

Reconciling Disconfirming Cases

Classic AI rests on three main pillars: (1) focusing on positive instances of an outcome, (2) identifying their shared antecedent conditions, and (3) assessing the substantive and conceptual implications of disconfirming cases. This chapter describes key analytic strategies involved in implementing the third pillar.

Despite its name, AI is both inductive and deductive (Hammersley and Cooper 2012; Katz 2001; Manning 1982). For example, the identification of causally relevant antecedent conditions depends both on case knowledge (the inductive aspect) and on theory and prior research (the deductive aspect). Likewise, the working definition of the outcome to be explained, while open to revision as the research proceeds (the inductive aspect) is initially based on the researcher's preexisting knowledge and interests (the deductive aspect). It is appropriately labeled "induction," however, primarily because a core principle of AI is that researchers should attend to, and try to reconcile, disconfirming cases—*those that display the outcome in question, but not the hypothesized antecedent conditions*. Recall, from chapter 1, classic AI's pivotal fifth step: "If cases are found that challenge a working hypothesis, then either the antecedent conditions or the outcome 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."

Consider a simple example: a researcher interested in terrorism examines causally relevant biographical details shared by a set of individuals who committed terrorist acts. The basic insight of AI is that the goal of understanding how outcomes happen dictates a focus on the causally relevant commonalities shared by positive instances. Studying negative cases (e.g., non-terrorists) provides little or no insight regarding how things come about (e.g., acts of terrorism). Suppose, in this

example, that the researcher assesses religious radicalization as a causally relevant commonality. Classic AI emphasizes the identification and evaluation of antecedent conditions that are uniform (i.e., invariant) across instances of an outcome. For example, if the researcher was able to identify terrorists who failed to display religious radicalization, then she would take this refutation of religious radicalization as a serious challenge to its importance as a causally relevant antecedent, and not simply treat the disconfirming cases as flukes worthy of demotion to the error vector. This focus on invariant connections provides researchers the motivation to progressively revise or refine their working hypotheses, as they simultaneously deepen their knowledge of their cases.

Suppose further that the disconfirming cases (terrorists who did not display the antecedent condition, religious radicalization) nevertheless experienced a process of secular ideological radicalization. The researcher might decide to realign her working hypothesis to accommodate the new evidence, positing “religious *or* secular ideological radicalization” as a shared antecedent condition. Enlarging the scope of antecedent conditions is one of the key reformulation strategies addressed in this chapter.

As an alternative to reformulating explanatory concepts, a researcher might choose instead to accommodate disconfirming cases by narrowing the scope of the outcome. The goal of this strategy is to exclude disconfirming cases from the analysis altogether. This exact tactic was used by Howard Becker (1953, 2015) in the AI classic *Becoming a Marijuana User*. As we saw in chapter 1, Becker found that most users traveled through a series of steps in the process of achieving the outcome—becoming a user. However, he also discovered that some users did not go through the standard set of steps (Becker 1998: 205). Further research solved the puzzle: users who went through the steps learned to use marijuana *for pleasure*, while those who did not go through the steps did not learn to get high and used marijuana simply to appear to be “cool.” By restricting his argument to those who used marijuana for pleasure, Becker was able to exclude the disconfirming cases from the purview of his working hypothesis.

ADDRESSING DISCONFIRMING CASES: A FORMALIZATION

The literature on AI emphasizes two main strategies, introduced above, for dealing with disconfirming cases. It is important to recognize, however, that there are several variants of each general strategy. The choice of strategies is based primarily on the researcher’s close inspection of the disconfirming cases and careful comparison of disconfirming cases with consistent cases.

As a backdrop for the discussion of strategies, consider table 2-1, which tabulates an outcome against a causally relevant antecedent condition. The outcome in this hypothetical example is presence/absence of mass protest against the

TABLE 2-1 Initial findings

	No severe austerity	Severe austerity
IMF protest	Cell <i>a</i> : disconfirming cases; <i>N</i> = 5	Cell <i>b</i> : consistent cases; <i>N</i> = 25
Negligible or no IMF protest	Cell <i>c</i> : alternate-outcome cases; <i>N</i> = 15	Cell <i>d</i> : alternate-outcome cases; <i>N</i> = 15

N = 60.

International Monetary Fund (IMF); the causal condition is presence/absence of the imposition of severe austerity measures, mandated by the IMF as conditions for debt restructuring. The analysis embraces sixty less developed, debtor countries.¹ AI focuses on the first row of the 2×2 table, where the outcome is present. The researcher's goal is to identify antecedent conditions that are shared by instances of the outcome, as indicated by an empty cell *a* and a well-populated cell *b*.² In this example, twenty-five of the thirty instances of the outcome share severe austerity as an antecedent condition—five cases short of perfect consistency.

Cases in cell *a* are treated as anomalies to be addressed by the investigator. These cases gain specificity in the course of the research, as the researcher resolves inconsistencies through close inspection of the evidence and systematic comparison of cell *a* cases with cell *b* cases.³ The primary focus is on strategies for emptying cell *a* of cases. There are two possible destinations for cell *a* cases: they can be moved to cell *b* or *c*. To move cell *a* cases to cell *b*, the researcher must reformulate the antecedent condition so that it is more *inclusive*. To move cell *a* cases to cell *c*, the researcher must reformulate the outcome so that it is more *restrictive*. There are two main variants of each strategy.

EXPANDING THE SCOPE OF THE ANTECEDENT CONDITION

The first variant of this strategy involves using the logical term *or* to join two (or more) related antecedent conditions, as in the terrorist radicalization example discussed above. In the context of the present example, IMF protest, assume that the researcher examined cell *a* cases and concluded that even though these cases were not subjected to severe austerity measures, there was still substantial IMF protest due to each country's heavy debt burden. The researcher combines these two conditions as shown in table 2-2, which illustrates the impact of treating "heavy debt burden" and "severe austerity measures" as substitutable antecedent conditions. The use of logical *or* to join two or more conditions entails a reconceptualization of the two conditions as a single, more abstract condition. In this example, the antecedent condition might be reformulated as "debt induced economic hardship."

TABLE 2-2 Using logical *or* to increase the scope of an antecedent condition*

	No severe austerity and no heavy debt burden	Severe austerity or heavy debt burden
IMF protest	Cell <i>a</i> : disconfirming cases; <i>N</i> = 0	Cell <i>b</i> : consistent cases; <i>N</i> = 30
Negligible or no IMF protest	Cell <i>c</i> : alternate-outcome cases; <i>N</i> = 12	Cell <i>d</i> : alternate-outcome cases; <i>N</i> = 18

*Compare with table 2-1.

TABLE 2-3 Lowering the threshold of the antecedent condition*

	Less than moderate austerity	Moderate to severe austerity
IMF protest	Cell <i>a</i> : disconfirming cases; <i>N</i> = 0	Cell <i>b</i> : consistent cases; <i>N</i> = 30
Negligible or no IMF protest	Cell <i>c</i> : alternate outcome cases; <i>N</i> = 11	Cell <i>d</i> : alternate outcome cases; <i>N</i> = 19

*Compare with table 2-1.

Comparing table 2-2 to table 2-1, the five cell *a* cases have moved to cell *b*, effectively emptying cell *a* of cases and establishing a pattern of results consistent with the goals of AI. Note that this reformulation of the antecedent condition also moves three cell *c* cases to cell *d*. Thus, from the viewpoint of a statistical assessment, there is only modest gain; however, from the viewpoint of AI, the researcher has successfully reformulated the working hypothesis and established an invariant connection.

The second variant of this strategy focuses on thresholds (table 2-3). The controlling principle is that the researcher assesses the degree to which the antecedent condition must be present for the outcome to be triggered. Assume that the researcher examined cell *a* cases and concluded that even though these cases were not subjected to severe austerity, they nevertheless experienced substantial IMF pressure in the form of moderate austerity measures. This pattern of results suggests that the initial threshold for the antecedent condition, severe austerity, was too high. Only moderate austerity was required. The researcher decides to relabel and recalibrate the antecedent condition, consistent with the discovery that moderate austerity engenders IMF protest. The five cell *a* cases relocate to cell *b*, thereby establishing a pattern of results consistent with the goals of AI. Note that several cases also shift from cell *c* to cell *d*, so once again there is only modest gain from a statistical viewpoint, but from the viewpoint of AI the reformulation is decisive—cell *a* is empty and an invariant connection has been established.

Restricting the Scope of the Outcome

The second general strategy involves narrowing the set of cases with the outcome, in an effort to relocate cell *a* cases to cell *c* in table 2-1. Suppose that, following close inspection of cell *a* cases and the comparison of cell *a* cases with cell *b* cases, the

TABLE 2-4 Qualifying the outcome, making it more restrictive*

	No severe austerity	Severe austerity
Broad-based IMF protest	Cell <i>a</i> : disconfirming cases; $N = 0$	Cell <i>b</i> : consistent cases; $N = 22$
Negligible or no broad-based IMF protest	Cell <i>c</i> : alternate-outcome cases; $N = 20$	Cell <i>d</i> : alternate-outcome cases; $N = 18$

*Compare with table 2-1.

TABLE 2-5 Raising the outcome threshold*

	No severe austerity	Severe austerity
Acute IMF protest	Cell <i>a</i> : disconfirming cases; $N = 0$	Cell <i>b</i> : consistent cases; $N = 19$
Non-acute or no IMF protest	Cell <i>c</i> : alternate-outcome cases; $N = 20$	Cell <i>d</i> : alternate-outcome cases; $N = 21$

*Compare with table 2-1.

researcher observes that in contrast to most cell *b* cases, the protest in cell *a* cases was not broad based. Instead, it was driven mostly by labor unions. This difference between cell *a* cases and most cell *b* cases provides an opportunity to reformulate the outcome in a way that excludes cell *a* cases.

Table 2-4 shows the impact of reformulating the outcome. It has been changed from “IMF protest” to “broad-based IMF protest,” and cases have been shifted to accommodate the reformulation. The five cell *a* cases now reside in cell *c*, and several cell *b* cases shift to cell *d* because they did not meet the revised outcome standard (broad-based IMF protest). Once again, from a statistical standpoint, there has been only modest gain; but, from the perspective of AI, there is now the clarity of an invariant connection between the causal condition and the outcome.

The second variant of the outcome-based strategy focuses on thresholds. In this instance, suppose that the researcher compares cell *a* cases with cell *b* cases and concludes that most cell *b* cases had widespread, violent protest against the IMF, while no cell *a* cases had such acute levels. She decides to raise the threshold for the outcome to “acute” IMF protest and reassigns cases to cells accordingly. The five disconfirming cases relocate to cell *c*, and several cell *b* cases are reassigned to cell *d*. Table 2-5 illustrates the impact of raising the outcome threshold. Consistent with the goals of AI, an invariant connection has been established, and once again, a reconciliation strategy that advances understanding from the perspective of AI registers only modest gain from a statistical viewpoint.

TABLE 2-6 Imposing a scope condition: low-income countries*

	No severe austerity	Severe austerity
IMF protest	Cell <i>a</i> : disconfirming cases; $N = 0$	Cell <i>b</i> : consistent cases; $N = 20$
Negligible or no IMF protest	Cell <i>c</i> : alternate-outcome cases; $N = 12$	Cell <i>d</i> : alternate-outcome cases; $N = 13$

* $N = 45$; compare with table 2-1.

TWO MORE STRATEGIES

While most treatments of AI emphasize reformulating the antecedent conditions or the outcome, as illustrated in tables 2-2 through 2-5, two additional strategies warrant attention: (1) stipulating a scope condition and (2) typologizing the outcome.

Stipulating a Scope Condition

The first strategy is to specify a “scope condition” that can be used as a filter to restrict the set of relevant cases (Walker and Cohen 1985; Goertz and Mahoney 2012: 205–17). For example, suppose that a researcher observes that the five cases in cell *a* of table 2-1 are all middle-income countries, while most cell *b* countries are low-income countries. The researcher speculates that the close connection between severe austerity and IMF protest may be specific to low-income countries. The researcher decides to use “low-income countries” as a scope condition and removes all middle-income countries from the analysis. The results are shown in table 2-6.

The results reveal an invariant connection between severe austerity and IMF protest, specific to low-income countries. While this pattern of results satisfies the requirements of AI, from the perspective of statistical analysis it does so at a substantial cost. Observe the number of cases in table 2-6 versus table 2-1. Stipulating a scope condition reduces the sample size from $N = 60$ to $N = 45$. Tests of statistical significance are very powerfully influenced by the number of cases—the fewer the cases, the more difficult it is to achieve significance. Thus, using a scope condition to narrow the set of relevant cases may jeopardize statistical significance. Dropping middle-income countries in this example impacts the count of cases in all four cells. Thus, from the perspective of statistical analysis, table 2-6 offers only slight gain, at best, over table 2-1, despite the fact that greater empirical clarity has been achieved via AI.

Specifying Subtypes of an Outcome

The final analytic strategy for dealing with disconfirming cases involves distinguishing subtypes of the focal outcome. This strategy is not part of the corpus of classic AI, but instead is a logical extension of basic principles of the approach, relevant to generalized AI (the focus of part II of this book). Like the third and fourth strategies discussed above, specifying subtypes involves reformulating the outcome as a way to cope with disconfirming cases (George and Bennett 2005).

As before, the focus is on cell *a*—cases that display the outcome but not the hypothesized causal condition(s). The researcher asks: Did cell *a* cases experience an outcome that differed qualitatively from the outcome experienced by cell *b* cases? If so, then the evidence may be reformulated in terms of outcome subtypes, with different causal conditions linked to different subtypes. From this viewpoint, the five cases in cell *a* of table 2-1 are not disconfirming, *per se*, but instead constitute the starting point of a separate application of AI, with an outcome that differs in kind from the outcome experienced by most cell *b* cases. Thus, the evidence in table 2-1 would prompt a reconceptualization of the outcome in terms of subtypes, and instead of simply demoting cell *a* cases to cell *c*, they would be made the focus of a separate analysis.

This strategy is similar, in some respects, to the outcome-oriented strategy depicted in table 2-4. In that example, cell *a* cases were reassigned to cell *c* because the IMF protests they exhibited were not broad based (the reformulated outcome). The same observation—that the protests exhibited by cell *a* cases were union based, while the protests exhibited by most cell *b* cases were broad based—is used in the present strategy to distinguish subtypes of IMF protest. The researcher's next step would be to remove cell *a* cases from table 2-1 (along with any cell *b* cases that were union based) and assess the antecedent conditions they share, in a completely separate application of AI. In the wake of their departures, these cases would leave behind an empty cell *a* and also a consistent antecedent condition (severe austerity) for a specific subtype of IMF protest: broad based.

It is important to note that the sixth strategy—typologizing the outcome—differs fundamentally from the practice of positing equifinality (Mackie 1965, 1980). To allow for equifinality (a core feature of qualitative comparative analysis) is to acknowledge that there may be different causal recipes for the same outcome. From the viewpoint of equifinality, cases in cell *a* of table 2-1 are simply cases that experienced a different, but causally equivalent, recipe for the outcome in question (IMF protest). It is the researcher's task to identify alternate but equivalent causal recipes. By contrast, the typologizing strategy accepts, in principle, that cases residing in cell *a* exhibit a different causal recipe, but adds the stipulation that the researcher should identify differences in the outcome that follow from differences in causal recipes. For example, union-based IMF protest might be limited to strikes and peaceful demonstrations, while broad-based IMF protest might include additional, more violent forms of protest (e.g., riots).

EXTENSIONS

The examples offered so far use simple 2×2 tables with only a single causal condition in each illustration (except for table 2-2, which joined two causal conditions using logical *or*). In practice, researchers are more likely to identify multiple antecedent conditions shared by instances of an outcome. When attempting to account for “how things happen,” it is useful to think in terms of causal

TABLE 2-7 Assessing a causal recipe

	Causal recipe not satisfied	Causal recipe satisfied*
IMF protest	Cell <i>a</i> : disconfirming cases	Cell <i>b</i> : consistent cases
Negligible or no IMF protest	Cell <i>c</i> : alternate-outcome cases	Cell <i>d</i> : alternate-outcome cases

*Causal recipe = severe austerity combined with government corruption, prior mobilization, and high inflation.

recipes—combinations of conditions joined by logical *and*—that generate the outcome in question. Thus, rather than cross-tabulating single conditions against an outcome, the researcher would instead focus on the relevant antecedent conditions shared by instances of an outcome, and assess the consistency of causal *recipes*. As before, the focus is on cells *a* and *b* of the cross-tabulation, but the column headings would indicate the presence/absence of a causal recipe.

Table 2-7 offers a simple illustration. Again, the outcome is protest against the IMF. The causal recipe has four ingredients: severe austerity combined with government corruption, prior mobilization, and high inflation. Disconfirming cases (cell *a*) are those that display the outcome but not the causal recipe; consistent cases reside in cell *b*; cases satisfying the causal recipe but not displaying the outcome (i.e., “alternate-outcome” cases) are in cell *d*; while cases residing in cell *c* are alternate-outcome cases that failed to satisfy the causal recipe in question.

Another limitation of the examples offered so far is that they rely on present/absent causal conditions and present/absent outcomes. Social scientists often deal with phenomena that vary by level or degree. For example, inflation can be precisely measured. To dichotomize it as “high” versus “not-high” in a causal recipe (as in table 2-7) may seem wasteful of useful information. Fortunately, there is a readymade solution, which is to calibrate causal conditions and outcomes that vary by level or degree as fuzzy sets (Zadeh 1965, 1972; Kosko 1993; Ragin 2000, 2006a, 2008; Ragin and Fiss 2017; Smithson 1987; Smithson and Verkuilen 2006; see also appendix B). With fuzzy sets, it is possible to evaluate the *degree* of membership of each case in each relevant set. Membership scores range from 0 to 1, with a score of 0 indicating full non-membership, a score of 1 indicating full membership, and a score of 0.5 (the crossover point) indicating maximum ambiguity in whether a case is more in or more out of the set in question. For example, a country might be assigned a membership score of 0.80 in the outcome set, IMF protest, indicating that it has strong but not quite full membership in the outcome. The calibration of fuzzy-set membership scores is heavily knowledge dependent and should be based as much as possible on external criteria, and not on inductively generated criteria such as means and standard deviations or percentiles (Ragin 2000; Ragin 2008: chaps. 4 and 5).

For illustration, consider figure 2-1, a scatterplot showing a hypothetical relation between degree of membership in a causal recipe (severe austerity combined with government corruption, prior mobilization, and high inflation) and degree of

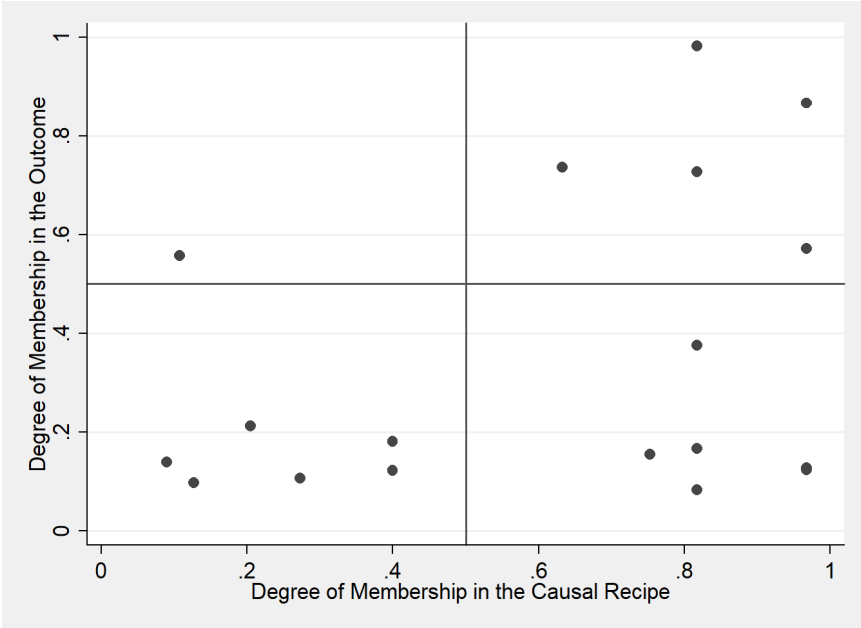


FIGURE 2-1. Illustration of the use of fuzzy-set membership scores.

membership in an outcome (IMF protest). Degree of membership in the causal recipe is calculated by first calibrating the four conditions as fuzzy sets and then selecting, for each case, the lowest of its four fuzzy membership scores, which becomes that case's degree of membership in the causal recipe. Using the lowest membership score directly implements fuzzy-set intersection, an operation that follows “weakest link” reasoning. A case can be assigned greater than 0.5 membership in a causal recipe only if it has greater than 0.5 membership in each component of the recipe.

The plot in figure 2-1 is divided into four quadrants using the two crossover points (scores of 0.5 on the causal recipe and on the outcome). The central focus of AI is the top half of the plot—cases that are more in than out of the set of cases with the outcome. Cases residing in the top-right quadrant are consistent cases—they share greater than 0.5 membership in the outcome and the causal recipe. Cases residing in the top-left quadrant are more in than out of the outcome set, but do not exhibit strong membership in the causal recipe. Thus, cases in this quadrant are disconfirming cases. It is the researcher's goal to reconcile these cases using the strategies described in this chapter. Cases residing in the lower-right quadrant share membership in the causal recipe but not in the outcome and are treated as “alternate-outcome” cases deserving of separate analytic attention (i.e., an assessment of “what happened instead”). Likewise, cases residing in the bottom-left quadrant are also alternate-outcome cases, and not directly relevant to AI.

DISCUSSION

The strategies for reconciling disconfirming cases described in this chapter involve close inspection of disconfirming cases and careful comparison of disconfirming cases with consistent cases. The choice of which reconciliation strategy to use is based fundamentally on the knowledge gained through case-oriented investigation. Because all strategies focus exclusively on emptying cell *a* of cases, they differ substantially from strategies rooted in conventional statistical methods. From the viewpoint of conventional statistical methods, the relationship observed in table 2-1 between severe austerity and IMF protest is simply probabilistic. There may be additional independent variables that could be added to the analysis that would increase the accuracy of prediction, but there is no specific focus on any one cell, nor is there any singular interest in establishing an invariant connection.

The strategies described in this chapter do not directly address the distribution of cases in the second row of the 2×2 table—where the outcome is absent. AI seeks to establish invariant connections between antecedent conditions and the presence of an outcome, paying relatively little parallel attention to the absence of the focal outcome. As explained in chapter 4, investigating the second row of table 2-1, especially cases in cell *d*, requires specification of “what happened instead?”—which could involve several alternate outcomes. Through the lens of AI, instances of the “absence” of an outcome are not “negative” cases; rather, they are positive cases of one or more alternate outcomes and are deserving of separate analytic treatment.

Finally, it is important to reiterate that the reconciliation strategies described in this chapter should be implemented in a completely transparent manner. These strategies entail close inspection of cell *a* cases and careful comparison of these cases with cases in cell *b*. The researcher has gained important insights from comparative case analysis, and she should indicate what she has learned and how she learned it.