

Framing an Analytic Question

I. INTRODUCTION

The difference between analyzing a subject and merely writing about it is the heart of what distinguishes the research paper from journalistic writing. Research papers revolve around clear analytic topics. Good topics are narrow enough to be “doable” yet broad enough to avoid a “who cares?” reaction from readers. Good topics are ambitious enough to be challenging but not so overly ambitious that their scope may prevent realistic completion of the research project. Analytic questions are empirical puzzles that cannot be answered merely by describing the history and present circumstances of the object of investigation. Rather, qualitative or quantitative data about that object must be analyzed. If your paper can be completed merely by giving a history or description, you do not have an analytic topic.

What is analysis? In contrast to descriptive or speculative writing, research papers pose one or more hypotheses, assemble evidence, and come to conclusions. For dissertations, these conclusions should make an original contribution to one’s discipline. A good research topic can be said to be analytic when four conditions are met:

1. *A dependent variable is identified.* At least one thing must be being explained. Moreover, this one thing must vary so that we can investigate what other things (independent variables) vary with it.
2. *A plausible explanation is posed.* There must be at least one idea of how one or more independent variables relate to the dependent vari-

able. This is the hypothesis, or a set of hypotheses, constituting a theory.

3. *Counterhypotheses are examined.* Analysis can show some explanatory models that are consistent with the data, but the same data may be consistent with many alternative models. Analysis must include investigation of alternative hypotheses and theories.
4. *Operationalization of all variables makes analysis concrete and specific.* All the terms appearing in the hypotheses must have a meaning that is defined well enough to allow us to gather evidence.

Note that the analytic questions associated with one's thesis are *not* usually the same as questions asked of respondents in an opinion survey (Steeves et al., 1996). Rather, the analytic questions associated with the thesis are ones the researcher poses based on theories in the discipline or scientific constructs. For instance, in a study in which the central analytic question revolves around whether additional years of education lead to greater job satisfaction controlling for differences in salary level, one would *not* ask respondents if they felt more satisfied on their jobs as a result of their education, disregarding differences in salary level. Rather, one would have separate items about degree of job satisfaction, level of income, and years of education. The researcher would use statistical inference to assess the validity of the analytic question's central hypothesis.

Under some circumstances, the researcher may want the respondent in a survey to consider the analytic question itself. This may be useful, for instance, in understanding differences among subjects in their perceptions of key terms. Respondents in the foregoing example might be asked to define "job satisfaction," for instance. Such a question may reveal intersubjective differences of perception, but the researcher should be very aware that respondents' answers may reflect not only their own experience but also popular, cultural, political, social, religious, and other types of influences. Moreover, the researcher must consider whether the very act of asking direct questions about the analytic query may contaminate the results. For example, does forcing respondents to clarify in their minds what job satisfaction is lead to changes in their thinking about something that, in truth, was much more ambiguous or otherwise different in respondents' minds before they were interviewed?

II. FRAMING AN ANALYTIC QUESTION

A. Dimensions of Analysis

Framing a good analytic topic is the key to success in research writing. The analytic process revolves around four dimensions:

1. *Having a clear dependent variable.* Decide specifically what you wish to explain. This involves identifying one or more dependent variables whose

variance the researcher wishes to explain it may also be helpful to be explicit about related phenomena for which the researcher will not try to account. If the researcher is writing about the effectiveness of a crime control program, for example, he or she will need to clarify what types of crimes are to be considered. It may be that homicides are very different causally from robberies. It may be that homicides of men differ causally from homicides of women. Refining the dependent variable is part of the analytic process, but eventually the researcher must stake out what it is he or she will be explaining and differentiate what is beyond the scope of the research effort. Less can be more. That is, it is best to have a limited scope that involves feasible research.

2. *Having clear hypotheses.* Based on literature review and brainstorming, the researcher should develop a list of all variables thought to affect the dependent variable(s) that has been selected. The researcher then identifies causal relationships among sets of variables. The literature review is essential and must not be skipped because, ideally, hypotheses arise from theory. Be specific. It is not enough, for instance, to say that A and B cause X, because this could mean several things: (1) that A causes B, which in turn causes X; (2) that A causes X, and B independently also causes X; or (3) that the joint interaction of A and B causes X. Specifying the direct, indirect, and interaction effects in the model can be clarified by diagrams in which variables become circles, effects become arrows, directions are plus and minus signs, and size of effect is shown by size of arrow or by a coefficient from 0 to 1.0 attached to the arrow. As the number of variables in the researcher's model exceeds three, such diagrams are almost a necessity.

3. *Having alternative hypotheses.* The default assumption in research writing is that more than one hypothesis or model will fit any given set of data. It is not enough to demonstrate that the data fit the hypotheses the researcher has advanced. Other theories may fit the data too. Ultimately, the researcher will never be able to prove his or her hypotheses to be "true," although it may be possible to disprove them. The best the researcher can do, and should do, is to compare the fit of the data to his or her model with the fit of the data to plausible alternative models, particularly those implied by the professional literature. In doing so, the researcher develops a list of hypotheses and alternative hypotheses to be investigated.

4. *Operationalization of variables.* Operationalization involves making a list of all variables mentioned anywhere in any of the hypotheses or alternative hypotheses for models to be investigated. All the items on the list must be translated into operational terms. For every term (e.g., *effectiveness*) the researcher must have at least one and ideally four or more indicators (such as favorable client response to evaluation items, differential cost of service compared with a reference, objective progress measures). Operationalization of variables is discussed more fully in a later section of this guide. Operationalization, of course,

will feed into later research processes, such as selection and application of appropriate methodologic procedures, also discussed in later sections.

If an analytic question involves hypotheses that cannot be confirmed or disproved by empirical means, then these are *metaphysical* propositions inappropriate for scientific and social scientific analysis. The philosopher Popper's "Falsifiability Criterion" should apply to quantitative analyses: propositions must be falsifiable at least in principle if they are to be investigated empirically. The empirical researcher must be agnostic on metaphysical propositions (the term "agnostic" has a root meaning of "unknown" or "unknowable"). For instance, the explanation that a set of labor riots occurred because of the Marxian process of synthesis and antithesis is not falsifiable and is, therefore, a metaphysical proposition. As a theory, synthesis/antithesis may give the researcher some insights, but ultimately it explains everything and is merely a vocabulary that can be used like similar vocabularies (such as systems theory) in the process of explaining something empirically by other means. Those other means must involve empirically falsifiable propositions.

Of course, difficulty in obtaining evidence is not a reason to consider a proposition metaphysical. For instance, once-untestable propositions in astronomy have become testable with the Hubble space telescope, but these were never metaphysical propositions. It was only that measurement technology had not been developed. Still, for practical purposes, whereas the theorist has more latitude, the empirical researcher must be content with propositions testable by today's available methods and techniques.

Empirical accuracy of a theory is not enough. The researcher should bear in mind the common observation that for any set of data, there will be two or more competing explanations that explain the data satisfactorily, sometimes to an equal degree. How then to decide among theories? The scientific method puts forward three additional criteria for good theory: (1) deductive fertility (the theory is best that generates the largest number of interesting, testable hypotheses); (2) generality (the theory is best that explains the largest range of phenomena); and (3) parsimony (the theory is best that explains the data using the smallest number of assumptions and axioms). A parsimonious, generalizable, fertile theory is the nirvana of empirical research.

In summary, a good analytic topic is one that centers on an empirical puzzle. Two or more strategies (theories) may unlock the puzzle. Each engenders at least one empirically testable proposition, allowing the researcher to see which better fits the data that may be collected. The puzzle is interesting. The solution is nonobvious. Usually the hypotheses involved in proposed solutions grow out of significant strands in the literature of the discipline. When the puzzle is solved, the solution will have many implications for theory and practice. If all this is true, one has selected an analytic topic very wisely.

B. Analytic Topic Checklist

1. Have you selected a topic capable of sustaining your interest over the length of time (usually twice as long as you initially expect!) needed to complete your paper, thesis, or dissertation?
2. Is your topic interesting enough to motivate you for the long run, but resistible enough that it does not consume you to the neglect of obligations in your studies or career?
3. Did you choose a topic early and plan your research and writing in a framework that enables you to meet deadlines, avoiding a last-minute writing crisis?
4. Have you selected a topic your audience will deem to be substantive? For dissertations, have you identified a topic that can be considered an original contribution to your field (because it confirms/refutes/amends a theory pertinent to your discipline, fills a gap in your discipline's literature, exemplifies new research methodologies, or establishes new baseline or trend data important to questions in your discipline)?
5. Have you carefully examined your central question, considering each key word for possible multiple meanings, and the differences between it and its various common synonyms?
6. Have you identified your dependent variable(s)?
7. Have you considered if your dependent variable is actually of two or more types, each requiring a separate causal explanation?
8. Have you consulted the professional literature to develop a list of relevant independent variables?
9. Have you avoided taking notes on everything related to your topic, instead taking notes in a format closely tied to the specific testable propositions you are analyzing?
10. In taking notes, have you kept an accurate record of full references so you will not have to go back (e.g., page numbers for direct quotations)? Do you know the exact reference requirements of the format (the APA format from the American Psychological Association is common) you will be using?
11. Have you related your dependent and independent variables to each other, possibly in the form of a diagrammatic model?
12. For each pair of variables that would be connected by an arrow in a diagrammatic model, have you considered whether there is a third variable that might intervene or be common antecedent causes of both?
13. Have you researched and developed possible alternative models for the same data?

14. Have you identified outliers (cases that deviate greatly from what your theories and propositions would lead you to predict) and whether they call for a second (or more) separate theory from the one you are investigating primarily?
15. Have you looked at other dissertations or journal articles in this area with a view to refining your model and its alternatives?
16. Have you considered interdisciplinary perspectives (economic, psychological, social, cultural, and political)?
17. Have you written out a formal research statement that summarizes your research objectives?
18. Are you prepared to have your dissertation or paper evaluated in terms of whether you have accomplished the purposes contained in your research statement?
19. Are you sure your research statement does not duplicate work already published?
20. Have you given particular care to selection of your title, which will play a significant role in what readers expect you to accomplish in your dissertation, thesis, or paper? Could a reader interested in your topic find your work doing an electronic search based on keywords in your title? Does your title refrain from implying a higher level of generalization of your findings than your data warrant?
21. Have you avoided subjects in which you have a personal stake or "axe to grind?" These are conflicts of interest to be avoided for their own sake and because your research must both be and have the appearance of being objective.

C. Analytic Research Examples

In pursuing an analytic topic, the researcher asks common sense questions that often arise from his or her review of the literature. These questions include: What relationships are discussed in this article and how do they relate to the propositions I am studying? What are the influences, constraints, and linkages among variables and agents in the model underlying the article, and how do they relate to my model? Which variables or information are left out, either in the article or in the researcher's model, that would make an important difference? What evidence is presented, and by what methods, and did the method of measurement influence the findings? What alternative theories are mentioned and do they apply to the researcher's model as well? Is there any discernible bias by the writer of the article, and has the researcher considered such possible biases in his or her own work? What assumptions are made in the article, and would other assumptions be plausible and lead to different conclusions?

Example 1: Electing Nixon. In specifying a subject as an analytic topic, it is common in undergraduate social science writing to write about something with-

out ever having an analytic topic. For example, a student may select the subject, "Why Richard Nixon was Elected President," then present historical information related to this topic. The essay may be well-organized, informative, and deserving of an "A." However, when framed this loosely, the reader is very likely to be left without the basis for a well-supported conclusion about more specific topics, such as "Was the margin of Electoral College votes resulting in Nixon's election accounted for by popular reaction against the Vietnam War under President Johnson, and not to Johnson's "Great Society" liberal agenda, civil rights, or his handling of the economy?" This more specific question meets the four necessary conditions of an analytic topic.

1. There is a dependent variable: Nixon's margin of votes.
2. There is a plausible explanation: a popular reaction against the Vietnam War in the time leading up to Nixon's election, and it is possible this was a dominant factor.
3. There are plausible counterhypotheses: popular reaction against Johnson's "Great Society" liberal agenda, civil rights, or his handling of the economy.
4. The variables can be operationalized: all the terms that appear in the hypotheses can be given concrete meaning in terms of such indicators as Electoral College votes on the one hand and public opinion poll items on the other.

Example 2: Environmental regulation. A second example is environmental regulation. The section on "Brainstorming" included an example of writing on the subject of "Environmental Regulation." Brainstorming resulted in this list of subjects:

- Air pollution.
- Noise pollution.
- Water pollution.
- Chemical pollution/hazardous waste.
- Endangered species.
- Is it effective?
- Cost/benefit analysis.
- EPA agency.
- EPA regulations versus property rights.
- Conservatives vs. liberals on environmental regulation.
- Al Gore.
- Politician's platforms on environment.
- Federal funding for environment.
- State funding.
- Private costs/funding.
- Private sector ads on environmental issues.

Public opinion on environment, causes of.
Smokey the Bear.
Conservation.
Conservation/environment in 1900s versus 1990s.

The next step would be to specify possible analytic topics derived from this list of subjects. This is a creative process with no specific right answer. However, good answers will meet the four criteria: (1) a dependent variable or set of dependent variables; (2) a hypothesis or theory relating one or more independent variables to the dependent variable(s); (3) plausible alternative explanatory models; and (4) variables that can be measured operationally.

Are corporate issue ads more effective than environmental group issue ads in impacting public opinion? One could create a small group experiment that exposed subjects to corporate and environmental group ads and measured reactions.

Was the Forest Service's "Smokey the Bear" campaign effective? If evidence could be obtained on variations in school systems' use of "Smokey the Bear" materials (based on Forest Service records for the order of such material) and on trends in forest fires before and since the "Smokey" campaign in the districts selected, then the effect of the campaign over and beyond district fire trends could be assessed.

Is the environmental movement of the 1990s less partisan than the conservation movement of the early 1900s? The political party affiliations and activities of board members of leading conservation/environmental groups in the two periods could be investigated to determine if connections to political parties (Progressive, Democratic, Republican) have the same pattern now as then. Explaining why would be the focus for yet another analytic topic.

Does federal funding of water quality supplement or displace state funding? Data could be obtained on state funding on municipal sewage treatment, then on selected areas, showing historical trends and variations, examining the extent to which federal funding in the same area seems supplementary (raising the total funding trend line but not affecting the state funding trend line) or displacing (causing reversals in the state trend line), or combinations of both.

Example 3: Comparative research. In a course on comparative politics, an analytic paper would go beyond merely describing chronological facts and events based on information from encyclopedias and newspaper articles. The typical analytic paper would pick at least two countries that have something in common (this will be the control variable) but that have one other variable that varies (this is the independent variable) and is thought to affect some dependent variable of interest such as regime stability, the status of women, or economic development policy. Typically, the dependent variable(s) vary among the nations being studied. For instance, one might study the impact of high oil prices in the 1970s on

the domestic political stability of African nations. To do so, it would be important to include at least one country where the oil crisis had destabilizing effects and one where it did not. The writer is then in a position to identify and assess one or more independent variables that explain the difference, with the control being that all countries in the study were similarly affected by the worldwide oil crisis (this might lead to excluding a nation with its own internal oil reserves).

In comparative research, one is looking at questions in which two countries have similar control factors but differ in outcomes because of some independent variables that the researcher is trying to analyze. For instance two countries like Taiwan and Singapore share the control variable of having a large middle class, yet in terms of an outcome dependent variable, such as type of government, the former country is more democratic and the latter more autocratic. The analytic question is why, and this question cannot be answered merely by describing the history of each. Rather, independent variables must be analyzed (e.g., differences in economic structure) comparing the nations. Of course, the more nations compared by the same methodology, the more confidence the researcher will have in the generalizability of his or her findings.

BIBLIOGRAPHY

- Brown, A., Dowling, P. (1998). *Reading Research/Doing Research*. Philadelphia, PA: Falmer Press. [A book for doctoral students in education who are considering their dissertation. Includes practical guidance on the development of a research question as well as related topics, such as doing literature reviews.]
- Krathwohl, D. R. (1988). *How to Prepare a Research Proposal : Guidelines for Funding and Dissertations in the Social and Behavioral Sciences*. Syracuse, NY: Syracuse University Press. [A bit dated but still helpful.]
- Levine, S. J. (1998). *Guide for Writing a Funding Proposal*. Available online at the website <http://www.canr.msu.edu/aec/dissthes/propint.htm>.
- Seltzer, R. A. (1996). *Mistakes That Social Scientists Make: Error and Redemption in the Research Process*. New York: St. Martin's Press. [A collection of short, true anecdotes by leading social scientists about mistakes they have made. Not limited to issues pertaining to framing analytic questions, this work is nonetheless an excellent one to read early in one's research process, particularly for doctoral students.]
- Silverman, D. (1999). *Doing Qualitative Research: A Practical Handbook*. Thousand Oaks, Ca: Sage Publications. [Covers narrowing down a topic, keeping a research diary, writing research reports, and presentation of research to different audiences.]
- Steeves, R. H., Kahn, D. L., Cohen, M. Z. (1996). Technical notes: Asking substantive theory questions of naturalistically derived data. *Western J Nurs Res* 18(2): 209–212.
- Stern, P. C., Kalof, L. (1996). *Evaluating Social Science Research*. 2nd ed. New York: Oxford University Press. [Generally relevant, but see especially chapter 1, "Scientific and Nonscientific Statements of Fact."]