

Teaching the Scientific Method in the Active Learning Classroom

Leanne C. Powner, *University of Michigan*

Undergraduate classroom presentations of the scientific method in the social sciences usually involve a lecture on the nature of the scientific enterprise: the discovery of the “truth” (if such a thing even exists) about the relationship between some set of concepts, events, or phenomena by means of generating falsifiable hypotheses and testing them against data. The tendency of faculty, however understandable, to convert the presentation into an abstract discussion of the philosophy of science often obscures the basic issues for students. I present here a simple, active learning classroom exercise which sensitizes students to the difficulties of hypothesizing about and testing unobservable phenomena by appealing to students’ intuitive understanding of the nature of knowledge and reality. I first describe the purposes and objectives of the activity in more detail, then discuss implementing this specific activity. Suggestions for activity adaptation and development by other instructors follow, and a brief discussion of sources and resources concludes.¹

Purposes of the Activity

First, this activity introduces students to the application of the scientific method in the social sciences using an inductive, active learning approach guided by the students’ own intuitive understandings of the conduct of scientific inquiry and the development of knowledge. Second, it alerts students to problems unique to scientific research in the social sciences, particularly the com-

lications posed by the need to find indicators for unobservable concepts. Third, it introduces basic notions of parsimony and falsifiability in a discussion of the theory development—hypothesis generation—hypothesis testing cycle and provides an opportunity for discussing inference in the social sciences. Finally, it allows instructors to introduce the distinction between positive and normative social science. The activity also provides broader benefits to students in helping them develop important critical thinking skills, specifically analysis and inference. The critical thinking skills combine with the conceptual vocabulary to allow students to be more sophisticated readers of academic literature.

Structure of the Activity

The activity begins with small groups of students (4–6 per group) selecting a Group Recorder and obtaining the question handout from the instructor. The instructor presents each group with a small, oddly-shaped object concealed in a fabric bag and asks the groups to theorize about their object *without touching it*, based only on the general characteristics it presents as it sits on their desks. Then, *without opening the bag*, groups try to identify, describe, and characterize their objects as thoroughly as possible by handling and manipulating the bag to collect data on their object. The Recorder notes the suggestions, evidence, arguments, and conclusions produced by the group’s discussion but does not participate in exploration or discovery.² When the groups have completed their investigations, the instructor asks them to present their findings, and through directed questioning, elicits specific information about the process of reaching those conclusions. What data did the group obtain from its object? What alternative hypotheses did the group adopt, test, and reject, and on what grounds were these rejected? Different questions are appropriate depending on the object and the instructor’s goals for the activity. The instructor next reveals the object and discusses with the class the accuracy of the group’s arguments and claims and any specific challenges that object

presents, and then relates the object to an appropriate concept (see below). With careful planning and question selection, students will discover the major concepts for themselves, even without substantial guidance from the instructor.

While seemingly simple, this activity has a substantial influence on student understanding of the challenges social scientists face in their work. The bagged objects represent the kinds of unobservable concepts social scientists use in their arguments. Major concepts in all subfields, such as power, globalization, anarchy, culture, and federalism, must be measured indirectly by their observable implications, and this introduces a notable amount of measurement error. Combined with our frequent inability to conduct controlled experiments to demonstrate causality, measurement error from indirect measurement often results in skepticism about causal claims since the causal mechanism—and indeed, the underlying concept itself—is only inferred rather than demonstrated directly. When students are exposed to the difficulties social scientists experience in measuring phenomena and making causal claims, they become more skeptical and critical consumers of the academic literature. By giving them the vocabulary and conceptual toolbox to discuss and critique theories, research design, and measurement, we provide them with an opportunity to engage with class material in ways that will help to develop their critical thinking skills more broadly.

Major Concepts of the Activity

This activity conveys concepts at two different levels: the individual bagged objects provide access to specific concepts, and the activity as a whole presents several broader elements. The largest concept introduced by this activity is the notion of measuring and characterizing unobservable phenomena. Much of social science revolves around concepts with no directly measurable characteristics, so we must look for observable implications that suggest our concept is causing or producing or otherwise involved in producing the effect we

Leanne C. Powner is a Ph.D. candidate at the University of Michigan. An earlier version of this essay was presented as a poster at the 2004 Annual Meeting of the American Political Science Association, Chicago, IL, and as part of a presentation at the 2006 American Political Science Association Teaching and Learning Conference, Washington, D.C. The author would like to thank poster session visitors, participants in the “Teaching Research Methods” track, and instructors at the University of Michigan for helpful comments on this activity. Comments are welcome and should be directed to LPowner@umich.edu.

are studying. The best analogy for students is shaking a wrapped gift to hear what sound it makes. Every person has tried this at some point, and most of us have guessed wrong at least once. We attributed the rattling sounds to marbles in the Chinese checkers game we really wanted, when the noise was actually from a jigsaw puzzle or a tin of peanuts. This concept of measuring unobservables pervades the activity and may be discussed at any point.

Instructors can also use this activity to introduce the notions of data, variables, and causality. Data is simply information—more precisely, systematically collected information about specific characteristics. The instructor might ask students what kinds of data they obtained about their objects and generalize the responses into variables of size, texture, hardness, weight, number and location of points or bumps, etc. Students obtain values of these variables, and link those values into a conclusion of the object's identity, nature, etc., based on their memories of other objects with similar characteristics. For some objects, students may even be able to suggest the characteristic(s) of the object that prompted them to think about the response they proposed. This concept is best conveyed during the discussion portion of the activity, before discussing theory generation and hypothesis testing.

The process of theory refinement is a continuous circle of theory generation, hypothesis generation, data collection, and hypothesis testing; the results of hypothesis tests are then used to refine the old theory or generate a new one, new or supplementary data is collected (often with refined measurement), and the process begins again. Work on a research program proceeds in this manner, conducted by a scholarly community. Old or unsupported hypotheses are rarely discarded, though; proponents splinter off and conduct a competing research program until one set of hypotheses is firmly rejected or another set gains the acceptance of the entire community. Bueno de Mesquita (2006) refers to this process of retaining old theories until they are decisively condemned as the “first theory of wing walking.” In an analogy to aerial stuntmen who walked on bi-winged planes, he argued that a field does not let go of a theory until another stronger or more encompassing alternative appears to replace it. During the discussion of a group's findings, instructors might query Recorders about alternative hypotheses the group proposed and ultimately rejected in favor of the final conclusion, and about the evidence which caused them to reject a hypothesis.

Objects and Related Concepts

This list includes elements the author and others have used in previous presentations of this lesson. It is not an exhaustive list; it is intended to provide other presenters with suggestions for translating objects into concepts, and where possible, connections to substantive questions asked in various subfields of the discipline. A detailed lesson plan is available at www-personal.umich.edu/~lpowner.

Unexpected Variants of Known Items

Finding something similar to a recognized item (but not exactly so) complicates the process of identifying and classifying events and behaviors. We recognize the general type, but the new items serve different purposes, or are of different types, or are things with recognized functions but which come in a different shape or packaging. Most of the data align into a pattern the researcher recognizes, but either some data appear anomalous when compared against the researcher's mental picture, or critical data are unobtainable through indirect measurement. Items in this category have included a dog chew constructed from a tennis ball with knotted rope through it, a comical and cartoonish looking stuffed cow, and a twisty-straw cup shaped like an ice cream cone. The ice cream cone cup in particular allows discussion of how even common phenomena are obscured to the point of possible misclassification when they cannot be observed directly. Students often miss its largest feature, the ice cream cone shape, though they identify all other important functions and components. Studying only data on political parties, scholars might doubt that Tanzania is a democracy since it only has had one party for most of its history, and so might place it in a category with other one-party states like the Soviet Union or the People's Republic of China. Democracy has other unobservable components, though, like the rule of law and observance of human rights, and Tanzania frequently scores reasonably well on these. This often prevents it from being placed into the non-democracy category, where its most salient features (freedom of expression, etc.) are essentially ignored.

Unknown Items

Scholars occasionally encounter phenomena we do not recognize. We can compare and relate these things to phenomena we do recognize which share some characteristics, but ultimately this

new thing will need to be classified and measured on its own. Such classification is contentious. Whoever authoritatively classifies that thing first usually wins acceptance of his or her definition of that phenomenon—revision is very difficult once the discipline begins activity on a topic because others will have used the initial definition to produce either supporting research or rebuttals. In the field of international relations, the Singer-Small list of wars has become standard, though most of the field acknowledges the list's inadequacy and incompleteness (especially in conflicts outside Europe). In particular, the list does not include wars between states and non-states, several of which occur during their period of study. By this criterion, the recent conflict between the United States and the Taliban and various United States attacks on al-Qaida do not constitute Singer-Small wars, since the Taliban and al-Qaida are non-state entities.³

Uncertainty and Contestability

Many phenomena have multiple potentially correct explanations, among which we can never distinguish with complete certainty. We must infer an actor's motivations or expectations from other behaviors which could be consistent with alternative theories or explanations. Some explanations are more probable than others, but even then any assertion of causality is contestable. In this activity, this concept usually appears when we discuss what the manufacturer labeled a “cell phone holder,” but which other students have identified as a beanie baby armchair, a remote control holder, and a paperweight. Was the Asian financial crisis a result of domestic economic mismanagement, herd behavior in international capital markets, a deliberate effort by individual investors (i.e., George Soros) to reshape the market, etc.? Did George W. Bush attack Iraq over oil prices, over real concern about weapons of mass destruction, over support for terrorism and ties to al-Qaida, as retaliation for the attempt on his father's life, or for sheer domestic political expediency to help him win in November 2004? How do we know which is “right”? How do we devise tests to discriminate between them? The answer lies in identifying observable outcomes that would occur under one theory and not under the others. With some prompting, students are often able to identify these types of critical tests.

The Importance of Assumptions

Most groups begin their investigation as scholars do, with the assumption that

they are, indeed, looking at *something*. Even if we know something exists or is happening, though, some events, processes or activities leave few observable traces. Unless the investigator looks very carefully, using a very specific set of measures or methods, he or she might miss the quarry altogether. Concepts like “trust,” “social capital,” or “identity” are highly attractive as explanations, yet they are devilishly difficult to measure convincingly or even to measure at all in a way that indicates the concept is truly present. The importance of assumptions is usually modeled by placing a scrap of fabric or cloth ribbon in one of the bags. Most student groups are convinced the bag is empty. The few who make a systematic investigation by flattening the bag, or who use an innovative investigative technique like holding the bag up to the light, are able to discern that something of a specific shape is in the bag; they can then infer other characteristics of the object *from* their inability to detect it through any other means. Be prepared, however, for the chance that students *do* discover the ribbon; the instructor might then turn the discussion into the importance of systematic activity in science.

Falsifiability

Falsifiability refers to our ability to say what evidence would demonstrate that a theory or hypothesis is incorrect. One bag usually contains a common 35mm film canister stuffed with cotton balls. Students have no difficulty identifying the film canister, but the cotton balls produce no observable implications: Students assume the canister is empty. With prompting, students will usually say that it could not contain film, because film would make noise if the canister were shaken. The absence of noise, then, would falsify a hypothesis that the canister contained film.⁴ Many constructivist hypotheses about international relations are notably difficult to falsify; the argument that an actor’s preferences changed (because of persuasion, social pressure, role redefinition, etc.) is often seen as *post hoc* and applicable to any case. Constructing falsi-

fiable hypotheses is a major challenge for those in the constructivist research program who wish to contribute to (or refute) the empirical literature.

Parsimony

As a characteristic of a good theory, parsimony means both explaining as much as we can with as little information as possible, and/or explaining a much broader scope of phenomena with a small amount of additional information. This is valuable in the social sciences because reliable data are scarce; most data are often blurry or imprecise as a result of weak measurement of unobservables. The ability to do more with less increases the power of the theory because it decreases the amount of error fed into the analysis by the weak measurement. This topic is often prompted by a doll-sized tea set. Students usually explore the cups and saucers first, but are unable to name them precisely until they identify the teapot; they will often admit that finding the teapot was all that really mattered in their characterization of their object and that the cups and saucers only served more as confirmation. How much do we need to see, and of what, before we can declare that globalization is indeed occurring? What are the key features that allow us to determine whether a political system is federal, and is South Africa federal by that definition?

Sources and Resources

By far the best sources for unique, difficult to identify objects are “dollar” stores, found in most cities and towns under a variety of names. Household decorations, seasonal items, and kitchen gadgets tend to work well. Hardware stores also provide a wide selection of strangely shaped items, though often the instructor is also pressed to name the item and/or describe its intended use or function. Consider using only one part of a two-piece object (i.e., just the straw and lid from a sports bottle), or attaching two related items to form a less-easily-

recognizable object (i.e., two batteries taped end-to-end).

Fabric drawstring bags are a very simple sewing project for most experienced sewing machine users. Dark cotton fabric is recommended to minimize the possibility of students seeing labels, etc., through the bag. Finished bags should be about 8 inches by 10 inches to accommodate a wide range of objects; cut fabric panels to about 9 inches by 12 inches for seam allowance and drawstring casing. For the less craft-inclined, Oriental Trading Company offers tie-dyed drawstring bags at a reasonable price; consult www.orientaltrading.com under Party Supplies. In this author’s experience, most drawstring plastic bags distributed by stores are both too thin and too pale-colored to make the object sufficiently unobservable; one possible exception is bags from Old Navy or The Gap, which are normally dark blue though about twice the recommended dimensions.

A short introduction to the scientific method in the social sciences is in Bruce Bueno de Mesquita’s textbook, *Principles of International Politics* (2006, 3rd ed.). A number of introductory methods books cover these concepts, but this author is unaware of any American or Comparative Politics texts which explicitly address the scientific method.

Conclusion

Students generally possess an intuitive understanding of how to investigate and evaluate arguments. This activity allows instructors to draw on that understanding to provoke student discovery of major challenges faced by scientific investigation in the social sciences. Careful selection of objects and design of questions provide opportunities to introduce a wide range of major concepts in social science research design and theory generation. Student participation in the learning and discovery process appears to produce both greater recall of content and greater confidence in student application of the concepts in later class activities.

Notes

1. A detailed lesson plan and a sample student reproducible page are available on the author’s web site, www-personal.umich.edu/~lpowner. Typical Recorder questions are in the Appendix.

2. An important part of the activity relies on the Recorder capturing the interim stages of group exploration and discovery. Experience suggests that when the Recorder participates in discovery, the group’s notes include only the

conclusions and omit the critical intermediate steps.

3. Several reviewers have asked for examples of objects which illustrate this concept. While I have several objects that I use for this, I am unable to describe them succinctly here as I really do not know their identities or intended uses. Curious readers may email the author for digital images of these items at LPowner@umich.edu.

4. I normally ask then if they would believe a claim that the canister contains cotton; they often say “yes” because I prepared the bags. I respond by asking if they would believe George W. Bush’s claims that he did not invade Iraq for the domestic benefits or the oil—which, after all, is what he says. I then lead the discussion around to primary sources, motivations, and incentives to dissemble.

References

Bueno de Mesquita, Bruce. 2006. *Principles of International Politics*. 3rd edition. Washington, D.C.: CQ Press.

Appendix: Sample Recorder Questions

To the Recorder: Please take as many notes as possible as your group members investigate the object in the bag. Note *any and all* suggestions made by group members, even ones the group ultimately rejects. Pay attention to the process by which your group reaches its conclusions as well as to the conclusions themselves.

1. Before the group begins examining the object, what is their theory about the nature of the item in the bag?
2. What hypotheses did your group suggest? How were competing hypotheses compared and rejected? Did your group develop intermediate hypotheses that were generally accepted as fact ("it's round," "it's squishy," etc.)?
3. What assumptions did your group make? How were they made—through a verbal process based on evidence or something more spontaneous? Did the group even notice the assumptions?
4. As an outside observer, did you agree with the assumptions that the group made? Did they seem reasonable given the evidence? Could you think of counter hypotheses? Were your ideas more or less accurate than the group's after the identity of the item was revealed?