

Research Prospectus

Your major writing assignment for the quarter is to prepare a twelve-page research prospectus.

A prospectus describes a research project that you would complete if you had the time and resources; you will not actually complete the empirical research this quarter. There is no magical formula for a prospectus and no template is appropriate to all projects. For a "normal science" research project, the following format is often appropriate: The prospectus [1] begins with a research question, [2] proposes the author's answer and presents alternative answers that other analysts have suggested, and [3] describes a research strategy to test the tenability of the author's answer against these alternatives. Your prospectus for this course should follow this pattern and include the four following elements:

1. **Research Question.** Perhaps the hardest part is formulating a successful research question. It is usually best to begin with some real world problem about which people actually care. That is, you must justify the importance of the question you ask. It should convince the reader that she or he should care about the question.

► Ideally, your research question addresses 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 opposition to Transbarbaria's oppressive military regime, you might broaden this to the question of the conditions that tend to bring down military regimes.

► Although an individual, non-recurring event may be the inspiration for the puzzle that interests you, the research question should not be a question about a unique or idiosyncratic event. Your research question should address some *recurring, patterned* set of events.

For example, if you are interested in the contribution of the Great Depression to the rise of Hitler, rather than ask "did Hitler come to power because of the depression?" you might ask "are fascist movements more successful in states undergoing greater social dislocations from economic depression?" The first asks about a single outcome, the second asks about a pattern of outcomes.

Another example: You are interested in the fact that most members of the Chumurt minority 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) you might turn this into an explicitly comparative question about a pattern of variation: 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?" 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 since the 1970s?" Still another possibility concerns cross-national variation: "why does the Chumurt minority in Permyakistan mobilize on class lines, but the Udvash minority in Komistan mobilize on ethnic lines?"

► 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.

►To summarize: The first, and perhaps most important, step in a prospectus 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. . . .” This may represent a fundamental shift in the focus with which you have become familiar in preparing for comprehensive examinations. In the latter we often focus on the independent variables in order to group authors into schools of thought.

2. **Literature Review.** Your literature review summarizes the major answers to your research question that other investigators have already offered.

►The literature review should not include every work under the sun that is distantly related to your topic; it should stress those works that present (or imply) different answers to your research question.

►The literature review should not simply present a series of summaries of relevant books and articles, it should organize the literature (grouping works where appropriate) into alternative approaches and alternative answers to your research question. The object of the literature review is to identify alternative hypotheses. Alternative theoretical approaches to your research question are usually rooted in some distinctive assumptions about what variables 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.

For example: Your research question is: Under what conditions do military regimes in the Third World survive longer? 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 dependency theory that explains political outcomes in the Third World by the intervention of foreign economic hegemons (usually advanced industrial states) in the internal affairs of the Third World. This theoretical approach is in tension with theories that explain the stability of Third World regimes by their level of institutional development or by the country's level of socioeconomic development.

These and other approaches lead to different hypotheses that specify the causal factors that influence the longevity of military regimes. The logical reasoning of dependency theory might lead you to the conclusion that this theory sees the longevity of military regimes as principally (or solely) the consequence of the degree of support offered by the foreign hegemon; in this view neither institutional nor socioeconomic development has an independent effect on longevity. The dependency hypothesis is: Military-regime longevity increases as foreign support for the military regime increases.

Note the form of this dependency hypothesis: It links dependent variable (longevity) with an independent variable (foreign support). 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 *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 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. When you

are writing the prospectus (that is, prior to your empirical research) this is your working proposition. If your research substantiates this working proposition, it becomes the thesis of your manuscript.

► 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.

An example: Your proposition (thesis) is the following: military regimes survive longer when they institutionalize the mechanisms of accountability between the armed forces and those officers who take governmental posts.

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

In both examples, you must explain to the reader that this working proposition is an application to a specific situation of a larger analytic tradition (in the second example, political institutionalism). You should explain the assumptions in this tradition that begin the chain of reasoning leading to your working proposition and you should lay out this logical chain for the reader.

4. **Research Design.** The purpose of the research design is to formulate a fair test of the alternative hypotheses (including your own working propositions) 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 dependency theories used above, how would you measure degree of external support? A possibility might be volume of military aid given by the foreign hegemon, 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 external support) 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).

A useful "reality check" when developing variables and operationalizations 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 one or two.

For example, some cross-national studies of ethnic conflict have relied on quantitative measures that are easily computed from available statistical sources such as the United Nations annuals or computerized data banks. The cases in these studies can be every independent country in the world ($n > 150$) or even every country in every year between 1950 and 1990 (40 years times 150 or more countries > 6000). Alternatively, some studies—for example, a study of the influence of internal structure of ethnic groups on their strategies—require information that can be gathered only through painstaking, close analysis. The costs of gathering information on each additional case in such studies mandate that the investigator focus on only a few cases.

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] Select cases that reflect variation on your independent variables.

For example, to test the hypothesis that socioeconomic development determines the longevity of military regimes, you must 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 regime longevity 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 socio-economic development.)

To test the competing hypotheses that socioeconomic development and external support determine the longevity of military regimes, you might use four cases (or only 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), there are four possible combinations (high socioeconomic development plus high external support; high development plus low support; 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 high external support might predict high longevity, similarly both high development and low support might predict low longevity. 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 high support and a second case involving low development and low support.

- [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: You select an Arab state as a country with high socioeconomic development and a Latin American state as a country with low socioeconomic development in order to explain longevity of military regimes. You find that, as expected, the regime in the country with low socioeconomic development survived longer than the other. Your hypothesis is confirmed! A cultural theorist might object that you have failed to hold constant the important effects of political culture on the survival of military regimes: the real reason for the difference in regime longevity in the two cases is differences in the orientation toward authority in Arab and Latin American cultures. 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.)

[3] Do not select cases in a manner that is correlated with 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 ethnic conflict might select only countries that have experienced severe ethnic 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.”

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.

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 such variation.

- c. ***Types of Evidence.*** Your research design should identify the types of evidence you will need to collect in order to measure the operational variables and run 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 interviews. Very successful studies in political science have used each of these types of evidence; the test of good evidence is once again appropriateness. In selecting evidence you must make a reality test when it comes to cost and availability. You must also be concerned about reliability and bias in sources.

In sum, the fourth part of your prospectus (the research design) should explain each of these choices—operationalizations, cases, and evidence. It should justify each of these choices in order to convince the reader that you have designed a good test of the competing theories and hypotheses.