

Geddes, Barbara. 2003. *Paradigm and Sand Castles*. Ann Arbor, MI: University of Michigan Press.

Chapter 1

Lesson: We should break the theory down into different parts and examining their merits each by each (less holistic).

We are always dealing with “big structures, large processes, and huge comparisons” (Tilly 1984).

Since most analyses found themselves in uncharted territory when it came to building theories that use political causes to explain political outcomes in developing countries, the result was the emphasis on contingency and voluntarism found in early literature. Previous scholars of developing countries, when studying democratization and economic reform, tended to focus more on individual cases and the details of decision-making process and found old theories wanting.

- The failure to recognize the problems caused by government strategies.
- Arguments in the dependency tradition continued to attribute these problems to structural features of the relationship between late developing countries and the international capitalist economy.
- Arguments in modernization structure continued to attribute developmental difficulties to culture (p.18).

Consequently, arguments, theories and even paradigms tend to rise and fall in rapid succession, leaving little behind to show that they ever existed.

- Inductive fact-gathering missions result in a disorganized mass of information.
- The failure to organize and store evidence in ways that make it accessible to others raises the costs of replication and thus also slows the theoretical progress.
- Many aspects in a given theory remain untested (e.g. the role of diffusion of values and technologies in transforming traditional European societies into modern ones).
- Selecting cases for study from one end of the outcome continuum we hope to explain (anecdotal cases for illustration instead of systematic tests) increases the likelihood of reaching mistaken conclusions. Anomalies that might have led to abandonment of some ideas and minor revision of others could instead be interpreted as challenging the whole paradigm.
- Analysts did not ignore evidence, but they used it selectively to develop, support, or test theories. Observers may note the existence of some trait or phenomenon, describe and discuss it in an ad hoc way, but never incorporate it or its implications into theories.
- Regression, rather than focused attention on how the causal factors work, puts more emphasis on the identification of causal factors.
- General inattention to evidence impeded the emergence of a better understanding of economic development that might have usefully informed policy choice.

Paradigm vs. approach

- Paradigm: Like ideologies, a paradigm simplifies the world in a compelling way by forming a coherent worldview. They need to be fruitful, explain previously unexplained regularities (*if there are any*), and lead to new questions and puzzles. Earlier theoretical understandings are first undermined by inconsistent facts and then replaced by newer theories that can accommodate both of old and new theories. They structure future research, determining what **facts** are theoretically salient; informing what constitutes a **paradox/puzzle** and what **questions** urgently require answers; identifying what **cases** to be examined and what **evidence** considered meaningful.
 - Kuhn (1970): The dominant understanding of a particular class of phenomena at a particular time. That is, paradigms reflect the time in which they develop.
 - Geddes (2003): Collections of theories, hypotheses, applications, and favored

methodologies.

Paradigms encompass a set of factual and explanatory knowledge claims, namely theories, widely accepted by adherents.

- Approach: A claim that argues certain factors deserve attention without articulating specific hypotheses about them and the belief that certain research methodologies are the most useful and appropriate means of gaining understanding. Examples include instead, rational choice (since it contains uncountable numbers of hypotheses and theories, many of which are inconsistent with each other) and historical institutionalism (new institutionalism).

Thought: *For many scholars of Chinese politics, since they heavily rely on fieldwork in a few sites to construct their "theories," scholars, after traveling to different sites, tend to reject existing theories promptly without deeply reviewing the logical connection between the new cases and the theories based on old cases.*

Chapter 2

Key lessons from Geddes' example: Coherent deductive arguments can be devised to explain constituent processes and hypotheses derived from the arguments can be tested.

- Disaggregate a big question into a series of processes (transition as the functioning of authoritarian regimes). List a set of topics (pp.45-6).
- Determine the focus of analysis (pp.47-50): Bargain between leaders' interests and intra-regime politics. In particular, transition involves different processes that in turn depend on individual types of authoritarian regimes. Different types also have different decision-making mechanisms. The break-up must be based on one's theories.
- Use simple models (e.g. coordination games; first-mover advantage - something unexpected learned from the model) to plan your arguments (p. 56, 59, 61).
- Draw testable implications from the argument (p.69).
- Define the concepts carefully (p.71). Gather an appropriate dataset and try different classification schemes.
- Policy implications: Where should the US government intervene?

Chapter 3

The close examination of an anomalous case with a particular outcome can also serve a useful role in either generating a proposed revision of current theory or suggesting domain conditions not previously understood. However, the revision should not be accepted until it has been tested and confirmed on a representative sample of cases (**pp.90-1**).

If cases are only selected with much reference to their position on the dependent variable, it will be unwarranted because the results may be "over-fitting," though it is almost inevitable at the early stage of theory-building (**p.95**). Also, over-confidence on the cases may tempt scholars to jump to the conclusion that any characteristic that the selected cases share is a cause, and inferred that relationships (or absence of relationships) between variables within the selected set of cases reflect relationships in the entire population of cases.

Therefore, to test any hypothesis, one must first identify *the universe of cases* to which the hypothesis should apply and then find or develop measures of the hypothesized causes and effects before one starts selecting cases for analysis from the pool of cases. The theory or hypothesis being tested determines the appropriate unit of analysis (e.g. individual, country, state, nation, and region) and the universe of potential observations.

Measurement and Coding (p.98)

The purpose of coding sheets and coding rules is to help make sure that the same factors are assessed in every case and that they are all judged using the same criteria. For a conceptual phrase (e.g. labor repression) no doubt has somewhat different connotations for scholars with different areas of expertise, but the person who has read the coding scheme will have a very clear idea of what is meant by the term that we referred here.

The point is not to demonstrate that the hypothesis that labor repression, based on the studies of the NICs in East Asia, contributes to growth is false. It may have a small positive effect. It might be that the addition of appropriate control variables (e.g. region) or an elaborate lag structure would make clear a relationship that does not show in the simple tests done here (p.105).

Time Series, Case Studies, and Selection Bias

Case studies can often be thought of as non-quantitative time-series research designs, usually examining a single country over a period of time for the purpose of explaining some outcome at the end or showing the effects of some change that occurred during the time examined.

While case studies are often criticized as single data points and hence incapable of revealing anything about cause-and-effect relationships, most can be more reasonably thought of as a series of observations of the same case at different times (e.g. not cross-sectional comparison).

Whether such a research design involves actual selection bias depends on whether the variation time within the case reflects the full range of outcomes in the relevant universe.

The key concern for the researcher is to identify the universe relevant to the question being asked or hypothesis being tested. As elsewhere, the question under examination determines the appropriate universe (p.118). Depending on the questions proposed, a case study of the particular place and time of interest may not provide the answer because it is possible that selecting multiple observations of the same case will have the effect of holding constant or near constant some of the true causes of the outcome of interest, even if the dependent variable spans a considerable range.

These causes of the within case variation can be less important causal variables that belong in a complete explanation, or they can be idiosyncratic factors that affect this case but not others and therefore do not belong in a general explanation.

If one knows quite a bit about the underlying causal model, a single-case time-series design can be a good way to assess the effect of one potential cause while holding many other things constant (because they do not vary within the single case), but it will be less useful in the more typical situation where the analyst does not know the underlying model (p.119).

The methodological point is that even if one cares only about one case, the best research strategy for discovering these causes of variation may require examining other cases. Important causes, such as the basic thrust of development strategy, may not change very much within a few decades in a single country. Consequently, analysts may overlook their importance and instead concentrate attention on less important causes or on conjunctural factors that turn out to have no general causal effect. Case studies generally help to explain zigzags in the trend line, but they sometimes offer little leverage for explaining the trend itself (p.123).

Chapter 4

Criteria for case selection (p.132):

1. First, cases should be representative of the domains of the theories they are intended to test.
2. Second, cases used for testing arguments should be different from the cases from which the

arguments were induced, because the analyst would not propose an argument in the first place if it did not fit the cases that generated it.

The Role and Appropriate Use of Case Study Evidence

Why case study? This is so mainly because much of the evidence used by comparative political scientists must be gathered through fieldwork or the study of archives country by country. Large-*N* data sources are becoming more common and easier to access, but available data sets have little relevance for many subjects and generally contain only superficial information (p.133).

Some common methodological pitfalls of case studies:

- The impossibility of testing hypotheses when variables outnumber cases.
- The inability to recognize and ignore idiosyncratic features of the cases chosen-called over-fitting theory to data in quantitative work.
- The standard and unanswerable criticism of studies based on single cases, or observations, is that they cannot be used to confirm or disconfirm hypotheses.
- With evidence from only a single case, one has no information at all about whether the universe of cases about which one wants to generalize would, if examined, form a line of points from the origin to the upper right quadrant, a line of points from the upper left to the lower right, an amorphous blob, or any other imaginable shape (see also KKV 94).

More generally, as long as the analyst has at least one more case than the number of hypothesized causes, estimates of relationships can be made. But an estimate based on only one more case than the number of possible causes will not be reliable. The more additional cases are included, assuming they have not been selected in ways that bias conclusions, the more reliable estimates become.

One simple way to increase the number of cases is to increase the number of observations within the small number of everyday language cases being studied (pp.136-7). One of the purposes of this chapter is to encourage analysts to make the common practice of including multiple observations in a single everyday language case self-conscious and systematic rather than ad hoc and causal (p.137).

Case Studies in Path-Dependent Arguments

Path-dependent arguments as used in comparative historical analysis face three methodological challenges to the knowledge claims they make (p.140-1):

1. How do we know that the proposed explanation of the outcome at a particular node is the correct one?
2. How do we know that the final outcome we observe (the legacy) was really caused by choices at the series of historical junctures that preceded it, rather than by something else?
3. If two path-dependent arguments set out to explain the same outcome, and one argument concludes that choices made at one historical juncture determined the final outcome while the other identifies a different juncture as critical, how can we tell which is correct?

They arise from two common features of the research design used in comparative historical studies: reliance on small numbers of cases, usually those from which the argument was induced; and vague or ambiguous definitions and measurements.

Whenever possible, the big argument should be broken at its branching points, and hypotheses that purport to explain choices at different nodes should be tested on additional cases that appropriate initial conditions. Efforts should be made to see if hypothesized legacies also occur in these other cases.

Non-quantitative Operationalization and Measurement

Most arguments that rely on case study evidence involve complex causes, outcomes, and limiting

conditions that must be operationalized and measured with great care if tests of the argument are to be persuasive.

- **Operationalization**, as used in quantitative research, refers to the choice of observable indicators that can be used as proxies for abstract and unobservable concepts (e.g. the level of development and GDP per capita).
- **Measurement** involves the assignment of particular cases to particular values or categories of the operationalized concept. In quantitative research, interval or at least ordinal levels of measurement are usually possible. If we take GDP per capita as an indicator of development, for example, each country can be given an interval level value on the indicator.

Conclusion

General advice for case studies, regardless of the number of cases involved (p.172):

- The inclusion of as many cases as possible in order to increase the robustness of findings.
- The use of “structured focused comparison,” that is, measuring the same causal factors and outcomes in the same way in each case.
- The use of cases other than those from which hypotheses have been induced to test arguments.