How to Read Social Science Research

PSCI 2220: Crisis Diplomacy
Professor Brenton Kenkel
Spring 2016

Introduction

Social science papers can be long, detailed, and jargon-filled, making it easy to lose sight of the main points the author is trying to make. In my experience, the best path to making sense of a piece of social science research is to ask these four questions while (and after) reading it:[1]

1. What is the central question?
2. What is the central answer?
3. What are the competing explanations?
4. Why are they wrong?

Answering these four questions is a starting point—not the end goal—for evaluating a piece of social science research. Once you've figured out the answers to these four questions, you know the basic structure of the argument that the author is trying to make. At that point, you can start digging into the details of the paper (or book) to evaluate the argument in light of the evidence the author musters.

What is the central question?

It is often tempting to summarize a paper in terms of its topic. As in, "This is a paper about nuclear deterrence," or "The topic of this book is the causes of war." But a work of social science isn't a mere collection of facts. It's a contribution

[1] I learned this method of summarizing and evaluating social science research from Professor Hein Goemans at the University of Rochester, who in turn took it from his advisors at the University of Chicago.

to a debate about how some facet of social life works. To understand the point of the work, we must know the terms of that debate.

The *central question*, or research question, is the controversy that a social science paper seeks to address.

- It is a source of disagreement, or at least potential disagreement, among scholars. Matters of mere fact like "Did World War II end in 1945?" do not constitute central questions.[2]
- It identifies the crucial *dependent variable* and *independent variable(s)* of the study.
  - The dependent variable is the phenomenon that the researcher wishes to explain variation in.
  - An independent variable is a phenomenon that may affect the dependent variable. In other words, we suspect that change in an independent variable may cause a change in the dependent variable.

  For example, consider the research question "Does nuclear proliferation make world war more likely?" The dependent variable here is world war: we wish to explain why there are more, or fewer, world wars. The independent variable is nuclear proliferation: we want to know whether greater levels of nuclear armament are associated with more (or fewer) world wars occurring.
- It is broad in scope. Do not be fooled into thinking a paper addresses a narrow central question simply because it draws on a narrow pool of evidence. For example, the central question of an article about the U.S. role in the Camp David Accords might be "How does the bias of an international mediator affect the chance that peace talks succeed?"
- It can be answered objectively. "Was the Iraq War a bad idea?" is a matter of opinion, not a central question for social science. On the other hand, "Do foreign occupations increase the likelihood of terror attacks against U.S. citizens?" is a valid research question.

[2] A good rule of thumb: if it's a question you can find the answer to on Wikipedia, it's not a central question.

What is the central answer?

The *central answer*, naturally, is the author's answer to the central question. It is the position that the author wishes to defend, or the argument that the article makes.

There are two components to a central answer. First, there is a statement about the relationship between the independent and dependent variables. Let's return to the example of the central question "Does nuclear proliferation make world war more likely?" The central answer will take a position on the nuclear proliferation–world war relationship, e.g.:

- Proliferation makes world war more likely.
- Proliferation makes world war less likely.
- Proliferation has no effect on the likelihood of world war.
- Proliferation makes world war more likely under this condition, but decreases the chance of world war under that condition.

This is just the beginning. The second component of the central answer is the *mechanism*: the reason for the relationship that the article posits between the independent and dependent variables. The task of social science is not just to find correlations out in the world, but to explain why they exist. Again returning to our example, here are some examples of potential central answers with mechanisms:

- Proliferation makes world war less likely because the fear of nuclear reprisal deters states from starting wars.
- Proliferation makes world war more likely because states will start preventive wars to keep others from attaining nuclear weapons.

What are the competing explanations? And why are they wrong?

It is a cliché—but true—that people are complicated. For any interesting research question about social phenomena, there are many plausible answers.

For an argument in favor of a particular answer to be convincing, it should contend with these potential alternatives.

We have seen that a central answer has two main components. Accordingly, there are two main types of alternatives to a given central answer. The first is that the relationship between the independent and dependent variables are different than claimed. To take a simple example, in order to argue that nuclear proliferation decreases the risk of world war, a scholar must address the possibility that it actually increases the risk. There are also more complex types of alternative explanations:

- *Omitted variable:* A different independent variable is primarily responsible for the variation in the dependent variable.
- *Spurious relationship:* Some outside factor is responsible for variation in both the independent and the dependent variables.
- *Reverse causation:* The dependent variable causes variation in the independent variable.

Scholars typically employ empirical evidence to argue in favor of the relationship that they claim holds, and against these kinds of alternative explanations. Using either case studies or statistical analysis, the goal is to demonstrate that the independent variable has the claimed effect on the dependent variable, holding all else equal. Which kinds of research designs do that convincingly is a topic we will return to throughout the rest of the course.

The other major type of competing explanation that a work of social science must contend with is that the mechanism is different. In other words, that the same relationship holds for a different reason than what the paper argues for. It is possible to make the right prediction for the wrong reasons. But if we want to extend and build on our findings, with the ultimate goal of building coherent theories of the social world, we must understand the reasons behind the relationships we find.

Showing that a particular mechanism drives a relationship is even more difficult than showing that the relationship holds in the first place. There are a few common steps to arguing in favor of a mechanism. You can show that there is a logically coherent argument in favor of the proposed mechanism.[3] Con-

[3]Making a logically coherent argument is more difficult than you might think. Thomas Schelling, whose work we'll read later in the course, does this masterfully.

versely, to rule out a proposed alternative mechanism, you can show that the theoretical argument for it makes no sense. Or, taking a more empirical tack, you can trace the history of one or two historical cases to find direct evidence that the proposed mechanism was at play.