

Assignment—Research Article (in Progress)

Your writing assignment for the quarter is to prepare a ten-page research article (in progress) that you would complete as a thirty- to forty-page manuscript for an article if you had time to undertake the empirical research. In its format your article (in progress) should include the following parts (more or less in this order):

1. A Statement of Your Research Question.
2. Examination of Alternative Answers to this Question Found in the Literature.
3. Your Answer or Thesis.
4. A Description of the Research Design That Will Adjudicate Among These Competing Answers.

1. **Research Question.** At the opening of your article you should define the research question that your article will answer. You will also explain to the reader why this question is important. Perhaps the hardest part of any project is formulating and defending a successful research question.

Your interest in a question may begin with the literature of the discipline or with the headlines of the newspaper. In either case, it is often fruitful as a first step to identify some pattern of variation in the real world that needs to be explained. Since you must convince the reader (and editor and potential employer) to care about your question, you may want to begin with some real world problem about which people actually do care. At the same time, it should be a question that political science can help us answer. So, ideally, your research question addresses both some real phenomenon that educated people outside your narrow specialty consider important *and* some puzzle that is significant in the context of the theories developed by specialists in the discipline of political science.

Be wary of the obscure topic that tempts you to make it your own private preserve. It is likely to become (and remain) this. Instead, link your burning interest in a particular case or topic to issues of concern to a larger audience.

An example: If your passionate interest is the “underdeveloped” nationalism of the Ruvakian people, you might broaden this to the question of conditions that tend to retard the development of national identities—the “dogs that don’t bark” around the world.

Another example: You note that the literature on bargaining to resolve secessionist conflicts has “overlooked” the importance of a variable that is mentioned in the general literature on bargaining—the level of alcohol consumption by each side. An article might define its objective as taking the next half-step in the literature on secessionist bargaining by testing hypotheses about the effects of relative alcohol consumption on the outcomes of bargaining. Before jumping into this topic suggested by the literature, however, ask whether you can convince a reader (and editor) that this is an important scholarly development or real-world problem.

Normally the research question poses some puzzle that concerns *causes* or *consequences*. That is, your question does not simply ask “what happened?” but it asks either “why?” or “so what?”

An example: If you are interested in political institutions and ethnic mobilization, you might ask how political institutions *influence* the effectiveness of various mobilizational strategies of ethnic groups. Or you might ask how the mobilizational strategies of ethnic groups *affect* governmental stability in parliamentary systems. The first question concerns causes (ethnic mobilization is the *dependent* variable); the second concerns consequences (ethnic mobilization is the *independent* variable).

Of course, most questions that are worth asking will be more complex than any of the previous examples.

A caveat: Although an individual, non-recurring event may be the inspiration for the puzzle that interests you, the research question normally should not be a question about a unique or idiosyncratic

event—unless the event is generally recognized to merit attention in and of itself. Normally your research question should address some *recurring, patterned* set of events.

An example: You are interested in the fact that most members of the Chumurt minority of Upper Udvashia entered politics through class rather than ethnic organizations. Rather than ask “why has the Chumurt minority mobilized along class rather than ethnic lines?” (a question about a single outcome with no variation and no obvious connection to other cases), you might turn this into an explicitly comparative question about a pattern of variation found in many countries.

One possibility is a question about variation at the individual level, “what distinguishes the Chumurts who have joined class-based organizations from the Chumurts who have joined ethnic organizations?” And defend this as a case study of a pattern of variation that occurs in many multi-ethnic societies such as Bolivia, Guatemala, and India—not just Upper Udvashia.

Another possibility is a question about variation across time: “why did Chumurts mobilize on class lines until the late 1960s, but increasingly on ethnic lines after 1970?” And defend this as a comparison that could have been applied to changing patterns of ethnic mobilization by ethnic minorities in many other countries.

Still another possibility concerns cross-national variation: “why does the Chumurt minority in Upper Udvashia mobilize on class lines, but the Chumurt minority in Lower Udvashia mobilize on ethnic lines?” And again, explain that this comparison will shed light on differences that are more universal than the differences between Upper and Lower Udvashia.

Another caveat: Political science cannot study the future. If you want to study the consequences of some current or proposed policy, you cannot study a case that has not yet reached some conclusion.

An example: You are interested in the question whether “soft partition” in Iraq would be a better solution to the communal conflict between Shiite and Sunni communities and to the secessionism of the Kurds than attempting to maintain a unitary state. That, however, is a question about the consequences of actions that have not even been undertaken. Even if soft partition had been introduced recently it is probably too early to attempt to analyze the consequences.

One possibility for reformulation of this question is to ask what lessons we might draw for Iraq from previous experiments with territorial decentralization following civil wars. Then identify all the most-nearly-comparable cases for which we have good evidence to track the consequences of territorial decentralization versus unitary government.

To summarize: The first, and perhaps most important, step in any project is to identify precisely the real-world differences that you want to explain. This often means stating clearly at the opening, “My dependent variable is. . . .”

2. ***How Others Have Answered This Question.*** Sometimes called the “literature review,” this section summarizes the major answers to your research question that other investigators have already offered and explains why some or all are inadequate.

Unlike the dreadfully dull literature reviews that one sometimes writes for seminars, this should not include every work written since humans first painted on cave walls that is distantly related to your topic. It should stress those works that present different answers to your research question.

Unlike the dreadful literature reviews this should not simply present a series of summaries of relevant books and articles. It should organize the literature (grouping works where appropriate) into alternative answers to the research question that your article poses.

Your objective in reviewing the literature review is to identify alternative hypotheses. Alternative theoretical approaches to your research question are usually rooted in some distinctive assumptions about the variables that deserve close attention. Theories present a logic or chain of reasoning that links larger concepts to the more specific question that you have posed. The last step in this line of reasoning is a hypothesis that links cause and effect. Your literature review should make clear the assumptions, reasoning, and hypotheses in each major theory. The literature review in a thirty-page article manuscript inevitably must be brief, since the bulk of the article (perhaps twenty of the thirty pages) will present your research findings.

For example, your research question is the following: Under what conditions do ethnic minorities press secessionist claims against their governmental leaders? You might discover two or three broad theoretical approaches in the literature (some of these approaches may be represented by the works of more than one author). One broad theoretical approach might be structural political economy that attributes the rise of secessionism to the economic exploitation of a culturally distinct periphery by the economic core of a country. This theoretical approach is in tension with political-institutionalist theories that attribute the rise of secessionism to the absence of constitutional avenues for ethnic minorities to press their demands on the government.

These and other approaches lead to competing hypotheses that specify different causal factors that explain variation in the levels of secessionism (among different ethnic groups in different countries at different times). The logical reasoning of structural political economy might lead you to the hypothesis: Where the economy of the culturally-distinct periphery produces raw materials and the economy of the core produces industrial goods, the cultural community in the periphery is likely to develop a national identity and press secessionist claims on the center.

Note the form of this “internal-colonialism” hypothesis: It links a dependent variable (secessionism) with an independent variable (a specific disparity in the economies of center and periphery). The *independent* variable is the cause or *explanatory* variable and, within the context of the hypothesis is *predetermined* (that is, its causes are not specified). The *dependent* variable is the effect that is determined by a cause specified within the hypothesis.

Also note that both dependent variable and independent variables in this dependency hypothesis actually *vary*—that is, each can take on at least two values (e.g., high vs. low, more vs. less, present vs. absent). A so-called “variable” that always assumes one value (that is always present, always low, etc.) is a constant and cannot be included in a causal hypothesis.

After you have outlined the alternative hypotheses, you should explain why the hypotheses advanced by others are inadequate. This critique justifies the theoretical contribution you will make: You are remedying a deficiency in the existing literature.

3. ***Your Answer or Thesis.*** Either you select one of the theories and hypotheses in the literature as the best prospect or you develop an alternative answer that has not yet been offered in the literature. Actually, until you complete your empirical research, this is only your working proposition. If your research substantiates this working proposition, it becomes the thesis of your article.

In developing your working proposition (thesis) you should do the following: [1] You should root this in some larger theoretical tradition. [2] You must then present the reasoning that links this tradition to specific expectations concerning the research question with which you began. [3] You should formulate your hypotheses. [4] You should compare your proposition with the other theories identified in your literature review. In the balance among these parts of your answer, in an article-length manuscript the first normally can get only brief attention, while the next three need the fullest development.

For example, your proposition (thesis) is the following: Ethnic movements are more likely to assume the form of political parties under proportional representation, but assume the form of interest groups under plurality voting. (Independent variable = voting system [PR vs. plurality]; Dependent variable = organizational form of ethnic movements [party vs. interest group])

You should explain to the reader that this working proposition is an application to a specific situation of a larger analytic tradition (e.g., political institutionalism). You should briefly identify the assumptions in this tradition but explain more fully the chain of reasoning that leads to your working proposition.

4. ***Research Design.*** The purpose of the research design is to formulate a fair test of the alternative hypotheses (including your own working proposition) and to identify appropriate empirical evidence for this. The test must be fair in that it cannot be biased in a way that favors your hypotheses over the alternatives. A research design normally specifies the following:

- a. **Operationalization of Concepts.** This is the process by which you identify empirical referents that measure the abstract concepts (the “true” variables) specified in your hypotheses. Abstract concepts seldom can be measured directly; operationalizations by definition always can be.

Please note that the word “measure” does not mean your operationalization must be a *continuous* variable like gross domestic product. It may be an *ordinal* measure of rank or even a *dichotomous* variable with only two values. (The most common dichotomous variables are “present vs. absent” and “more vs. less.”) Operationalization may require that you create new measures not found in existing compendia of statistics or that you use *multiple indicators* to measure a single concept.

The ideal *operational* variable is perfectly correlated with the *true* variable (or concept) and can be measured by other researchers at a later date. That is, your operationalizations should be *valid* (measure the true variable as closely as possible), *reliable* (yield consistent values across a series of observations), and *replicable* (yield the same values for other researchers).

For example, in the example of the internal colonialism hypothesis used above, how would you measure the disparities in economies of core and periphery? A possibility might be proportion of workforce employed in industry, but you should recognize that this is only a surrogate (and possibly a flawed indicator) for the concept in your hypothesis.

That is, to measure a “true” variable (such as economic disparity in the internal colonialism hypothesis) you must develop a coding scheme or an index that permits you to classify cases with as little ambiguity as possible. In an individual case the most accurate (valid) measure might well be your “gut instinct,” but this subjective measure is unacceptable because it is unlikely to yield as precise a measure for all cases (reliability) and it certainly cannot be replicated by other researchers (replicability).

When developing variables and operationalizations, a useful “reality check” is the mental exercise of actually using these to code some real cases. If all your cases fall into one category or many cases simply cannot be coded unambiguously, you possibly have a bogus variable or flawed operationalization.

- b. **Case Selection.** Your choice of cases should be dictated by the research question rather than vice versa. (Of course, if your research question is derived from an empirical puzzle, your selection of cases may be constrained or even predetermined.)

The *number* of cases you study will depend on the research costs of each additional case: Where you need little information about each case and that information is easily collected, you might include every case in the population. Where you need more information or that information is harder to collect, you will examine fewer cases—perhaps as few as two. (And the cases may all come from one country.)

Where you must select a sample of cases from the larger population, the *appropriateness* of cases will be determined by the theories you test. When selecting cases, keep in mind three considerations:

- [1] Try to select cases that reflect variation on your independent variables.

For example, to test the hypothesis that socioeconomic development determines the level of secessionism, you should try to select at least one county at a low level of socioeconomic development and another at a high level of development. (If your proposition is that the relationship is curvilinear, with secessionism reaching a peak at intermediate levels of development, then you would need a minimum of three cases—one each at a low level, intermediate level, and high level of socioeconomic development.)

To test the competing hypotheses that socioeconomic development (high versus low) and regime-type (democracy versus authoritarianism) determine the levels of secessionism, you might use four cases (or as few as two cases if you select appropriately). In such tests of competing hypotheses it is often useful to go through an exercise like the following: Treating the two independent variables as dichotomous (high vs. low development, democracy versus authoritarianism), there are four possible combinations (high socioeconomic development plus democracy; high development plus authoritarianism; etc.). You could represent this in a simple two-by-two matrix. You may want to

have one case for each of the four combinations. Yet, if each new case represents very costly research, you may possibly trim your cases to two: In two of these combinations the hypotheses are likely to predict similar outcomes—for example, both low socioeconomic development and authoritarianism might predict high levels of secessionism, similarly both high development and democracy might predict low levels of secessionism. The cases you select for your study should be those in which your competing hypotheses predict different outcomes—that is, one case involving high development and authoritarianism and a second case involving low development and democracy.

- [2] Hold constant excluded factors that might contaminate your results. The cases you select may differ from one another on many other dimensions apart from the independent variables that interest you and these “excluded” variables (the variables not taken into account in your analysis) could be the real reason your cases vary on the dependent variable. You should select cases so as to “hold constant” or “control for” the most obvious alternative causes.

An example: In a study of the treatment of ethnic minorities by governments you select an Arab state as a country with high socioeconomic development and a Latin American state as a country with low socioeconomic development. You find that, as expected, the regime in the country with high socio-economic development extended more political rights to its minorities. Your hypothesis is confirmed!

Reviewers of your manuscript might object that you have failed to hold constant some important effects. First, the less developed nationalism of the majority in the Arab case made the government less likely to see minorities as threatening to the nation and so the Arab government was more likely to tolerate ethnic minorities. The reason was not differences in level of development, but differences in the strength of majority nationalism. Second, minorities constitute a much larger share of the population of your Latin American case and so the majority tends to see these as a threat to their predominance and finds it harder to extend them rights. The reason was not level of development but the size of the minority population. To “control for” these other effects, you should have selected two countries that are alike on all dimensions except the one that interests you.

Note that you can sometimes achieve the same type of control by examining the same country at two different points in time, if the country has changed relative to the explanatory variable that interests you. You may also achieve the same level of “control” over other factors by comparing cross-sectional cases (e.g., provinces, economic sectors) within the same country in the same period of time.

- [3] Do not select cases in a manner that precludes variation in the dependent variable.

The most common example of this is the study that attempts to identify the causes of some phenomenon and then examines only instances where that phenomenon has occurred. For example, a flawed study of the causes of secessionism might select only countries that have experienced secessionist conflict. This study selects its cases in a manner that is correlated with the dependent variable and includes no cases “where the dog didn’t bark.”

You may select cases that you know to differ on the dependent variable such as one country with high levels of secessionism and one with low levels of secessionism. This will permit you to discover whether the pattern of outcomes (the dependent variable) correlates with your independent variable (such as level of socio-economic development). A well-designed research project that selects for variation on the *independent* variables does not guarantee variation in the *dependent* variable, but does select cases with the *a priori* expectation of finding such variation.

Perhaps even more important is a closely related admonition: Beware of selecting only cases that you know conform to the pattern you hypothesize in your working proposition.

- c. **Types of Evidence.** You should explain and defend your decision to use specific types of evidence in measuring your operational variables and running your “tests.” The list of possible sources is extensive. It might include statistical sources, secondary sources such as monographic histories, memoirs, newspapers, government documents, public opinion surveys, and in-

interviews. The measure of good evidence is once again appropriateness. In selecting evidence you must also make a reality test when it comes to cost and availability. You must also be concerned about reliability and bias in sources.

In sum, your “research article (in progress)” for this seminar will include the part of a completed article manuscript that normally appears just before you dive into the details of your empirical evidence. Your manuscript will end with an explanation of your operationalizations, cases, and evidence; it will not go on to the actual research.