this class because you want to produce excellent research. Do you have a good sense for how to know when a piece of writing is research versus when it is not research? What about excellence? How would you know that your research is excellent? This week, we read some authors who take stock of evolving standards for what makes for excellent political science research. From these readings, you should gain a sense of how standards in political science have evolved over the last two decades and what some of the concerns about the changes in those standards are.

Week exercise due before class on canvas: Find an article published in a peer-reviewed journal about which you have strong emotions: you love this article, you hate this article, this article makes you sad or delighted. Explain in no more than one page why this article matters for political science or social science in general: after reading this article, how should other scholars change their own thinking and writing? Come to class prepared to say the name of the article, one sentence about your reaction, and one sentence about why the article matters.

Required Readings

Samii, Cyrus. 2016. Causal Empiricism in Quantitative Research. *The Journal of Politics* 78 (3): 941–55. https://doi.org/10.1086/686690.

Torreblanca, Carolina, William Dinneen, Guy Grossman, and Yiqing Xu. 2025. The Credibility Revolution in Political Science. Preprint. https://doi.org/10.31235/osf.io/w2kmc_v1.

Spirling, Arthur, and Brandon M. Stewart. 2025. What Good Is a Regression? Inference to the Best Explanation and the Practice of Political Science Research. *The Journal of Politics* 87 (4): 1587–99. https://doi.org/10.1086/734280.

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. https://doi.org/10.1353/pbm.0.0088

Korinek, Anton. 2025. "AI Agents for Economic Research: August 2025 Update to 'Generative AI for Economic Research: Use Cases and Implications for Economists,' published in the Journal of Economic Literature 61(4). https://doi.org/10.1257/jel.20231736 (Follow the link to the updated version)

Additional Readings

Gelman, Andrew and Eric Loken. 2014. The Statistical Crisis in Science. *American Scientist* 102(6): 460.

Christensen, Garret, and Edward Miguel. 2018. Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature* 56.3: 920–80. https://doi.org/10.1257/jel.20171350

Voelkel, Jan G., Michael N. Stagnaro, James Y. Chu, Sophia L. Pink, Joseph S. Mernyk, Chrystal Redekopp, Isaias Ghezae, et al. 2024. Megastudy Testing 25 Treatments to Reduce Antidemocratic Attitudes and Partisan Animosity. *Science* 386.6719: eadh4764. https://doi.org/10.1126/science.adh4764

Ruggeri, Kai, Friederike Stock, S. Alexander Haslam, Valerio Capraro, Paulo Boggio, Naomi Ellemers, Aleksandra Cichocka, et al. 2024. A Synthesis of Evidence for Policy from Behavioural Science during COVID-19. *Nature* 625.7993: 134–47. https://doi.org/10.1038/s41586-023-06840-9

Korbmacher, Max, Flavio Azevedo, Charlotte R. Pennington, Helena Hartmann, Madeleine Pownall, Kathleen Schmidt, Mahmoud Elsherif, et al. 2023. The Replication Crisis Has Led to Positive Structural, Procedural, and Community Changes. *Communications Psychology* 1.1: 3. https://doi.org/10.1038/s44271-023-00003-2

2. January 27: Political Science Inquiry. What Does It Mean to Compare and Explain?

Week aim: Different research traditions in political science have different standards for what counts as a compelling comparison or explanation. By the end of this week, you should be able to articulate what these different standards are and why they matter. When a reviewer or colleague asks why your research design is credible, you need to have an answer.

Week exercise: Choose a research question from your own substantive area. How would you know if you had answered it well? What would a skeptic demand as evidence? With reference to this week's readings, write one page describing what a convincing answer would look like.

Required Readings

Rosenbaum, Paul R. 2020. *Design of Observational Studies, 2nd Ed*. New York, NY: Springer. (Chapter 1)

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton and Oxford: Princeton University Press. (Chapter 1)

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press. (Chapter 1)

Ashworth, Scott, Christopher R. Berry, and Ethan Bueno de Mesquita. 2021. *Theory and Credibility: Integrating Theoretical and Empirical Social Science*. Princeton: Princeton University Press. (Chapters 1 and 2)

Gerring, John. 2004. What Is a Case Study and What Is it Good for? *American Political Science Review* 98.2: 341-354.

Additional Readings

Humphreys, Macartan, and Alan M. Jacobs. 2023. *Integrated Inferences: Causal Models for Qualitative and Mixed-Method Research*. New York: Cambridge University Press. (Chapters 1–4)

Blair, Graeme, Alexander Coppock, and Macartan Humphreys. 2023. *Research Design in the Social Sciences: Declaration, Diagnosis, and Redesign*. Princeton: Princeton University Press. (Chapters 2 and 3)

3. February 3: Theory and Observation

Week aim: Theory is not the same as description, and more detail is not the same as more insight. This week you should learn to distinguish genuine theoretical arguments from elaborate descriptions dressed up as theory. A good theory simplifies; a bad theory just adds words.

Week exercise: Identify a theoretical claim in your substantive field. In one page and with reference to the assigned readings for this week, evaluate whether it is genuinely a theory or merely a description. What does it rule out? What would it take to prove it wrong?

Required Readings

Lave, Charles A. and James G. March. 1975. *An Introduction to Models in the Social Sciences*. Lanham: University Press of America. (Chapters 1 and 2)

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press. (Chapter 2)

Healy, Kieran. 2017. Fuck Nuance. *Sociological Theory* 35.2: 118-127.

Sutton, Robert I. and Barry M. Staw. 1995. What Theory is Not. *Administrative Science Quarterly* 40.3: 371-384.

Additional Readings

Hal Varian. 1989. "What Use is Economic Theory?" Unpublished paper. https://people.ischool.berkeley.edu/~hal/Papers/theory.pdf

Rory Smead. 2013. "A Brief Introduction to the Basics of Game Theory," Northeastern University, https://joelvelasco.net/teaching/5330/GameTheoryBasics.pdf

Rosenbaum, Paul R. 2017. *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press. (Chapter 7)

Ashworth, Scott, Christopher R. Berry, and Ethan Bueno de Mesquita. 2021. *Theory and Credibility: Integrating Theoretical and Empirical Social Science*. Princeton: Princeton University Press. (Chapter 4)

4. February 10: Explanation and Causality

Week aim: Political scientists ask and answer "why" questions using the language of causality. If you use such explanations, you need to understand what you **mean** when you talk about cause and effect. Notice that we are here talking about **both** the theoretical explanations that articulate causal mechanisms **and** the evidence that a comparison tells us something clear about a theorized causal relationship.

Week exercise: Find a published article in your field that makes a causal claim. What do the authors mean by "cause"? If they are using a counterfactual conception, which of Holland's assumptions are they using? If they are not using a counterfactual understanding, how are they thinking about causal relations? If you have space on your page, explain at least one of the challenges that anyone would face in providing evidence in favor of and/or against this causal explanation.

Required Readings

Holland, Paul W. 1986. Statistics and Causal Inference (with Discussion). *Journal of the American Statistical Association* 81.396: 945-970.

Brady, Henry E. 2008. Causation and Explanation in Social Science. In *The Oxford Handbook of Political Methodology* (Oxford Handbooks of Political Science).

Rosenbaum, Paul R. 2017. *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press. (Chapter 3)

Daoud, Adel and Devdatt Dubhashi. 2023. Statistical Modeling: The Three Cultures. *Harvard Data Science Review*, 5(1). [link]

Additional Readings

Mackie, J. L. 1965. Causes and Conditions. *American Philosophical Quarterly* 2.4: 245–64.

Fearon, James. 1991. Counterfactuals and Hypothesis Testing in Political Science. *World Politics* 43: 169-195.

Mahoney, James. 2000. Strategies of Causal Inference in Small-N Analysis. *Sociological Methods & Research* 28(4): 387–424.

Breiman, Leo. 2001. Statistical Modeling: The Two Cultures. *Statistical Science* 16.3: 199-231.

See also the commentaries in *Observational Studies* Vol 7, Issue 1, 2021). [link]

Shmueli, Galit. 2010. To explain or to predict? *Statistical science*: 289-310.

Gerring, John. 2010. Causal Mechanisms: Yes, But. . . *Comparative Political Studies* 43.11: 1499-1526.

Goertz, Gary and James Mahoney. 2013. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences.* Princeton University Press.

Freese, Jeremy, and J. Alex Kevern. 2013. Types of Causes. In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, 27–41. Handbooks of Sociology and Social Research. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3

Barringer, Sondra N., Scott R. Eliason, and Erin Leahey. 2013. A History of Causal Analysis in the Social Sciences. In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, 9–26. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_2

Slater, Dan, and Daniel Ziblatt. 2013. The Enduring Indispensability of the Controlled Comparison. *Comparative Political Studies* 46(10): 1301–27. https://doi.org/10.1177/0010414012472469

5. February 17: Conceptualization and Measurement

Week aim: Before you can measure something, you must decide what it is. Measurement choices are not neutral—they shape your conclusions. This week you should learn to recognize the differences between an abstract concept and a concrete indicator, to provide a justified basis for making choices about concepts and indicators, and to understand that your operationalization is a choice you must defend. There is a key question that everyone must be able to answer about their measurements: "How do I know that this observation (a number in an indicator, the response of a person to a question, etc) **means** what I say it means?"

Week exercise: Identify a concept central to your research. In one page and with reference to the readings, explain what it means, how you would measure it, and how you would know if your measure is valid. What alternative operationalizations exist, and why might they give different answers?

Required Readings

Jackman, Simon. 2008. Measurement. *The Oxford Handbook of Political Methodology* (Oxford Handbooks of Political Science). Janet M. Box-Steffensmeier, et al., eds. Oxford and New York, NY: Oxford University Press. (Chapter 6)

Adcock, Robert and David Collier. 2001. Measurement Validity: A Shared Standard for Qualitative and Measurement Validity. *American Political Science Review* 95.3: 529-546.

Slough, Tara. 2020. 10 Things to Know About Measurement in Experiments. Evidence in Governance and Politics. https://egap.org/resource/10-things-to-know-about-measurement-in-experiments/

Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. Democracy and Dictatorship Revisited. *Public Choice* 143 (1-2): 67-101.

Coppedge, Michael and John Gerring, et al. 2011. Conceptualizing and Measuring Democracy: A New Approach. *Perspectives on Politics* 9 (2): 247-67.

Additional Readings

Carmines, Edward G., and Zeller, Richard A. 1979. *Reliability and Validity Assessment*. Quantitative Applications in the Social Sciences 07–017. Sage.

Elman, Colin. 2005. Explanatory Typologies in Qualitative Studies of International Politics. *International Organization* 59(2): 293-326.

Goertz, Gary. 2008. Concepts, Theories and Numbers: A Checklist for Constructing, Evaluating and Using Concepts or Quantitative Measures. *The Oxford Handbook of Political Methodology* (Oxford Handbooks of Political Science). Janet M. Box-Steffensmeier, et al., eds. Oxford and New York: Oxford University Press. (Chapter 5)

Seawright, Jason and David Collier. 2014. Rival Strategies of Validation: Tools for Evaluating Measures of Democracy. *Comparative Political Studies* 47 (1): 111-38.

Goertz, Gary. 2020. *Social Science Concepts and Measurement*. New and Completely Revised Edition. Princeton University Press.

6. February 24: Elite Interviews, Ethnography, Participant Observation

Week aim: Some things can only be learned by being there. This week you should understand when and why field-based qualitative data collection—interviews, ethnography, observation—can answer questions that surveys and experiments cannot. You should also recognize that field research has its own standards of rigor and its own challenges.

Week exercise: Find an article in your field that uses direct interviews and/or observations with or without statistical analysis of large datasets. In one page, describe how those authors engaged (explicitly or implicitly) with the standards of rigor of such work. Does this piece either generate compelling new explanations and/or hypotheses or provide compelling evidence for theorized causal mechanisms? Does it do something else?

Required Readings

Wood, Elisabeth Jean. 2007. Field Research. In Carles Boix and Susan Stokes (eds.), *The Oxford Handbook of Comparative Politics*. Oxford, UK, Oxford University Press, pp. 123-146.

Cramer, Katherine J. 2016. *The Politics of Resentment : Rural Consciousness in Wisconsin and the Rise of Scott Walker*. The University of Chicago Press. Chapter 2 (pp. 26-44).

Fujii, Lee Ann. 2015. Five Stories of Accidental Ethnography: Turning Unplanned Moments in the Field into Data. *Qualitative Research* 15.4: 525-539.

From Cyr, Jennifer and Sarah Wallace Goodman. 2024. *Doing Good Qualitative Research*. Oxford University Press:

- Li, Lantian. Interviewing Elites, pp. 183-194.
- Cyr, Jennifer. Focus Groups, pp. 222-232.

- Pisano, Jessica. Ethnography, pp. 233-244
- Thorson, Emily and Emily M. Fariss. Supplementing Qualitative Work with Surveys, and Vice Versa, pp. 245-254.

Additional Readings

Aberbach, Joel and Bert A. Rockman. 2002. Conducting and Coding Elite Interviews. *PS: Political Science and Politics* 35.4: 673-676.

Cramer, Katherine. 2012. Putting Inequality in Its Place: Rural Consciousness and the Power of Perspective. *American Political Science Review* 106.3: 517-532.

Yanow, Dvora and Peregrine Schwartz-Shea. 2016. Encountering Your IRB 2.0: What Political Scientists Need to Know. *PS: Political Science and Politics* 49.2: 277-286.

Lewis, Janet I. 2016. How Does Ethnic Rebellion Start? *Comparative Political Studies*, 50(10), 1420-1450. https://doi.org/10.1177/0010414016672235

7. March 3: Experimental Logic

Week aim: Randomization allows us to use statistical tools to directly engage the fundamental problem of causal inference. This week you should understand why this is true, what it buys you, and what it does not (i.e., the limitations of randomization). You should also learn that experimental design involves real choices about efficiency, ethics, and what quantities you can estimate.

Week exercise: Find an article in your field that uses a randomized experiment to answer a causal question. In one page, describe the treatment, the randomization, and the outcome. What would you learn? What remained uncertain or unclear?

Required Readings

Gerber, Alan S. and Donald P Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* New York, NY: W.W. Norton. (Chapters 1 & 2)

Rosenbaum, Paul R. 2017. *Observation and Experiment: An Introduction to Causal Inference.* Cambridge, MA: Harvard University Press, (Chapters 1-4)

Fisher, R.A. 1935. *The Design of Experiments*. Edinburgh: Oliver and Boyd. (Chapters 1 & 2)

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.

Additional Readings

Gaines, Brian J. and James H. Kuklinski. 2011. Treatment Effects. In James N. Druckman et al, eds. *Cambridge Handbook of Experimental Political Science*. New York, NY: Cambridge University Press: 445-458.

Bowers, Jake. 2011. Making Effects Manifest in Randomized Experiments. In James N. Druckman et al, eds. *Cambridge Handbook of Experimental Political Science*. New York, NY: Cambridge University Press: 459-480.

Deaton, Angus S. 2009. Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development. National Bureau of Economic Research. http://www.nber.org/papers/w14690.

Deaton, Angus S. 2020. Introduction: Randomization in the Tropics Revisited, a Theme and Eleven Variations. In Florent Bédécarrats, Isabelle Guérin, and François Roubaud, eds. *Randomized Control Trials in the Field of Development: A Critical Perspective*. Oxford University Press: 29-46.

Dietrich, Simone, Heidi Hardt, and Haley J. Swedlund. 2021. How to Make Elite Experiments Work in International Relations. *European Journal of International Relations* 27.2: 596-621.

8. March 10: Survey Experiments

Week aim: Survey experiments let you randomize at scale, and many lab experiments are actually survey experiments, but they come with tradeoffs. This week you should learn to think carefully about what your treatment manipulates, whether your vignettes are realistic enough to matter, whether findings from surveys generalize to real-world behavior, and whether such generalization should matter from the perspective of asking and answering "why" questions.

Week exercise: Find a survey experiment in your field. In one page, evaluate the treatment. What exactly does it change or manipulate? Is that the right thing to study via a randomized manipulation? If so, why? If not, why not? How does this experiment help us learn about a theoretical explanation? Would you expect the effect to appear outside the survey context? Should the reader care about generalization in this case? Why or why not?

Required Readings

Gaines, Brian J., James H, Kuklinski, and Paul J. Quirk. 2007. The Logic of the Survey Experiment Reexamined. *Political Analysis* 15.1: 1–20.

Dafoe, Allan, Baobao Zhang, and Devin Caughey. 2018. Information Equivalence in Survey Experiments. *Political Analysis* 26.4: 399–416.

Brutger, Ryan, Joshua D. Kertzer, Jonathan Renshon, Dustin Tingley, and Chagai M. Weiss. 2023. Abstraction and Detail in Experimental Design. *American Journal of Political Science* 67.4: 979–95.

Mummolo, Jonathan, and Erik Peterson. 2019. Demand Effects in Survey Experiments: An Empirical Assessment. *American Political Science Review* 113.2: 517–29.

Weitz-Shapiro, Rebecca, and Matthew S. Winters. 2017. Can Citizens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil. *Journal of Politics* 79 (1): 60–74.

Boas, Taylor C., Daniel Hidalgo, and Marcus André Melo. 2019. Norms Versus Action: Why Voters Fail to Sanction Malfeasance in Brazil. *American Journal of Political Science* 63.2: 385- 400.

Additional Readings

Sniderman, Paul M. 2011. The Logic and Design of the Survey Experiment: An Autobiography of a Methodological Innovation. In *Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 1st ed., 102–14. Cambridge University Press. https://doi.org/10.1017/CBO9780511921452.008.

Sniderman, Paul M. 2018. Some Advances in the Design of Survey Experiments. *Annual Review of Political Science* 21.1: 259–75. https://doi.org/10.1146/annurev-polisci-042716-115726.

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. Analyzing Causal Mechanisms in Survey Experiments. *Political Analysis* 26.4: 357–78.

Barabas, Jason and Jennifer Jerit. 2010. Are Survey Experiments Externally Valid? *American Political Science Review* 104.2: 226–242.

Incerti, Trevor. 2020. Corruption Information and Vote Share: A Meta-Analysis and Lessons for Experimental Design. American Political Science Review 114.3: 761-774.

March 17 — NO CLASS — SPRING BREAK

9. March 24: Natural Experiments, Discontinuities, and Instrumental Variables

Week aim: When you cannot randomize, you may look for situations where nature or policy did something **close to randomization**, where some process in the world breaks the relationship between a theoretically important intervention ("treatment") and alternative explanations ("confounders"). This idea that an intervention is "exogenous" or "controlled" motivates a set of not-primarily-randomized research design ideas such as the use of discontinuities or other devices as "instruments". This week you should connect what we learned about the benefits and limits of actual, known randomization to research designs where researchers **hope** that differences in observations isolate a presumed causal mechanism from confounders rather than **know** that this arises from randomization. The claim that assignment was as-if random is an empirical claim that requires evidence and argument.

Week exercise: Find a study that uses a natural or quasi-experiment. In one page, state the as-if-random assumption and evaluate it. What could violate it? How confident are you?

Required Readings

Angrist, Joshua D., and Jõrn-Steffen Pischke. 2015. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton, NJ: Princeton University Press. (Chapters 2–4) (Chapter 2 is background)

Rosenbaum, Paul R. 2017. *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press, (Chapter 6, 8, & 13)

Dunning, Thad. 2012. Natural Experiments in the Social Sciences: A Design-Based Approach. New York: Cambridge University Press. (Chapters 1–3)

Sekhon, Jasjeet S. and Rocío Titiunik. 2012. When Natural Experiments Are Neither Natural nor Experiments. *American Political Science Review* 106.1: 35-57.

Additional Readings

Angrist, Joshua D., and Jõrn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. (Chapters 4 and 6)

Cattaneo, Matias D, Nicolas Idrobo, and Rocıo Titiunik. 2024. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. New York: Cambridge University Press.

Sovey, Allison J., and Donald P. Green. 2011. Instrumental Variables Estimation in Political Science: A Readers' Guide. American Journal of Political Science 55 (1): 188–200. https://doi.org/10.1111/j.1540-5907.2010.00477.x.

Zubizarreta, José R, Dylan S. Small, and Paul R. Rosenbaum. 2014. Isolation in the Construction of Natural Experiments. *The Annals of Applied Statistics* 8.4: 2096–2121.

Mellon, Jonathan. 2024. Rain, Rain, Go Away: 194 Potential Exclusion-restriction Violations for Studies Using Weather as an Instrumental Variable. *American Journal of Political Science* https://doi.org/10.1111/ajps.12894.

Lal, Apoorva, Mackenzie Lockhart, Yiqing Xu, and Ziwen Zu. 2024. How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies. *Political Analysis* 32(4): 521-540. https://doi.org/10.1017/pan.2024.2

Bove, Vincenzo, Riccardo Di Leo, and Marco Giani. 2024. Military Culture and Institutional Trust: Evidence from Conscription Reforms in Europe. *American Journal of Political Science* 68.2: 714-729.

Mo, Cecilia Hyunjung and Katharine Conn. 2018. When Do the Advantaged See the Disadvantages of Others? A Quasi-Experimental Study of National Service. *American Political Science Review* 112.4: 721-741.

10. March 31: Working with Observational Data in the Absence of an Exogenous Instrument

Week aim: Most political science research uses observational data without a natural experiment. This week you should understand the assumptions required for making causal claims with observational data. These assumptions are often untestable and frequently violated: to the extent that some important figures in the discipline say that we cannot learn anything from observational data. You should also learn that before you estimate anything, you need to define your estimand—what exactly are you trying to learn? The readings for this week should make you uncomfortable. Observational research is harder than it looks. Researcher degrees of freedom generate false positives. Most published work is underpowered. Regression coefficients may not mean what you think. This week you should confront these problems honestly and ask what they mean for your own work.

Week exercise: Find a study in your field that relies on an observational data analysis to try to make causal claims. In one page, discuss whether the authors clearly stated their identifying assumptions. If they did, do you believe that they are met? If they did not, what assumptions would need to hold true in order for them to give a causal interpretation to their estimates?

Required Readings

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.

Gerber, Alan S., Donald P. Green, and Edward H. Kaplan. 2014. The Illusion of Learning from Observational Research. In *Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences*. Dawn Langan Teele, ed. New Haven, CT: Yale University Press. (Chapter 2, pp. 9–32)

Breznau, Nate et al. 2022. Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty. *Proceedings of the National Academy of Sciences* 119.44: 1-8.

Keele, Luke, Randolph T. Stevenson, and Felix Elwert. 2020. The Causal Interpretation of Estimated Associations in Regression Models. *Political Science Research and Methods* 8.1: 1–13. https://doi.org/10.1017/psrm.2019.31.

Rosenbaum, Paul R. 2017. *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press, (Chapter 11)

Additional Readings

Angrist, Joshua D., and Jõrn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. (Chapter 2)

Morgan, Stephen L, and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Second Edition. New York: Cambridge University Press. (Chapters 1 - 6)

Rosenbaum, Paul R. 1999. Choice as an Alternative to Control in Observational Studies (with discussion). *Statistical Science* 14.3: 259–304.

King, Gary, and Langche Zeng. 2006. "The Dangers of Extreme Counterfactuals." Political Analysis 14 (2): 131–59. https://doi.org/10.1093/pan/mpj004.

Westreich, Daniel, and Sander Greenland. 2013. The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients. *American Journal of Epidemiology* 177.4: 292–98. https://doi.org/10.1093/aje/kws412.

Sekhon, Jasjeet S. 2009. Opiates for the Matches: Matching Methods for Causal Inference. *Annual Review of Political Science* 12: 487-508.

Arel-Bundock, Vincent, Ryan C Briggs, Hristos Doucouliagos, Marco Mendoza Aviña, and TD Stanley. 2024. Quantitative Political Science Research Is Greatly Underpowered. *Journal of Politics*. https://doi.org/10.1086/734279.

Kam, Cindy D., and Marc J. Trussler. "At the nexus of observational and experimental research: Theory, specification, and analysis of experiments with heterogeneous treatment effects." Political Behavior 39, no. 4 (2017): 789-815.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. A general approach to causal mediation analysis. Psychological methods, 15(4), p.309.

11. April 7: Observational Data Approaches Based on Over-Time Patterns in the Outcome Variable

Week aim: Having multiple observations of "treated" and "untreated" units over time opens the door to additional model-based analytical methods. As with all research designs, to give a causal interpretation to resulting estimates, we have to believe that a set of assumptions hold. In recent years, a number of scholars have come to question whether the assumptions necessary for the workhorse difference-in-differences model are met in many of the scenarios where they have been applied. This week, you should become familiar with the underlying logic of difference-in-difference models and other closely related models.

Week exercise: Find a paper that uses difference-in-difference assumptions or a two-way fixed effects specification. Discuss in one page the comparisons that the authors are trying to make and whether or not they are credible comparisons for estimating a causal effect,

Required Readings

Angrist, Joshua D., and Jõrn-Steffen Pischke. 2015. Mastering 'Metrics: The Path from Cause to Effect. Princeton, NJ: Princeton University Press. (Chapter 5)

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. Comparative Politics and the Synthetic Control Method. American Journal of Political Science 59(2): 495–510. https://doi.org/10.1111/ajps.12116

Imai, Kosuke, In Song Kim, and Erik H. Wang. 2023. Matching Methods for Causal Inference with Time-Series Cross-Sectional Data. American Journal of Political Science 67 (3): 587–605. https://doi.org/10.1111/ajps.12685.

Chiu, Albert, Xingchen Lan, Ziyi Liu, and Yiqing Xu. 2025. Causal Panel Analysis under Parallel Trends: Lessons from a Large Reanalysis Study. American Political Science Review: 1–22. https://doi.org/10.1017/S0003055425000243

Additional Readings

Angrist, Joshua D., and Jõrn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press. (Chapter 5)

Chaudoin, Stephen, Jude Hays, and Raymond Hicks. 2018. Do We Really Know the WTO Cures Cancer? British Journal of Political Science 48.4 (2018): 903–928.

Imai, Kosuke, and In Song Kim. 2019. When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data? American Journal of Political Science 63 (2): 467–90. https://doi.org/10.1111/ajps.12417.

Card, David, and Alan B. Krueger. 1994. Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania. American Economic Review 84 (4): 772–93. https://doi.org/10.3386/w4509.

12. April 14: Cases and Samples

Week aim: How you select your cases, respondents, or subjects determines what you can learn. Selection on the dependent variable is deadly for inference. Convenience samples may not generalize. The way that regression coefficients reflect a specific–but typically hidden–weighing of the units in your sample means that you need to take care in generalizing from them. This week you should learn to think carefully about who or what is in your study and who or what is not—and what that means for your conclusions.

Week exercise: Find an article in your substantive area. Describe the selection and/or sampling process for cases. In one page, explain who or what is included, who or what is excluded, and what this means for internal validity and also external validity.

Required Readings

Geddes, Barbara. 1990. How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics. *Political Analysis* 2: 131-50.

Aronow, Peter M., and Cyrus Samii. 2016. Does Regression Produce Representative Estimates of Causal Effects? *American Journal of Political Science* 60 (1): 250–67. https://doi.org/10.1111/ajps.12185

Findley, Michael G, Kyosuke Kikuta, and Michael Denly. 2021. External Validity. *Annual Review of Political Science* 24:365–93.

Stantcheva, Stefanie. 2023. How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible. *Annual Review of Economics* 15: 205-34.

Westwood, Sean J. 2025. The Potential Existential Threat of Large Language Models to Online Survey Research. *Proceedings of the National Academy of Sciences* 122 (47): e2518075122. https://doi.org/10.1073/pnas.2518075122.

Additional Readings

Egami, Naoki, and Erin Hartman. 2023. "Elements of External Validity: Framework, Design, and Analysis." American Political Science Review 117 (3): 1070–88. https://doi.org/10.1017/S0003055422000880.

Slough, Tara, and Scott A. Tyson. 2023. "External Validity and Meta-Analysis." American Journal of Political Science 67 (2): 440–55. https://doi.org/10.1111/ajps.12742.

Groves, Robert M. and Emilia Peytcheva. 2008. The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly* 72.2: 167-189.

Coppock, Alexander and Donald P. Green 2013. Assessing the Correspondence between Experimental Results Obtained in the Lab and Field–A Review of Recent Social Science Research. *Political Science Research and Methods* 3.1: 113–131.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge University Press.

Dryzek, John S. 2025. For Example: How to Use Examples in Political Science. *American Political Science Review* 119(1): 449–461.

13. April 21: Content Analysis and Measurement

Week aim: Text, images, audio, and satellite data can all be evidence, but evidence of what? The computer does not solve your conceptualization problem—it just lets you make the same mistake at scale. This week you should learn to think about validation: how do you know your measures capture what you intend?

Week exercise: Find an article in your substantive area that uses some kind of content analysis (text, audio, video, etc). In one page, describe the way that the researcher worked to convince the reader about the construct validity and general reliability of the measurement strategy.

Required Readings

Grimmer, Justin, Margaret E. Roberts, and Brandon M. Stewart. 2022. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton, NJ: Princeton University Press. (Chapters 2, 25, 26 and 27)

Benoit, Kenneth, Scott De Marchi, Conor Laver, Michael Laver, and Jinshuai Ma. Forthcoming. Using Large Language Models to Analyze Political Texts Through Natural Language Understanding. *American Journal of Political Science*. https://kenbenoit.net/pdfs/Benoit_etal_2025_AJPS.pdf

Casas, Andreu and Nora Webb Williams. 2019. Images that Matter: Online Protests and the Mobilizing Role of Pictures. *Political Research Quarterly* 72.2: 360-375.

Dietrich, Bryce J., Matthew Hayes, and Diana O'Brien. 2019. Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech on Women. *American Political Science Review* 113.4: 941-962.

Livny, Avital. 2021. Can Religiosity be Sensed with Satellite Data? An Assessment of Luminosity during Ramadan in Turkey. *Public Opinion Quarterly* 85.S1: 371-98.

Additional Readings

Neuendorf, Kimberly A. 2002. *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications. (Chapters 1 and 7)

Young, Lori and Stuart Soroka. 2012. Affective News: The Automated Coding of Sentiment in Political Texts. *Political Communication* 29.2: 205-231.

Dietrich, Bryce J. and Melissa Sands. 2021. Seeing Racial Avoidance on City Streets. *Nature of Human Behavior*.

Ornstein, Joseph T., Elise N. Blasingame, and Jake S. Truscott. 2025. How to Train Your Stochastic Parrot: Large Language Models for Political Texts. *Political Science Research and Methods* 13(2): 264–81. https://doi.org/10.1017/psrm.2024.64

Bermejo, V.J., Gago, A., Gálvez, R.H. et al. 2025. LLMs outperform outsourced human coders on complex textual analysis. *Scientific Reports* 15: 40122.. https://doi.org/10.1038/s41598-025-23798-y

14. April 28: What are we doing when we do political science?

Week aim: How does political science make progress? This week you should step back from methods and ask big questions about knowledge accumulation, paradigms, and what counts as scientific advance. These debates shape how you position your own work within the discipline. After all, any one researcher can only do so much: we rely on a community of other researchers over decades and across the world in order to build understanding of politics.

Week exercise: Reflect on the semester. In one page and with reference to one or more theories of how science advances, describe the most important thing you learned about research design and how it will change what you do.

Required Readings

Godfrey-Smith, Peter. 2021. *Theory and Reality: An Introduction to the Philosophy of Science*. Second Edition. University of Chicago Press. (Chapters TBD)

Additional readings on knowledge accumulation to be determined (e.g., work by Tara Slough, Ana Wilke, the EGAP Metaketa teams)

Additional Readings

Popper, Karl R. 1979. Objective Knowledge: An Evolutionary Approach. Revised Edition. New York: Oxford University Press. (Chapter 1)

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. (Chapters 1 - 9)

Lakatos, Imre. 1969. "Criticism and the Methodology of Scientific Research Programmes." *Proceedings of the Aristotelian Society, New Series* 6:149–86.

Feyerabend, Paul. 1975. *Against Method: Outline of an Anarchistic Theory of Knowledge.* New Left Books.

Elman, Colin, and Miriam Fendius Elman. 2002. How Not to Be Lakatos Intolerant: Appraising Progress in IR Research. *International Studies Quarterly* 46.2: 231–62. http://www.jstor.org/stable/3096070

15. May 5: In-Class Final Test

Acknowledgments:

We have already acknowledged that this syllabus has emerged as a joint effort over the years between Bowers, Gaines, Winters, Wong and Kuklinski with special debt to Jim Kuklinski for creating it in the first place and to the many students who have offered feedback.

We used Claude AI to draft the AI Policy, the Week Aims and Week Exercises; used Perplexity to suggest more recent engagement with the prediction versus explanation debates articulated by the Breiman 2001 reading; and also used Perplexity's agent to help with formatting the document.