Research Training: Scope and Methods/Research Design in Political Science
POLS4011/POLS8058

Richard Frank
March 3, 2026

WEEK 2: THE ART OF THE POSSIBLE

PART 1: OVERVIEW

This week serves as a bridge between the introductory framework of Week 1 and the two-week block on concepts that begins next week. My central question this week is: *how do you move from a broad intellectual interest to a tractable, defensible research project?* This is especially pressing because the research design memo is due at the end of Week 3, meaning you need to begin narrowing your focus now.

Today's title is deliberately chosen. Research design is the art of the possible, not the pursuit of the ideal. Every project involves trade-offs between ambition and feasibility, between internal validity and external validity, between the question you want to answer and the question the available evidence will let you answer. The goal this week is to help you internalise this: a good research design is not one that eliminates all threats to inference but one that makes defensible choices given real-world constraints.

Plan for today

1. Overview and connecting last week to this week
2. Readings: formulating questions, hypotheses, and feasible designs
3. Group activity: stress-testing research questions
4. Workshop time: your design memo

Key themes for this week

- Research questions under data, ethical, and feasibility constraints
- Why many "interesting" questions should not become theses
- The difference between an important question and an answerable one
- How to deal with setbacks, blind alleys, and changing questions
- When are questions unanswerable with available evidence?

The differentiated expectation from Week 1 carries forward here. Honours students should be able to identify ***whether*** a question is answerable given available evidence and methods. MA/PhD students should be able to articulate ***why*** a question is or is not answerable and what trade-offs different formulations of the same underlying interest entail.

PART 2: READINGS

Required readings

1. Van Evera, Stephen (1997), *Guide to Methods for Students of Political Science*, Chapter 1 ("Hypotheses, Laws, and Theories").

2. Booth, Colomb, and Williams (2016), *The Craft of Research*, Chapters 1–2
3. Hyde, Susan (2007), "The Observer Effect in International Politics: Evidence from a Natural Experiment," *World Politics* 60(1): 37–63.

1. Van Evera (1997), Chapter 1

Van Evera's opening chapter provides a concise, polisci-specific treatment of what hypotheses are and what makes them useful. The central argument is that good research begins with well-formed hypotheses, and that hypotheses vary in their value depending on several testable properties. This is a practical complement to KKV's more abstract discussion of inference from Week 1: where KKV tell you *what* good research looks like in principle, Van Evera (1997) starts telling you *how* to formulate it.

What makes a hypothesis valuable?
Van Evera (1997) identifies several criteria. A hypothesis should have large explanatory power (i.e., it should account for important phenomena). It should be falsifiable in a meaningful sense (i.e., generate predictions that could be shown to be wrong). And it should prescribe policy or provide actionable insight. The key point is that not all hypotheses are created equal. A hypothesis that explains a trivial outcome or that cannot be distinguished from its competitors is not worth building a thesis around, no matter how technically well-specified it is.

The hierarchy of theories
Van Evera (1997) distinguishes between hypotheses, laws, and theories. A hypothesis is a conjectured relationship; a law is an empirical regularity that has survived repeated testing; a theory is an integrated explanation of why laws hold. Most students tend to stick around the hypothesis level. The important point is clarity about what level of claim you are making and what evidence would be needed to support it.

Connection to the "art of the possible."
Van Evera's (1997) criteria for hypothesis quality are mostly about feasibility. A hypothesis with no observable implications is untestable. A hypothesis whose implications overlap entirely with those of a rival hypothesis cannot be tested in a way that differentiates itself with its rival. These are constraints on what is possible. Put simply, can my hypothesis actually be tested with the data and methods available to me and will the test tell us anything important?

2. Booth, Colomb, and Williams (2016), Chapters 1–2

These chapters address the most practical problem you face at this stage of the semester: how do you move from a topic to a research question, and from a research question to a research problem worth investigating? Booth et al. (2016) provide a structured process that complements the more discipline-specific advice in Van Evera (1997) and the inferential logic of KKV (1994).

From topics to questions
The first move Booth et al. (2016) describe is narrowing a broad area of interest ("I'm interested in democratic backsliding") into a specific question ("Why do some democracies with strong judiciaries still see executive overreach?"). The key thing is that topics are not questions. A topic is a territory; a question is a specific location within that territory. Students often confuse the two, and this confusion leads to projects that are unfocused and unmanageable.

From questions to problems
The second move is explaining why anyone should care about the answer. Booth et al. (2016) argue that a research question becomes a research *problem* when you can articulate what is at stake: what we will not understand, what we will get wrong, or what we will fail to do if this question goes unanswered. This is about motivation and justification: the "so what?" question we will keep coming back to. It maps onto KKV's (1994) emphasis on questions that are important in the real world and contribute to scholarly literature, but Booth et al. (2016) provide more practical scaffolding for how to make that case.

Practical value for the research design memo
You could use Booth et al.'s (2016) framework directly in drafting your research design memos. The memo asks you to articulate a research question and explain its significance to the field. Booth et al. (2016) give you a concrete method for doing this: state your question, explain the cost of not answering it, and show how your answer will change what we know or what we do.

3. Hyde (2007)

Hyde starts with a seemingly intractable question: does international election monitoring actually reduce fraud? The naive approach (comparing fraud levels in monitored versus unmonitored elections) runs straight into selection bias, since monitors are not randomly assigned. Hyde (2007) finds a creative solution in Armenia's 2003 presidential election, where monitors were quasi-randomly assigned to polling stations within a single election, allowing her to compare vote shares across monitored and unmonitored stations while holding country- and election-level factors constant.

Why this article?
Hyde's article is a great case for the "art of the possible." The broad question (does monitoring reduce fraud?) is important but in its most general form is very difficult to answer. Hyde's (2007) contribution was not to answer the general question but to find a specific empirical setting where a clear answer was possible. The article demonstrates how a researcher moved from a broad, important question to a feasible design by exploiting a specific empirical opportunity. This is exactly the move you need to make for your own projects.

The natural experiment design
The key design feature in this article is the quasi-random assignment of monitors to polling stations. Because the OSCE observation mission did not have enough monitors to cover all voting stations, and because the assignment was largely logistical rather than strategic, Hyde argues the assignment approximates random allocation. This allows her to treat the presence of monitors as an as-if random treatment and compare vote outcomes across monitored and unmonitored stations. The result (monitored stations showed significantly lower vote shares for the incumbent) provides credible evidence that monitoring constrains fraud.

Important trade-offs
The article highlights the trade-off between internal and external validity. The natural experiment is clean within the Armenian context, but is Armenia in 2003 generalisable? Would monitoring have the same effect in a country with different institutional arrangements, a different type of authoritarian regime, or different incentives for fraud? This is the scope question that will become central in Weeks 5 and 6. It is also worth noting that Hyde's (2007) design was opportunistic in the best sense: she found a setting that happened to produce quasi-

random variation. You could think about whether similar opportunities exist for your own questions—what Thad Dunning (2012) later calls natural experiments in the social sciences.

Reading discussion questions

Honours students

1. Van Evera (1997) argues that good hypotheses must be falsifiable and generate observable implications. Take a hypothesis from your own project: what would you need to observe to conclude the hypothesis is wrong? If you cannot answer this question, what does that tell you about the hypothesis?
2. Booth et al. (2016) distinguish between a research topic and a research problem. In your own words, what is the difference? Can you articulate the "cost" of not answering your research question? That is, what will we get wrong or fail to understand without your project?
3. What made Hyde's (2007) research question answerable in the specific Armenian context? What features of that setting allowed her to make a causal claim that would not be possible with a cross-national comparison?

MA/PhD

4. Van Evera's (1997) criteria for hypothesis quality include explanatory power, falsifiability, and policy relevance. Are these criteria always compatible? Can you think of a hypothesis that scores highly on one criterion but poorly on another, and what does that tension mean for how you design your research?
5. Hyde's (2007) natural experiment has strong internal validity but uncertain external validity. If you were advising an honours or undergraduate student who wanted to build on Hyde's (2007) finding, what design would you recommend to assess whether the monitoring effect generalises beyond Armenia in 2003? What new threats to inference would that design introduce?
6. Booth et al. (2016) argue that the key part in formulating a research problem is articulating what is at stake if the question goes unanswered. How does this map onto KKV's (1994) notion of "importance" from Week 1? Are there cases where a question is important in Booth et al.'s (2016) sense but not in KKV's (1994), or vice versa?

Entire class

7. All three readings in their own ways address the gap between the question you want to answer and the question empirical evidence will let you answer. Gerring and Seawright emphasised the discovery phase; KKV emphasised inferential logic; Van Evera emphasises hypothesis quality; Booth et al. emphasise the practical process of narrowing; and Hyde demonstrates what a successful narrowing looks like in practice. Where in this sequence does your own project currently sit, and what is the next concrete step you need to take?
8. Hyde's (2007) design was opportunistic: she found a context where quasi-random variation happened to exist. Van Evera (1997) would say the value of her hypothesis depends partly on its explanatory power and generalisability. Is there a tension between designing research around empirical opportunities (as Hyde [2007] did) and designing research around theoretically important questions (as Van Evera [1997] suggests)? How should we navigate this trade-off?

PART 3: GROUP ACTIVITY

Stress-testing research questions

Please form some small groups (3–4 students) according to whether you are an honours, MA, or PhD student. Each student presents their current research question to the group in two minutes.

The group then has five minutes to ask the following questions:

4. Is this a **topic** or a **question**? If it is a topic, what specific question within it would you prioritise?
5. What would count as **evidence** for or against this claim? If you cannot tell what negative evidence looks like, the question may need reformulating.
6. What data or **evidence** would you need, and does it exist or could it be collected within the scope of an honours thesis or a HDR chapter?
7. What is the most obvious alternative **explanation** for whatever pattern you expect to find? Can your design distinguish between your explanation and this rival?
8. Using Booth et al.'s (2016) framing: what is the cost of **not answering** this question? What will we get wrong or fail to understand?

After the small group discussions, please share a particularly interesting case where the diagnostic questions revealed a problem or led to a sharper formulation with the entire class. The point is not to embarrass anyone but to normalise the process of refining questions through critique. This is exactly what happens in academic workshops and seminars.

Remember the brief for the first assignment.

- A clearly articulated research question
- An explanation of its significance to the field
- Identification of the theoretical framework or approach
- A description of the proposed methodology including data sources, analytical strategies, competing designs, and scope conditions
- Identification of key risks
- Demonstration of command of alternative methodological approaches

Each student should write a single paragraph (no more than 100 words) stating their research question and why it matters. Students should then read each other's paragraphs and answer two questions: (1) Can I identify exactly what this person is studying? (2) Can I explain to someone else why it matters? If either answer is no, discuss what is missing.

A final note on scope and ambition

In my experience, students often err in one of two directions: questions that are far too broad to be answerable in a thesis ("Why do democracies fail?") or questions that are so narrow they have no broader significance ("Why did one council vote in Yass, NSW in 1985 go a particular way?"). The readings this week give you tools for finding the middle ground.