PART ONE

The Logic of Analytic Induction

1

Classic Analytic Induction

Analytic induction (AI) was a popular technique in U.S. sociology during the early decades of empirical social research. The method was first formalized by Florian Znaniecki (1934) in his book *The Method of Sociology*. Znaniecki believed AI to be more scientific than "enumerative induction" (known today as correlational analysis) because of AI's emphasis on "universals"—invariant connections between antecedent conditions and outcomes (Tacq 2007). The basic idea was that the researcher should pinpoint antecedent conditions uniformly shared by instances of an outcome. Thus, the method focuses on *positive* instances of an outcome and attempts to provide an account of the outcome's etiology based on an analysis of shared antecedent conditions.[1]

AI's focus on the antecedent conditions shared by instances of an outcome is rooted in John Stuart Mill's method of agreement. He argues that if two or more instances of the phenomenon under investigation have only one circumstance in common, that one circumstance is the cause (or effect) of the given phenomenon (Mill 1967). In short, his method of agreement dictates close inspection of the antecedent conditions shared by instances of the phenomenon under investigation. While he frames the definition of the method of agreement in terms of a single shared condition ("only one circumstance"), his argument can be easily extrapolated to situations where there is more than one shared circumstance. Together, multiple shared conditions can be understood as contributing causes in situations where their combination is seen as a causal formula or recipe.

Both AI and Mill's method of agreement are formalizations of a very common technique for deriving empirical generalizations. Humans look for connections in everyday experiences and draw conclusions from repeated observations. For example, the observation that I must leave home for work by 7:00 a.m. in order to avoid heavy automobile traffic is an empirical generalization, based on repeated experiences. A consistent antecedent condition for the avoidance of heavy morning traffic is on-time departure for my commute to work. Of course, the consistency of the connection may be far from perfect, but still consistent enough to guide my behavior.

While commonplace, the search for antecedent conditions shared by positive instances can be the basis for prizewinning research. Consider, for example, Elinor Ostrom's (1990) research reported in *Governing the Commons*. Her main target was a widely held view of common-pool resources: that, absent state oversight and management, such resources are likely to be abused and rendered unsustainable through overuse.[2] To counter this view, Ostrom studied a variety of *self-governing* common-pool resources, where there were successful collective efforts to achieve sustainability, orchestrated by the surrounding communities. Ostrom observed that these *positive* cases shared a number of characteristics, including, for example, rules that clearly defined who gets what, good conflict-resolution methods, users who monitor and punish violators, and so on. In short, she established important preconditions for community-based resource sustainability based on her analysis of positive cases. She won the Nobel Prize in Economics for her research.[3]

Another example of this strategy in comparative research is Daniel Chirot's *Modern Tyrants* (1996). Examining thirteen tyrants, drawn from diverse settings, Chirot writes that "tyrannies have come to power in states both big and small; in rich industrial and very poor agrarian societies; in countries with many centuries of statecraft in their tradition, and in brand new ones; in culturally united nations with a firm sense of identity, and in ethnically split states with almost no basis for common solidarity" (Chirot 1996: 403). He asks, "What generalizations can be drawn from these thirteen sad and diverse histories?" While acknowledging that his conclusions are probabilistic in nature (418), he offers eight generalizations based on his study of thirteen tyrants, noting, for example that "the more chaotic the economy and political system, the more they seem to be failing, the more likely it is that a tyrant will emerge as a self-proclaimed savior" (409).

AI is often overlooked as a formal technique because it is simultaneously ubiquitous and rare. It is ubiquitous because it is based, as just described, on a very common method of generalizing about empirical regularities from equivalent observations (Bernard et al. 2017). Why formalize or even cite a method that seems like common sense? By contrast, applications of *classic* AI are somewhat rare because of its requirement that researchers demonstrate *invariant* connections between outcomes and antecedent conditions. All exceptions to working hypotheses must be addressed and resolved. As detailed in this and subsequent chapters, this feature of classic AI mandates both in-depth knowledge of cases and conceptual agility on the part of researchers. For some analysts, the invariance requirement dictates a determined pursuit of disconfirming cases—positive cases of the outcome that do not exhibit the antecedent conditions specified in a working hypothesis (Katz 1983; Denzin 2006; Athens 2006).

This book, while building upon classic AI, ultimately relaxes several of its defining features in order to lay the foundation for *generalized* AI. For example, the invariance requirement is unrealistic for the work of many researchers and research projects. Typically, researchers have a fixed set of collected data and little or no opportunity to return to the cases for more evidence or to seek out new cases that might challenge a working hypothesis. Another example: classic AI has little use for frequency criteria because a single disconfirming case can torpedo a working hypothesis. For most empirically minded social scientists, however, the weight of the empirical evidence matters, and frequency criteria are considered not only informative, but often decisive (Goertz and Haggard 2022; Miller 1982).

This chapter provides an extensive discussion of classic AI, focusing on the logic of the approach. First, I examine several classic examples of the approach and then formulate the method as a series of steps. Classic AI is both dynamic and iterative. It is a research approach that builds empirical generalizations on the basis of in-depth case knowledge. Second, I examine classic AI's understanding of causation, contrasting it with more conventional forms of analysis.

SOME EXAMPLES OF CLASSIC AI

Early, exemplary studies utilizing classic AI include Alfred R. Lindesmith's *Addiction and Opiates* (1947 [titled *Opiate Addiction*], 1968), Donald R. Cressey's *Other People's Money* (1953, 1973), and Howard S. Becker's *Becoming a Marihuana User* (1953, 2015). All three studies offer detailed portrayals of AI as a research process that builds a coherent argument based on in-depth analysis of cases.

Drawing on his interviews with more than sixty addicts, Lindesmith attempted to identify the antecedent conditions linked to opiate addiction. He argued that users become addicts only when they consciously use the substances to diminish the effects of withdrawal (Lindesmith 1968: 191). In other words, there is an important cognitive component to opiate addiction. Addicts must realize that this is why the effects are happening, and that no other physical ailment explains the painful withdrawal symptoms (1968: 191). If they do not attribute the withdrawal as such to their opiate use, and they believe that some other physical deficiency causes the side effects, they do not become addicts. Lee and Fielding (2004) summarize Lindesmith's argument as a specification of the process of becoming an addict: these individuals (a) use an opiate; (b) experience distress due to withdrawal of the drug; (c) identify or recognize the symptoms of withdrawal distress; (d) recognize that these symptoms will be alleviated if they use the drug; and (e) take the drug and experience relief (see also Becker 1998: 197).

The purpose of Cressey's study was to look at the sequence of conditions that lead an individual in a trusted financial position to embezzle money (Cressey 1973: 12). He gathered interview data from 210 convicted embezzlers, asking them about their experiences before, during, and after they were caught violating trust. After a lengthy process of reformulating hypotheses, identifying themes, and connecting them to a general concept, Cressey concludes that there are three necessary conditions: the individual (1) perceives that a personal, non-shareable financial problem has occurred, (2) rationalizes a reason for taking entrusted funds, and (3) believes that this is the only way to solve the non-shareable problem (1973: 139). It is important to point out that Cressey allowed for the possibility that a necessary condition could be satisfied in more than one way. For example, he identified three circumstances in which embezzlers "rationalized" their behavior: (1) they needed or wanted to borrow money, (2) they felt that the funds belonged to them, or (3) they felt it was a one-off situation (1973: 101–12).

Becker interviewed fifty recreational marijuana users in an effort to specify the process of becoming a user. His stated goal was to document the necessary conditions for recreational marijuana use (Hammersley 2011). He argued that there are three universal conditions or steps that must occur, at some point, in order for one to become a recreational marijuana user: (1) smoking it properly to induce a high, (2) recognizing and understanding the effects caused by the drug, and (3) learning "to enjoy the sensations" (Becker 1953: 242). Without satisfying all three conditions, an individual will not be able to become a recreational marijuana user. Becker draws an important distinction between those who use marijuana for pleasure and those who use the drug, but not for pleasure, and restricts his account of marijuana use to the former.

Based on these early applications, it is clear that AI is a discovery-oriented, abductive tool (Diesing 1971; Tavory and Timmermans 2014). It is also clear that because of its requirement of causal invariance, applications of classic AI tend to focus on antecedent conditions that are proximate to the outcome in question. Indeed, the antecedent conditions identified in these exemplary AI studies could be seen as constitutive of their outcomes (Turner 1953), which in turn suggests, to Lindesmith (1952), that the conditions are not only necessary but also sufficient.

In *Poor People's Lawyers in Transition*, Jack Katz (1982) offers a detailed illustration of AI's dynamic nature, especially the process of "double fitting" the conceptualization of causally relevant conditions with the conceptualization of the outcome. More recent applications of classic AI include the work of Hicks (1994), Gilgun (1995), Monaghan (2002), and Bansal and Roth (2000). In political science, there are several notable examples of work utilizing principles of AI. In addition to Chirot's *Modern Tyrants*, these include Guillermo O'Donnell and Philippe Schmitter's *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies* (1986), Crane Brinton's *The Anatomy of Revolution* (1938), O'Donnell's (1973) work on the origins of bureaucratic-authoritarian regimes in South America, and Juan Linz and Alfred Stepan's *The Breakdown of Democratic Regimes* (1978).

CLASSIC AI: STEPS

Various authors (e.g., Robinson 1951; Cressey 1973; Hammersley and Cooper 2012: 131–32) have attempted to capture classic AI's dynamic, iterative character by formalizing the method in terms of a series of steps:

1. Specify the outcome to be explained. Typically, the outcome is qualitative in nature. For example, it might be a "happening" or an occurrence like becoming an embezzler (Cressey 1973) or becoming a marijuana user (Becker 1953). The happening also can be meso- or macro-level (Katz 2001)—for example, episodes of mass protest against an authoritarian regime.
2. Collect evidence on a number of cases in which the outcome occurred. Usually, these are very clear instances of the outcome in question (Goertz 2017: 63–66). Some versions of classic AI (e.g., Lindesmith 1968) restrict the initial investigation to a single case, then add more cases one at a time (see also Robinson 1951; Lee and Fielding 2004). This restriction ensures that each case will be subjected to an in-depth assessment. However, for many investigations, this restriction is neither feasible nor warranted.
3. Identify the causally relevant antecedent conditions shared by these initial instances of the outcome. Formulate a working hypothesis on the basis of observed commonalities. Existing theory and substantive knowledge regarding relevant causal conditions for the outcome serve as preliminary guides. The commonalities identified by the researcher must make sense, on either substantive or theoretical grounds, as antecedent conditions.
4. Seek out and collect evidence on additional instances of the outcome. It is more important that the selected cases are diverse than that they are representative of a population (Goertz and Mahoney 2012: 182–85). Researchers should identify and study instances of the outcome that challenge their working hypothesis.
5. If cases are found that challenge a working hypothesis, then either the antecedent conditions or the outcome (or both) must be reformulated in some way. Typically, if the outcome is reformulated, its scope is narrowed so that the nonconforming cases are excluded from the purview of the working hypothesis. If the antecedent conditions are reformulated, the causal argument is altered so that the nonconforming cases are embraced in some way, typically through a strategy of conceptual realignment. In either situation, the process of reformulation should be both public and transparent.
6. Continue steps 4 and 5 until the evidence derived from additional instances no longer prompts reformulations of the working hypothesis or its empirical scope. The research has reached a point of theoretical saturation and an invariant connection has been established.

As this summary of classic AI's steps makes clear, AI's "dependent variable" is not a variable, but a constant; it is an outcome that is more or less the same across selected cases. This type of analysis is beyond the purview of conventional quantitative methods, which are focused on explaining variation in dependent variables by using variation in independent variables. Worse yet, examining only positive cases is viewed in the quantitative literature as an extreme form of "selecting on the dependent variable"—a great sin to be avoided, according to some authors (e.g., King et al. 1994). AI has little use for the analysis of the covariation of variables. Instead, the goal is to explain a constant, the outcome, with other constants—their shared antecedent conditions. The end result is a specific type of empirical generalization, one that is set-analytic, as opposed to correlational, in nature.

For example, the observation that social revolutions share peasant insurrections as an antecedent condition (Skocpol 1979) casts social revolution as a subset of instances of peasant insurrection. In this example, a *connection* between two sets (the set of countries with social revolutions and the set of countries with peasant insurrections) provides the basis for an empirical generalization. This observed connection stands on its own, without reference to variation in the presence versus the absence of either social revolution or peasant insurrection. Instead, the presence of peasant insurrection is linked to the presence of social revolution. It does not matter that there are many instances of peasant insurrection not linked to social revolution. By contrast, most empirical generalizations in the social sciences today are based on correlations between variables. For example, a researcher might offer an empirical generalization based on a positive correlation between social inequality and social unrest. In general, social scientists have not acknowledged connections between sets as a separate type of empirical generalization, distinct from those based on covariation.

THE LOGIC OF AI

AI was challenged as a technique for studying causation in 1951 by W. S. Robinson, in his article "The Logical Structure of Analytic Induction," published in the *American Sociological Review*, then and now the flagship journal of the discipline. His basic argument is that the method is fundamentally flawed because it can only identify necessary conditions, and therefore is not suitable for prediction. If it is used at all, it must be complemented with or followed by "enumerative induction" (i.e., correlational analysis) to certify that the causal factors identified using AI are in fact predictive (Miller 1982; Goldenberg 1993).

To fully grasp the substance of Robinson's critique, it is important to consider the essential differences between correlational analysis (Robinson's favored technique; see also Miller 1982) and set-theoretic analysis. The core principle of correlational analysis is the assessment of the degree to which two series of values parallel each other across comparable cases.[4] The simplest form is the 2 × 2 table

TABLE 1-1 Correlational approach to causation

	Cause absent	Cause present
Outcome present	Cell *a:* cases in this cell contribute to error	Cell *b:* many cases should be in this cell
Outcome absent	Cell *c:* many cases should be in this cell	Cell *d:* cases in this cell contribute to error

TABLE 1-2 Set-analytic approach to causation

	Cause absent	Cause present
Outcome present	Cell *a:* cases in this cell contradict necessity	Cell *b:* cases in this cell are consistent with both necessity and sufficiency
Outcome absent	Cell *c:* cases in this cell are not directly relevant to either necessity or sufficiency	Cell *d:* cases in this cell contradict sufficiency

cross-tabulating the presence/absence of a cause against the presence/absence of an outcome (table 1-1). Correlation is strong and in the expected direction when there are as many cases as possible in cells *b* and *c* (both count in favor of the causal argument, equally) and as few cases as possible in cells *a* and *d* (both count against the causal argument, again, equally).

Because cases in cell *c* are as hypothesis-confirming as cases in cell *b*, researchers must guard against including irrelevant cases in their analyses. Irrelevant cases would likely reside in cell *c* (cause absent/outcome absent) and thus spuriously confirm the researcher's hypothesis and contribute to a Type I error. In short, researchers who utilize the correlational template (which embraces the bulk of conventional quantitative social science) for their analyses must ensure that the cases they include are all valid candidates for the outcome in question (see chapter 3).

The set-analytic approach to this same 2 × 2 table differs substantially from the correlational approach (Miller 1982), as demonstrated in table 1-2. Each of the four cells has a different interpretation (Goertz 2017). The analytic focal point is cell *b*, which captures the cases that exhibit both the cause and the outcome (2017: 63–66). But is the cause a necessary condition for the outcome? If so, then cell *a* should be empty.[5] Is the cause sufficient for the outcome? If so, then cell *d* should be empty. Thus, the set-analytic approach to the 2 × 2 tabulation of outcome by cause is to separate the two causal relationships embedded in the table. After all, a cause can be sufficient but not necessary, and it can be necessary but not sufficient.[6] Note also that cases in cell *c* (cause absent/outcome absent) are not involved in either assessment. While cell *c* cases are integral to correlational analysis, computationally equal in importance to cases in cell *b* (cause present/outcome present), they play no direct role in the set-analytic approach.

Thus, two very different kinds of disconfirming cases are represented in table 1-2 (see also Ragin 2008). Cell *a* contains cases where the outcome is present, but the hypothesized cause is absent; cell *d* is the opposite—the hypothesized cause is present, but the outcome is absent. Robinson (1951) is correct in noting that classic AI focuses primarily on necessary conditions. Classic AI's central concern is the first row of table 1-2, especially the challenges to a working hypothesis posed by disconfirming cases in cell *a*.[7] Furthermore, addressing and reconciling cell *a* cases is the primary means of theoretical advancement in classic AI. Thus, the technique has little interest in cases that occupy cells *c* and *d*. Cases in cell *c* (cause absent/outcome absent) are not directly relevant to the assessment of either necessity or sufficiency and thus can be safely set aside. Some AI researchers (e.g., Cressey 1973: 31) do utilize hypothetical cases in cell *c* by arguing that their cell *b* cases were in cell *c* before they experienced the relevant causal conditions associated with the outcome. In effect, these cases traveled from cell *c* to cell *b* once the right causal conditions were present.

The issue of disconfirming cases in cell *d* (cause present/outcome absent), however, deserves further attention. Cases in cell *d* could be seen as AI's blind spot, because it is standard AI practice to focus on cases with the focal outcome—meaning that cases in cells *c* and *d* are routinely bypassed. However, there are several factors to consider regarding AI's apparent disinterest in cell *d* cases:

1. It is important to note that AI focuses primarily on questions regarding how outcomes happen. As explained in detail in chapter 3, AI views outcomes as happenings and seeks to account for happenings in terms of their shared antecedent conditions. Cases in cell *d* fail to exhibit the focal outcome and thus can provide very little useful information regarding how it came about (see also point 6 below). Cases in cell *a*, by contrast, experienced the outcome but not the hypothesized causes and thus offer important raw material for clarifying the outcome's etiology.
2. From the viewpoint of AI, cell *d* cases experience a different outcome, compared to cell *b* cases. The cell *d* outcome is deserving of separate investigation, culminating in a specification of its etiology (Kidder 1981). In short, the cell *d* outcome, if there is one that is shared by these cases, should not be treated simply as instances of the absence of the focal outcome (i.e., as mere negative cases), but as instances of an alternate outcome that is worthy of separate analytic attention. For example, if Cressey (1973: 31) found that his cell *d* cases resorted to suicide, not embezzlement, once confronted with a non-shareable financial problem, that outcome would become the focal point of a separate investigation.
3. Typically, however, cell *d* cases display a wide variety of nonfocal outcomes and are thrown together only by the fact that they did not experience the focal outcome (see chapter 4). Of course, the researcher may choose to

document the different outcomes and identify the antecedent conditions specific to each, but this effort would be secondary to understanding the focal outcome, displayed by cell *b* cases.

4. Cases in the first row of table 1-2 (cells *a* and *b*) comprise a relatively well-defined and circumscribed set of cases—they are all instances of the focal outcome. The second row, which contains cases lacking the focal outcome, is not so well defined (see chapter 4). Presumably, cases in the second row are or were viewed as candidates for the focal outcome; otherwise, their inclusion in the analysis would not be justified. However, the definition of candidacy for the focal outcome may be arbitrary, which in turn makes the decision regarding which cases to include in the second row contestable (Ragin 1997, 2009). By focusing on the first row, AI bypasses the problem of circumscribing the set of valid negative cases—cases that might have experienced the outcome, but did not.
5. In general, AI focuses on shared antecedent conditions for an outcome. Very often, the "cause" in table 1-2 is not a single condition, but a combination or sequence of conditions. The greater the number of antecedent conditions the researcher is able to identify, the less likely there will be cases in cell *d*. In short, as more antecedent conditions are added to the mix, the number of cell *d* cases that meet them all may be correspondingly diminished. Full articulation of relevant antecedent conditions could easily lead to an empty cell *d*, which would provide evidence consistent with an argument of causal sufficiency (Ragin 2008: 17–23). The fact that cell *d* can be emptied of cases as the researcher specifies more antecedent conditions explains, in part, why Lindesmith (1952) responded to Robinson's (1951) critique by arguing that the conditions identified by classic AI were not just necessary, but necessary *and sufficient*.
6. Cases in cell *d*, if they exist, have a potentially useful role—they can help the researchers refine their articulation of the etiology of the focal outcome. Because cell *d* cases display the causal conditions but not the focal outcome, close inspection of cell *d* cases can lead to the identification of conditions that either neutralize one or more of the antecedent conditions manifested in cell *b* cases or block the outcome altogether. However, because cases in cell *d* are likely to be heterogeneous (see point 3 above) and their inclusion as negative cases contestable (see point 4), they may offer only limited analytic leverage.
7. Classic AI's invariance requirement tends to favor the identification of antecedent conditions that are proximate to and constitutive of the outcome. Turner (1953: 608) goes so far as to argue that applications of classic AI culminate in constitutive definitions of the outcome, not causal explanations. Consequently, any case that displays the antecedent conditions specified by

> the classic AI researcher may automatically display the outcome. As a result, cell *d* cases (condition present/outcome absent) may be extremely rare, if they exist at all.

Given these considerations, AI's apparent disinterest in cases in cell *d* is understandable. Note also that several of these considerations upend Robinson's (1951) critique of AI. His critique focuses on AI's inability to predict an outcome, based on its failure to take into account cases in the bottom row of table 1-2. But Robinson failed to consider (1) the goal of AI—to explain how an outcome happens, not to predict its distribution in a population or a sample; (2) the problematic nature of the definition of relevant negative cases—that it is often arbitrary and contestable; (3) the heterogeneity of negative cases—that they may include many alternate outcomes, each suggesting a possible avenue for further investigation; and (4) the fact that there may be no cases in cell *d*, due to a comprehensive specification of relevant antecedent conditions.

LOOKING AHEAD

Chapter 2 presents analytic strategies for addressing disconfirming cases, building on table 1-2 as a template. Altogether there are six main strategies, all focused on emptying cell *a* of cases. In general, the strategies are consequential from a set-analytic perspective because their goal is to document an invariant relationship between one or more antecedent conditions and an outcome. By contrast, these strategies typically yield only very modest gains from a conventional variable-oriented perspective.